Operations Research Letters 65 (2026) 107409

Contents lists available at ScienceDirect

Operations Research Letters

journal homepage: www.elsevier.com/locate/orl

Certified inventory control of critical resources ⋆

Ludvig Hult *, Dave Zachariah, Petre Stoica

Department of Information Technology, Uppsala University, Box 337, 751 05, Uppsala, Sweden

ARTICLE INFO

Keywords:
Inventory control
Forecasting
Policy cost inference
Policy certification

ABSTRACT

Inventory control is subject to service-level requirements, in which sufficient stock levels must be maintained despite unknown demand. We propose a data-driven order policy that certifies any prescribed service level under minimal assumptions on the demand process. The policy achieves this by adding an adjustment to any base policy. We further propose a method for forecasting the policy's operational costs that is valid in finite samples. Properties and guarantees of the method are illustrated using both synthetic and real-world data.

1. Introduction

Inventory control using discrete-time models is a well-studied problem, where orders of items to hold in stock must anticipate future demand [1,2]. By defining the costs of insufficient stocks, it is possible to find cost-minimizing policies using dynamic programming [3–5]. In practice, however, it is often preferred to encode this as a constraint on service levels rather than as penalty costs [6]. Under certain restrictive assumptions on the demand process - such as memoryless and identically distributed demand - there are explicit formulations of the duality between service levels and costs [7]. Efforts to relax such assumptions can be found in [8,9]. When the unknown demand distribution is learned from data, it is possible to provide probabilistic guarantees on the service level of an order policy in the special case of no stock being held between time periods [10,11].

In this letter, we formulate an inventory control problem relaxing most assumptions on the demand process and allowing for arbitrary time dependence of this process, while providing order policies with certifiable service-level guarantees. The relaxation is relevant in several critical inventory control problems where service levels are important, such as hospital inventory control [12,13].

We consider the management of a single item type with stock level X_t at time instant t by determining an *order* of Q_t additional units. Between t and $t+1$, items in stock are consumed by an unknown demand process D_t. A *critical stock event* occurs if

$$X_t \leq x_c \tag{1}$$

for a fixed safety stock x_c. (If $x_c = 0$, this is also known as a 'stockout' event.) When (1) is not satisfied, the normal stock level is insufficient to cover demand. In this letter, we will describe an *order policy* that guarantees a certain service level by bounding the number of critical stock events over a specified period. Maintaining the policy from time t to $t+H$ will be associated with a operating cost denoted C_t^H. Forecasts of C_t^H may inform the order policy itself or budgeting decisions to ensure the maintenance of the policy. We develop a method for constructing prediction intervals C^β that will cover C_t^H at a specified coverage level $1-\beta$. Fig. 1 provides an illustrative example of the joint order policy and cost inference method proposed in this letter.

Notation. We employ the following notation: $a \wedge b = \min\{a,b\}$, $a \vee b = \max\{a,b\}$, $(x)^+ = \max\{0,x\}$. The indicator function $\mathbb{1}(A)$ takes the value 1 if A is true, 0 otherwise.

2. Problem formulation

The stock of an inventory system follows a discrete-time dynamical equation

$$X_{t+1} = \left(X_t + Q_t - D_t\right)^+, \tag{2}$$

where order $Q_t \geq 0$ is placed at time instant t and a demand $D_t \in [0, D^{\max})$ is subsequently observed. The initial stock is $X_0 \geq 0$. For ease of exposition and without loss of generality, we will work with $x_c = 0$, so that critical stock events are indicated by $\mathbb{1}(X_t = 0)$. The generalization to $x_c > 0$ is straight forward.

At any t, we have access to historical stock data, demand data, and side information Z_t, collected in the variable $\mathcal{Z}_t = \{X_{0:t}, D_{0:t-1}, Z_{0:t}\}$. For any order policy

$$Q_t = \mu(\mathcal{Z}_t),$$

we define its *service level* as the fraction of non-critical stock events over a period $t = 1, 2, \ldots, T$. Our primary problem is to construct a policy $\mu^\alpha(\cdot)$

⋆ This work was supported in part by the Swedish Research Council under contract 2021-05022.

* Corresponding author.
E-mail addresses: ludvig.hult@it.uu.se (L. Hult), dave.zachariah@it.uu.se (D. Zachariah), ps@it.uu.se (P. Stoica).

https://doi.org/10.1016/j.orl.2026.107409
Received 24 May 2024; Received in revised form 25 June 2025; Accepted 5 January 2026
Available online 22 January 2026
0167-6377/© 2026 The Author(s). Published by Elsevier B.V. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

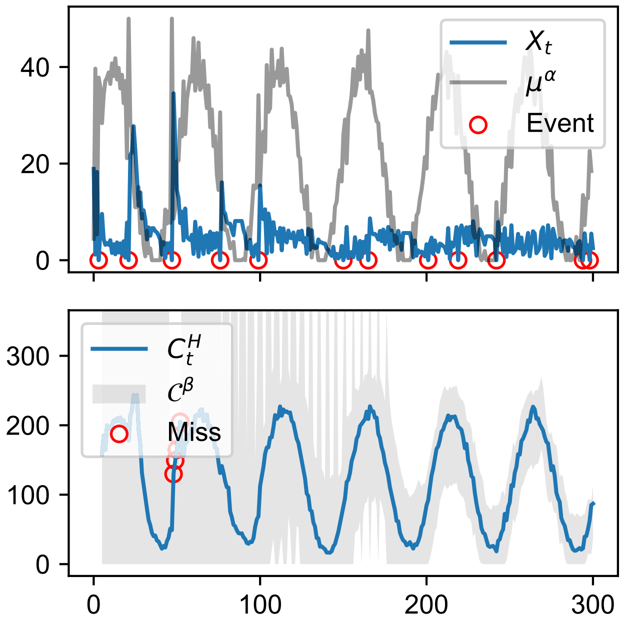

Fig. 1. Inventory control under an unknown (periodic) demand process. Top: Stock level and orders by policy μ^α that is certified to have a service level of at least $1-\alpha = 95\%$. The empirical service level is greater than 96%. Bottom: Policy operating costs C_t^H for future $H = 5$ time steps and a forcasting prediction interval C^β with a certified coverage level of $1-\beta = 95\%$. The empirical coverage level above 98%. For a full experimental description, see Section 4.1.1.

that ensures a service level of at least $1-\alpha$, that is

$$\boxed{\frac{1}{T}\sum_{t=1}^{T} \mathbb{1}\left(X_t > 0\right) \;\geq\; 1-\alpha} \tag{3}$$

with a user-defined rate $\alpha \in [0,1]$, assuming only that D_t is a bounded sequence. All policies that ensure (3) are said to be *certified* with a service level $1-\alpha$. The trivial policy $\mu^\alpha(\mathcal{Z}_t) \equiv D^{\max} - X_t$ is certified but may hold a wastefully large stock beyond the specified service level. To pursue certified policies with low operating costs requires a specification of their operating costs.

Let C_t denote the operating cost of policy at time t. For instance, $C_t = hX_{t+1} + cQ_t + K\mathbb{1}(Q_t > 0)$ is a cost that is linear in the held stock, with both linear and fixed ordering costs. The operating cost over time horizon H starting at t is then

$$C_t^H = \sum_{\tau=t}^{t+H-1} C_\tau. \tag{4}$$

The choice of cost function is application dependent. The current period cost C_t is available in next period side information Z_{t+1}. Operating costs depends critically on the demand process D_t, which is unknown. We therefore seek to construct prediction intervals $C^\beta(\mathcal{Z}_t)$ that provide an online forcast of the future operating costs C_t^H with a coverage level $1-\beta$. That is,

$$\boxed{\frac{1}{T-H+1}\sum_{t=0}^{T-H+1} \mathbb{1}\left(C_t^H \in C^\beta(\mathcal{Z}_t)\right) \geq 1-\beta} \tag{5}$$

where $\beta \in [0,1]$ is a user-defined rate.

3. Method

The method takes as a starting point any order policy $\mu^\star(\cdot)$, that we call *base policy*.

Example 1. Consider a demand predictor on the form

$$\widehat{D}(\mathcal{Z}_t) = \phi^\top(\mathcal{Z}_t)\theta_t, \tag{6}$$

where $\phi(\mathcal{Z}_t) = [1\ D_{t-1}\ \cdots\ D_{t-d_D}\ X_t\ \cdots\ X_{t-d_X}]^\top$. This is an autoregressive exogenous input (ARX) model. Then the data-driven policy

$$\mu^\star(\mathcal{Z}_t) = \left(\widehat{D}(\mathcal{Z}_t) - X_t\right)^+ \tag{7}$$

can serve as a base policy. If the demand predictor is accurate, (7) may achieve low operating costs but is does not have any certified service level under a general demand process.

Example 2. A commonly used order policy is the periodic review (s, S)-policy:

$$\mu^\star(\mathcal{Z}_t) = \begin{cases} S - X_t & \text{if } X_t < s \\ 0 & \text{else} \end{cases}, \tag{8}$$

which orders up to level S when the stock falls below level s. This is a stationary base policy in that it does not account for any information except the current stock. For certain cost functions, this policy minimizes the expected cost under the appropriate choice of (s, S) which depends on the unknown demand process [14]. It does not, however, certify any given service level.

We will introduce a form of correction to any base policy $\mu^\star(\cdot)$ such that it becomes *certified* with a specified service level $1-\alpha$. The choice of $\mu^\star(\cdot)$ will therefore only influence the cost, but not the service level guarantee (3). After deriving such a certified order policy, we turn to the problem of forecasting its operation cost C_t^H. Both the control and inference problems will be tackled by introducing nonlinear gain functions that bound an error process.

The number of 'errors' made (critical stock events or non-covering prediction intervals, repectively) will relate to the quality of the models. When there is little information available about demand, the error rate will be high, and any error corrector should account for this. We consider this to be a problem of 'burn-in time'; a certain amount of time is needed to learn the demand process sufficiently well to rely on the models.

Our proof technique is inspired by [15], which develops a one-step-ahead method for time-series prediction. However, the cited paper does not consider interventions and their feedback effects, nor the problem of providing multi-step-ahead prediction guarantees for future operating costs of a dynamic ordering policy.

3.1. Error bound and gain functions

Let the integer E_t denote an *error process* such that $E_0 = 0$ and $E_{t+1} \in \{E_t, E_t + 1\}$. For instance, it may count the number of critical stock events: $E_t = \sum_{t=1}^{T} \mathbb{1}(X_t = 0)$. We seek a function that bounds the error process.

Definition 1 (Error bound function). Any nondecreasing function $b(t)$ that satisfies $0 \le b(t) \le \alpha T$.

We will design certified policies that respect any specified error bound function $b(t)$.

Example 3. A piecewise linear error bound function is

$$b(t) = \begin{cases} 0 & \text{if } t \le T_* \\ b_* + (\alpha T - b_*)\frac{(t-T_*)^+}{(T-T_*)} & \text{else} \end{cases} \tag{9}$$

with parameters $0 \le b_* \le \alpha T$ and $0 \le T_* < T$. The parameter T_* specifies an initial duration under which we tolerate no errors. If the base policy requires some burn-in time, we can set T_* so that the certified policy compensates for the poor initial performance of the base policy. After the burn-in time, (9) tolerates a linear increase in errors from b_* to αT. Fig. 2 illustrates how this error bound function relates to an error process in a concrete example.

To ensure that an error process E_t is bounded by $b(t)$, we consider an associated gain function.

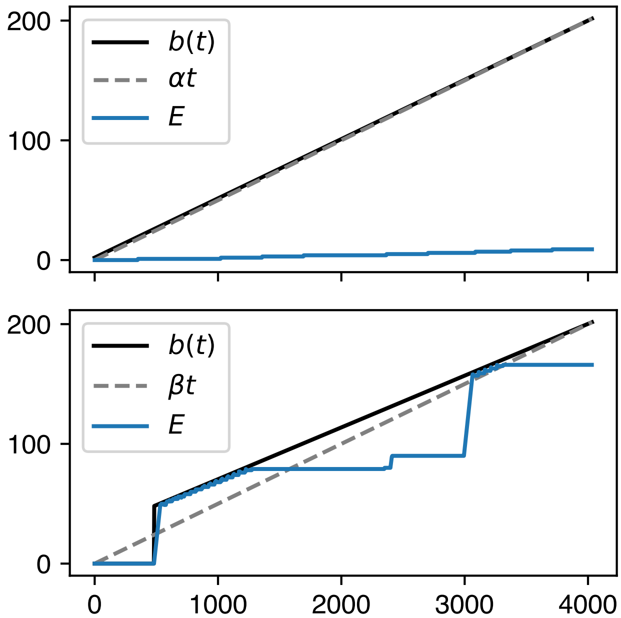

Fig. 2. Error processes E_t for experiments on the Elec2 dataset. The error processes E_t is bounded by $b(t)$ in (9) with parameters T_*, b_*. Top: E_t is the number of critical stock events. The values $b_* = 2\,T_* = 0$ are used since historical demand data was used to train the base policy, and no further burn in was required. Bottom: E_t is the number of observed plus potential miscoverage events in the cost inference, The parameters $b_* = H$, $T_* = 480$ were used since no historical costs data is available, and the nominal prediction interval model requires burn in time. See Section 4.2 for more experimental details.

Definition 2 (Associated gain). An error function $b(t)$ has an associated gain function $g_t(E)$ if it satisfies

$$E_t + 1 \geq b(t) \;\Rightarrow\; g_t(E_t) \geq g_\circ, \tag{10}$$

where $g_\circ \in \mathbb{R} \cup \{\infty\}$ is a saturation level.

Lemma 1. *Consider any error bound function $b(t)$ with an associated gain $g_t(E)$. If a saturated gain, $g_t(E_t) \geq g_\circ$, implies no error growth, i.e., $E_{t+1} = E_t$, then the error process is bounded as*

$$E_t \leq b(t) \leq \alpha T. \tag{11}$$

Proof. The proof is by induction. Base case $E_0 = 0 \leq b(0)$ holds by nonnegativity of the error bound function. Consider the inductive assumption $E_{t-1} \leq b(t-1)$. Two cases arise:

Case i) $E_{t-1} + 1 \geq b(t-1)$. Then $g_{t-1}\left(E_{t-1}\right) \geq g_\circ$ and $E_t = E_{t-1} \leq b(t-1) \leq b(t)$.

Case ii) $E_{t-1} + 1 < b(t-1)$. Then $E_t \leq E_{t-1} + 1 < b(t-1) \leq b(t)$.

Finally, $E_t \leq \alpha T$ follows by the definition of E_t. □

Lemma 1 justifies the term ‘error bound function’ for $b(t)$.

We will now turn to formulating a nonlinear gain function that adds error correction to the base policy so as to certify its service level. The methodology will then be extended to infer the future operating costs.

3.2. Certified policy

We present how to construct a *certified policy* with service level $1 - \alpha$ starting from any given base policy, by taking the number of critical stock events at t to be the error process E_t and designing a policy so that $E_T \leq \alpha T$.

Theorem 1. *Given any error bound function with associated gain g_t with saturation level $g_\circ \geq D^{\max}$, for any base policy $\mu^\star(\cdot)$, the order policy*

$$\mu^\alpha(\mathcal{Z}_t) = \left(\mu^\star(\mathcal{Z}_t) + g_t\left(\sum_{\tau=1}^{t} \mathbb{1}(X_\tau = 0)\right)\right)^+ \tag{12}$$

is certified with service level $1 - \alpha$. That is, it satisfies (3).

Proof. The cumulative number of critical stock events $E_t = \sum_{\tau=1}^{t} \mathbb{1}\left(X_\tau = 0\right)$ is an error process. Using (2) and (3), we can express the error process as:

$$E_{t+1} - E_t = \mathbb{1}\Big(\left(\mu^\star(\mathcal{Z}_t) + g_t\left(E_t\right)\right)^+ \leq D_t - X_t\Big) \tag{13}$$

A sufficient condition for $E_{t+1} = E_t$ is therefore $g_t\left(E_t\right) > D_t - X_t - \mu^\star\left(\mathcal{Z}_t\right)$. Since the base policy and the stock must be nonnegative, it is sufficient that $g_t\left(E_t\right) > D^{\max}$ to know that $E_{t+1} = E_t$. Since $g_\circ \geq D^{\max}$, we have established that $g_t(E_t) \geq g_\circ$ implies that $E_{t+1} = E_t$. Lemma 1 provides the conclusion that $E_T \leq \alpha T$, thus completing the proof. □

Remark 1. The certified policy of Theorem 1 can be replaced by $\mu^\alpha\left(\mathcal{Z}_t\right) \vee (D^{\max} - X_t)$ without affecting its certification. This may provide a policy with lower cost in cases where the saturation level of the gain function is large.

Next, we propose two gain functions for use in (12) that yield low costs in simulations. Verification that they are associated gain functions according (10) is direct and thus omitted.

Corollary 1. *The nonlinear gain*

$$g_t(E) = \begin{cases} \tan\left(\frac{\pi}{2}\,\frac{E+1}{b(t)}\right) & \text{if } E + 1 \in [0, b(t)] \\ \infty & \text{else} \end{cases} \tag{14}$$

can be used to provide a certified policy. It is the associated gain to any error bound function $b(t)$, and has saturation level $g_\circ = \infty$. Specifically, one may use the error bound of (9) with $T_* = 0$ and $b_* = 2$.

Using (14) yields low gain when there is a large tolerance for critical stock events ($E + 1 \ll b(t)$). On the other hand, the gain is increased indefinitely as the error process grows closer to the error bound.

Other nonlinear gain functions may provide lower costs than (14). For instance, in the presence of fixed ordering costs the nonlinear gain should allow for zero sized orders to not incur the fixed cost.

Corollary 2. *The piecewise linear gain*

$$g_t(E) = \begin{cases} 0 & E + 1 \in [0, b(t)/2] \\ D^{max}\left(2\frac{E+1}{b(t)} - 1\right) & E + 1 \in [b(t)/2, b(t)]\,. \\ D^{max} & else \end{cases} \tag{15}$$

can be used to provide a certified policy. It is the associated gain to any error bound function $b(t)$, and has saturation level $g_\circ = D^{\max}$. Specifically, one may use the error bound of (9) with $T_* = 0$ and $b_* = 2$.

3.3. Valid cost inference

We now turn to forecasting the operating costs (4) of a certified policy $\mu^\alpha(\cdot)$. Similarly to the policy construction that starts with a base policy, the valid cost inference method starts with any nominal prediction interval $\left[\hat{C}^\beta(\mathcal{Z}_t), \check{C}^\beta(\mathcal{Z}_t)\right]$.

Example 4. Nominal intervals produced by a quantile estimator. Let $\bar{C}(\mathcal{Z}_\tau)$ be a simple point predictor of the cost and define its empirical error as $\epsilon_\tau = C^H_\tau - \bar{C}(\mathcal{Z}_\tau)$. Then a quantile estimator is given by:

$$\bar{C}_a(\mathcal{Z}_t) = \bar{C}(\mathcal{Z}_t) + \inf\left\{p : \frac{1}{t - H + 1}\sum_{\tau=0}^{t-H} \mathbb{1}\left(\epsilon_\tau \leq p\right) \geq a\right\}. \tag{16}$$

and using it we can define a nominal prediction interval as

$$\left[\hat{C}^\beta\left(\mathcal{Z}_t\right), \check{C}^\beta\left(\mathcal{Z}_t\right)\right] = \left[\bar{C}_{\frac{\beta}{2}}(\mathcal{Z}_t), \bar{C}_{1-\frac{\beta}{2}}(\mathcal{Z}_t)\right]. \tag{17}$$

The nominal interval will be adjusted by introducing a gain as we show next.

Theorem 2. *Assume non-negative cost over a finite horizon, $C_t^H \in [0, \widetilde{C}]$, where the upper bound may be infinity. Consider a horizon $H \geq 2$ and any nominal prediction interval $[\hat{C}^\beta(\mathcal{Z}_t), , \check{C}^\beta(\mathcal{Z}_t)]$. Define an adjusted prediction interval*

$$C^\beta(\mathcal{Z}_t) = \left[\left(\hat{C}^\beta(\mathcal{Z}_t), -q_t\right)^+, \left(\check{C}^\beta(\mathcal{Z}_t) + q_t\right) \wedge \widetilde{C} \right], \tag{18}$$

where

$$q_t = g_t\left(\sum_{\tau=t-H+1}^{t-1} \mathbb{1}\left(C^\beta(\mathcal{Z}_\tau) \neq [0, \widetilde{C}] \right) + \sum_{\tau=0}^{t-H} \mathbb{1}\left(C_\tau^H \notin C^\beta(\mathcal{Z}_\tau) \right) \right) \tag{19}$$

and

$$g_t(E) = \begin{cases} \tan\left(\frac{\pi}{2}\left(2\frac{E+1}{b(t)} - 1\right) \right) & \text{if } 0 < b(t) \text{ and } E + 1 \leq b(t) \\ \infty & \text{else} \end{cases} \tag{20}$$

is a gain associated with the error bound function $b(t)$ in (9) using β in lieu of α. Then $C^\beta(\mathcal{Z}_t)$ has a coverage level $1 - \beta$ according to (5).

Remark 2. A similar result holds for the special case of $H = 1$.

Remark 3. For $b(t)$ we use $b_* = H$ as a default value and the burn-in period T_* is problem dependent.

Proof. Define

$$C^\beta(\mathcal{Z}_t) = \{ v \in [0, \widetilde{C}] : \left(v - \check{C}^\beta(\mathcal{Z}_t)\right)^+ + \left(\hat{C}^\beta(\mathcal{Z}_t) - v\right)^+ \leq q_t \} \tag{21}$$

Let the number of observed miscoverage events at time t be define as $\widetilde{E}_t = \sum_{\tau=0}^{t-H} \mathbb{1}(C_\tau^H \notin C^\beta(\mathcal{Z}_\tau))$ that we want to bound by $\beta\widetilde{T}$ where $\widetilde{T} = T - H + 1$. Define the error process to be the number of known as well as potential miscoverage events:

$$E_t = \widetilde{E}_t + \sum_{\tau=t+H+1}^{t-1} \mathbb{1}\left(C^\beta(\mathcal{Z}_\tau) \neq [0, \widetilde{C}] \right) \tag{22}$$

so that $q_t = g_t(E_t)$. Let $g_\circ = \widetilde{C}$ be the saturation level. Then a direct verification shows that $g_t(E_t) \geq g_\circ \Rightarrow E_{t+1} = E_t$. We can therefore use Lemma 1 to show that $E_{\widetilde{T}} \leq \beta\widetilde{T}$. Since $\widetilde{E}_t \leq E_t$, we have also proven that

$$\widetilde{E}_{\widetilde{T}} \leq E_{\widetilde{T}} \leq b(\widetilde{T}) \leq \beta\widetilde{T} \tag{23}$$

and therefore that $C^\beta(\mathcal{Z}_t)$ has a coverage level $1 - \beta$. □

4. Numerical experiments

The proposed method of an obtaining certified order policy and the forecasting for its future operational costs, are illustrated in a series of simulations using synthetic and real data. In all cases, we fix the prescribed service levels to $1 - \alpha = 95\%$ and the coverage level to $1 - \beta = 95\%$.

We explore three different settings: 1) synthetic data with continuous demand, no fixed ordering costs, 2) real world data but otherwise retaining the cost function of the synthetic cases, and 3) a synthetic data example where there are fixed ordering costs.

4.1. Synthetic data

The operating cost is taken to be a cost of purchase plus that of holding stock, i.e., $C_t = Q_t + hX_t$. For simplicity, we use $h = 1$. We use Corollary 1 and Remark 1 to transform the base policy in (7) into a certified policy (12).

The parameters θ_t of the demand predictor (6) are tracked by recursive least squares (RLS) [16, eq. 9.12]. To reduce the burn-in time, the RLS parameters are initialized by tracking on historical data $D_{-B:0}$, $Q_{-B:0}$, $X_{-B:0}$ recorded under an alternative policy, namely the empirical $(1 - \alpha)$-quantile of demand.[1] We use model orders $d_X = d_D = 2$ in (6), and set the RLS forgetting factor to $\lambda = 0.99$.

A direct verification shows the finite horizon operating cost is upper bounded by $C_t^H = HD^{\max}(1 + h)$. To construct the nominal prediction interval of the cost, we use (17) with another linear-in-parameters autoregressive model with parameters tracked by RLS,

$$\bar{C}(\mathcal{Z}_t) = \varphi^\top(\mathcal{Z}_t)\vartheta_t.$$

We use the feature vector of a linear AR-5 model for H steps ahead prediction $\varphi(\mathcal{Z}_t) = \begin{bmatrix} 1 & C_{t-H-4}^H & C_{t-H-3}^H & \dots & C_{t-H}^H \end{bmatrix}^\top$. This model predicts the future costs of the specific policy, and there are no data on this policy for $t \leq 0$. Therefore, no pre-training is done. The parameters are initialized with $\hat{\vartheta}_0 = [C/2, 0 \dots, 0]$.

We explore the order policy and cost inference methods on three rather different demand processes. All simulations on synthetic data use $T = 300$, $H = 10$ and $B = 150$, and the demand is upper bounded by $D^{\max} = 50$.

4.1.1. Periodic demand

The unknown demand process is assumed to follow

$$D_t = 20 + 20\sin(2\pi t/50) + e_t \tag{24}$$

where $e_t \sim \mathcal{N}(0, 1)$ and the demand is clipped to the interval $[0, 50]$. This captures seasonality with random shocks. In the cost inference, we use the RLS forgetting factor $\lambda = 0.99$ and set a burn-in time of $T_* = 40$ samples. Fig. 1 shows the periodic orders resulting from the policy. We also see that the cost inference starts to become informative around $t = 150$.

4.1.2. Spiking demand

The unknown demand for items assumed to be proportional to the size of a latent infected population I_t,

$$D_t = 50I_t,$$

which models episodic demand that arises from infectious decreases. Specifically, I_t follows a stochastic 'Susceptible-Infected-Removed' or SIR-type model:

$$\begin{aligned} S_t &= S'_{t-1} - 0.5S'_{t-1}I'_{t-1} \\ I_t &= I'_{t-1} + 0.5S'_{t-1}I'_{t-1} - 0.2I'_{t-1} \\ R_t &= (1 - e_t)R_{t-1} + 0.2(I_{t-1} + 0.001e_t) \end{aligned} \tag{25}$$

In (25), we have that

$$\begin{aligned} S'_{t-1} &= S_{t-1} + (R_{t-1} - 0.001)e_t \\ I'_{t-1} &= I_{t-1} + 0.001e_t \end{aligned}$$

and $e_t \sim \mathrm{Ber}(0.03)$. In each time step, with 3% probability, the population loses immunity and 0.1% of the population gets infected.

In the cost inference, we use the RLS forgetting factor $\lambda = 0.995$ and set a burn-in time of $T_* = 50$ samples. Fig. 3 shows a spiking order pattern resulting from the policy. Note how the orders drop to zero during a longer period only to shoot up at the end.

4.1.3. Feedback demand

The unknown demand is assumed to depend on the stock level and therefore introduces feedback into system. Specifically, the demand process is

$$D_{t+1} = (5 + X_t + e_t) \vee 49.999, \tag{26}$$

where $e_t \sim \chi_1^2$. In the cost inference, we use $\lambda = 0.95$ and set the burn-in period to $T_* = 30$ samples. Fig. 4 shows a considerable tightening of the cost inference after $t = 150$.

[1] More specifically, $\mu^\circ(\mathcal{Z}_t) = \inf\left\{ p : \sum_{\tau=-B}^{t-1} \mathbb{1}(D_\tau \leq p) \geq (1 - \alpha)(B - t + 1) \right\}$.

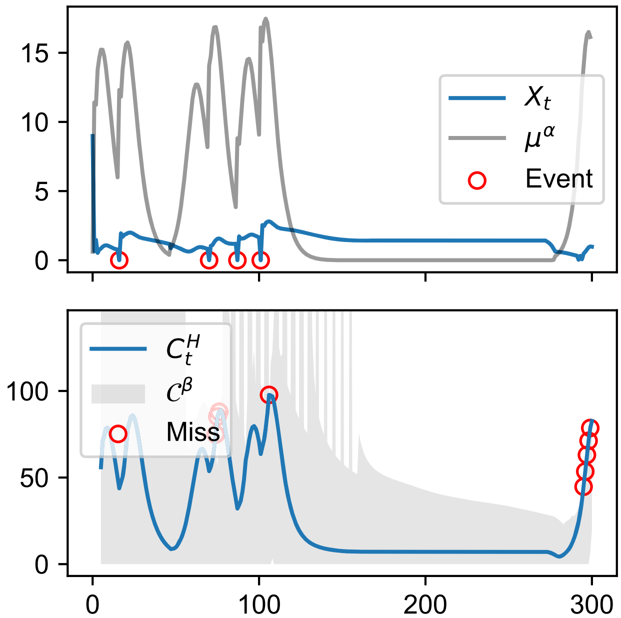

Fig. 3. Stock, purchase and cost estimate for the SIR model. The empirical service level is 98.7% exceeding the prescribed $95\% = 1 - \alpha$. The cost prediction interval covers the true cost in 97% of cases, exceeding the prescribed $95\% = 1 - \beta$.

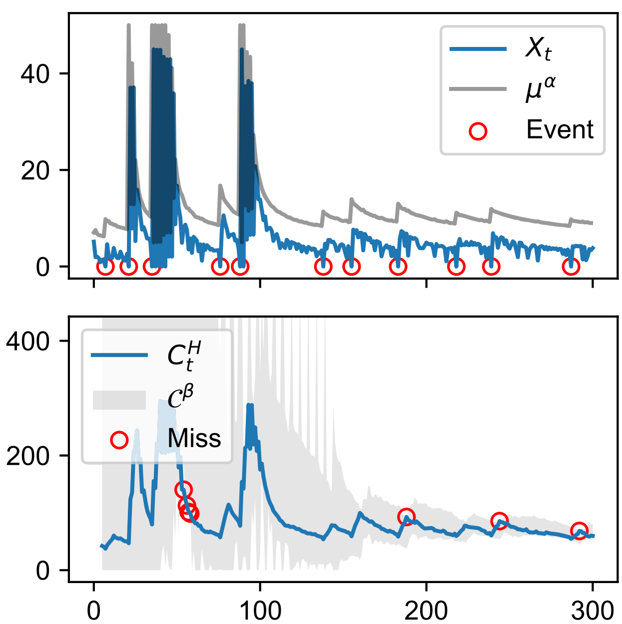

Fig. 4. Stock, purchase and cost estimate for the feedback demand model. The empirical service level is 96.3% exceeding the prescribed $95\% = 1 - \alpha$. The cost prediction interval covers the true cost in 97.7% of cases, exceeding the prescribed $95\% = 1 - \beta$.

4.2. Electricity dataset

The results on synthetic data validate the guarantees of the order policy and cost inference methodology for three rather different demand processes. As an example of a real-world demand process, we apply the same method to a problem driven by the electricity demand variable D_t (NSWDemand) in the Elec2 dataset [17]. This process represents seasonality and shocks as well as other distribution shifts. It is therefore both interesting and challenging to analyze. The full dataset records the demand every 30 min (48 samples per day) for ca 900 days, normalized so that $D^{\max} = 1$. We set the cost horizon to $H = 1 \times 48$ corresponding

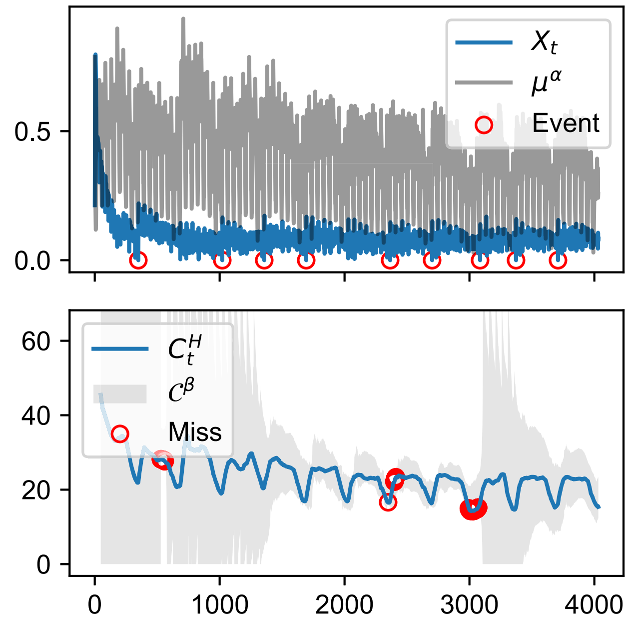

Fig. 5. Stock, purchase and cost estimates using the demand recorded in the Elec2 dataset (and is normalized to [0,1].) The empirical service level is 99.8% exceeding the prescribed $95\% = 1 - \alpha$. The cost prediction interval covers the true cost in 97% of cases, exceeding the prescribed $95\% = 1 - \beta$.

to 1 day. The pretraining window length to 3 days ($B = 3 \times 48$) and algorithm is run over 12 weeks of time periods ($T = 12 \times 7 \times 48$). The first $B + T$ datapoints are used for selecting suitable hyperparameters. The algorithm is then run on the subsequent $B + T$ samples.

The base policy, the nominal cost prediction interval, the error bounds functions and the nonlinear gain functions were as in the synthetic data experiments, with the following changes. We chose demand model orders $d_X = 0$, $d_D = 48$ in (6), and set the RLS with forgetting factor to $\lambda = 0.99$. The cost inference feature map φ was set to include an autoregressive model of order 24 with additional Fourier coefficients representing periods of 3, 6, 12, 24 h, and 7 days. The burn-in time was $T_* = 10 \times 48$ and the forgetting factor was $\lambda = 0.995$.

Fig. 5 shows strong periodic variations as well as sudden spikes in the orders. Despite significant variability in demand, the order policy ensures the prescribed service level $1 - \alpha$. We also note that the cost inference tightens as in the previous case, but around $t = 3\,000$ demand fluctuations lead to several miscoverage events. These inflate the prediction intervals, which subsequently shrink again.

We use this example to illustrate in Fig. 2 the error processes E_t associated with the order policy and cost inference, respectively. These processes are bounded by $b(t)$, which determines the associated gains.

4.3. Fixed ordering costs

This numerical experiment showcases how the certified policies in (12) work equally well with other demand processes and base policies. Specifically, consider the cost function $C_t = K\mathbb{1}(Q_t > 0) + cQ_t + h(X_t + Q_t - D_t)^+ + p(D_t - Q_t - X_t)^+$, i.e., fixed and linear ordering costs plus a linear holding cost and penalty costs for shortages. The cost parameters of the problem were chosen to be $K = 64$, $h = 5$, $p = 32$. The unit purchase cost $c = 0$ for simplicity. The demand is drawn from an exponential distribution

$$D_t \sim \exp(1/10)$$

and truncated to a maximal demand $D^{\max} = 100$.

It is known that under an i.i.d. demand process and an infinite time horizon $T \to \infty$, there exists an (s, S)-policy (8) that minimizes the expected cost. We therefore use (8) as a base policy, and set its parameters

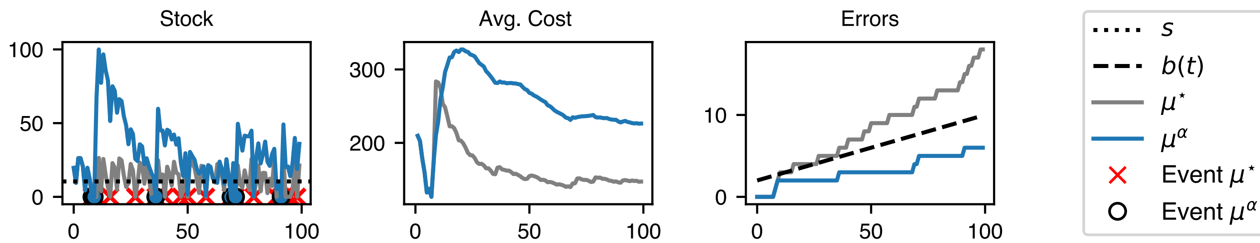

Fig. 6. Simulation of base (s, S)-policy (8) and certified policy, using Corollary 2 for 100 time steps. From left to right, it compares the two policies in terms of stock X_t, average cost $\frac{1}{t}\sum_{s=1}^{t} C_s$, and the number of critical stock events, which is the error process E_t. As reference, the order level s and the error bound function $b(t)$ are presented. Critical stock events are indicated in the stock plot. The certified policy has an empirical service rate of 94% (where $1-\alpha = 90\%$) compared to only 82% for the base policy.

s and S using a gradient based optimization [18,19]. This is transformed into a certified policy (12) with service level $1-\alpha = 90\%$ using Corollary 2. For an illustration of how the certified policy differs from the base policy, we run both over $T = 100$ time periods as shown in Fig. 6. The results illustrate a fundamental trade-off between cost and certified service level. For while the certified policy results in higher costs than the base policy, it by contrast satisfies the specified service level.

5. Discussion

We have considered an inventory control problem with an unknown demand process in which order policies can be constructed with certified service levels and their operating cost can be forecast with a prescribed coverage level. These guarantees hold under very weak assumptions on the unknown demand process. We have verified the methods numerically and illustrated them on synthetic and real data.

Because our focus is on servicing critical resources, we are interested in constraining policies to fulfill specified service levels deterministically. This problem formulation differs from many other inventory control problems in which critical stock events are more tolerable and instead associated with penalty costs; then the goal is to find policies that achieve some minimal expected costs no matter how well they serve.

The method presented herein starts from any given base policy and transforms it into a certified policy by a nonlinear gain function of the number of critical stock events. Similarly, the cost estimation problem is addressed by taking a nominal prediction interval and adjusting it to obtain a valid cost inference method. Several design choices are possible here. In the simulation studies we used classical linear autoregressive models, but the use of more sophisticated predictors, such neural networks, is rather straightforward.

The choice of error bound function - which control the tolerable rate of errors - and the associated gain function affect the conservativeness of the method. In our experiments, we used the error bound function (9) with a simple parametrization, where we adjust the burn-in time T_* reflecting that different models need more data to predict well. More complex error bound functions are conceivable, which depend not only on time but also on the accuracy of the nominal predictor models.

The only assumption we have made about the demand process is that it is a bounded sequence. The upper bound is required to ensure that there is a fall back ordering quantity that prevents critical stock events. If an upper bound on the demand is not available, an arbitrary large upper bound may be used, but this may lead to very large order quantities. We suggest that the upper bound is set based on available data and domain expertise.

Further exploratory work may study the impact of the choice of prediction models, as well as error bound functions and associated gains, on the operating costs and the tightness in cost inference. This may lead to finding alternative certifiable policies with lower cost, and valid cost inference methods with tighter prediction intervals.

CRediT authorship contribution statement

Ludvig Hult: Writing – original draft, Software, Investigation, Formal analysis, Conceptualization; **Dave Zachariah:** Writing – review & editing, Supervision, Funding acquisition; **Petre Stoica:** Writing – review & editing.

Data availability

No data was used for the research described in the article.

References

[1] K.J. Arrow, S. Karlin, H.E. Scarf, M.J. Beckmann, J. Gessford, R.F. Muth, Studies in the Mathematical Theory of Inventory and Production, Stanford mathematical studies in the social sciences, Stanford university press, 1958.
[2] S. Axsäter, Inventory Control, 225 of *International Series in Operations Research & Management Science*, Springer International Publishing, 2015. https://doi.org/10.1007/978-3-319-15729-0
[3] D.P. Bertsekas, Dynamic Programming and Stochastic Control, 125, Academic Press, 1976.
[4] D.P. Bertsekas, Dynamic Programming and Optimal Control, Athena Scientific, 1995.
[5] A. Bensoussan, Dynamic Programming and Inventory Control, IOS Press, 2011.
[6] F.Y. Chen, D. Krass, Inventory models with minimal service level constraints, Eur. J. Oper. Res. 134 (1) (2001) 120–140. https://doi.org/10.1016/S0377-2217(00)00243-5
[7] G.J. Van Houtum, W.H.M. Zijm, On the relationship between cost and service models for general inventory systems, Stat. Neerl. 54 (2) (2000) 127–147. https://doi.org/10.1111/1467-9574.00132
[8] C.E. Larson, L.J. Olson, S. Sharma, Optimal inventory policies when the demand distribution is not known, J. Econ. Theory 101 (1) (2001) 281–300. https://doi.org/10.1006/jeth.2000.2772
[9] X. Yan, X. Chao, Y. Lu, Optimal control policies for dynamic inventory systems with service level dependent demand, Eur. J. Oper. Res. 314 (3) (2024) 935–949. https://doi.org/10.1016/j.ejor.2023.11.005
[10] J. Huber, S. Müller, M. Fleischmann, H. Stuckenschmidt, A data-driven newsvendor problem: from data to decision, Eur. J. Oper. Res. 278 (3) (2019) 904–915. https://doi.org/10.1016/j.ejor.2019.04.043
[11] R. Levi, G. Perakis, J. Uichanco, The data-driven newsvendor problem: new bounds and insights, Oper. Res. 63 (6) (2015) 1294–1306. https://doi.org/10.1287/opre.2015.1422
[12] E. Saha, P.K. Ray, Modelling and analysis of inventory management systems in healthcare: a review and reflections, Comput. Ind. Eng. 137 (2019) 106051. https://doi.org/10.1016/j.cie.2019.106051
[13] M. Bijvank, I.F.A. Vis, Inventory control for point-of-use locations in hospitals, J. Oper. Res. Soc. 63 (4) (2012) 497–510. https://doi.org/10.1057/jors.2011.52
[14] H. Scarf, The Optimality of (S,s) Policies in the Dynamic Inventory Problem, Stanford University Press, 1960, pp. 196–202.
[15] A.N. Angelopoulos, E. Candes, R. Tibshirani, Conformal PID control for time series prediction, in: Thirty-seventh Conference on Neural Information Processing Systems, 2023. https://openreview.net/forum?id=zPYeYv6YYs.
[16] T. Söderström, P. Stoica, System Identification, Prentice-Hall, 1989.
[17] M. Harries, Splice-2 Comparative Evaluation: Electricity Pricing, techreport UNSW-CSE-TR-9905, University of New South Wales, 1999.
[18] W. Dai, J.-Q. Hu, C. Zhu, A gradient-based method to calculate (s,s) policies, Oper. Res. Lett. 51 (4) (2023) 468–473. https://doi.org/10.1016/j.orl.2023.06.008
[19] A.F. Veinott, H.M. Wagner, Computing optimal (s,s) inventory policies, Manage. Sci. 11 (5) (1965) 525–552. https://doi.org/10.1287/mnsc.11.5.525