IIETA International Information and Engineering Technology Association
Advancing the World of Information and Engineering

Journal Européen des Systèmes Automatisés
Vol. 59, No. 10, October, 2026, pp. 2979-2997
Journal homepage: http://iieta.org/journals/jesa

Adaptive Wasserstein Distributionally Robust Markov Decision Process for Dynamic Emergency Department Scheduling

Rina Widyasari[1], Saib Suwilo[2*], Mardiningsih[2], Sutarman[2]

[1] Graduate School of Mathematics, Faculty of Mathematics and Natural Science, Universitas Sumatera Utara, Medan 20155, Indonesia
[2] Department of Mathematics, Faculty of Mathematics and Natural Science, Universitas Sumatera Utara, Medan 20155, Indonesia

Corresponding Author Email: saib@usu.ac.id

Copyright: ©2026 The authors. This article is published by IIETA and is licensed under the CC BY 4.0 license (http://creativecommons.org/licenses/by/4.0/).

https://doi.org/10.18280/jesa.591018

Received: 18 July 2026
Revised: 25 September 2026
Accepted: 4 October 2026
Available online: 31 October 2026

Keywords:
adaptive robustness, automated decision-making, distribution shift, distributionally robust Markov decision process, emergency department scheduling, Wasserstein distance

ABSTRACT

Emergency department (ED) scheduling is affected by substantial uncertainty in patient arrivals, so policies optimized for nominal conditions may be sensitive to demand shifts. This study develops a data-calibrated adaptive Wasserstein distributionally robust Markov decision process (MDP) for dynamic ED patient scheduling. The model represents five triage classes and physician availability in a finite-state framework and applies Wasserstein robust value iteration to handle transition-distribution uncertainty. Unlike fixed-radius formulations, the proposed Data-Calibrated Adaptive Wasserstein Robust Value Iteration (DC-AWRVI) links the observed deviation of the arrival distribution from the nominal model to the level of robust protection. For each hour, the ambiguity radius is the 95th percentile, over state–action pairs, of the Wasserstein-1 distance between hour-specific and nominal transition kernels, bounded by a baseline and a maximum, and it selects a policy from an offline library of robust policies. The nominal arrival rate was 7 patients per hour, with hourly variation calibrated from NHAMCS data; radii ranged from 0.00127 under near-nominal demand to 0.01529 under high upward shifts. Numerical checks confirmed consistency of the nominal and robust Bellman solvers; robustification changed decisions only in selected states. DC-AWRVI reproduced nominal-policy performance under nominal demand but, under NHAMCS-calibrated and controlled-surge scenarios, did not achieve the lowest mean waiting cost. Based on 2,000 Monte Carlo replications per policy–scenario pair, it incurred a premium of 1.39% over Nominal Value Iteration (NVI) under the NHAMCS-calibrated profile and 1.25% under controlled surge, with equal or marginally higher queue measures and no improvement in $CVaR_{0.95}$. Across constant arrival rates of 7–12 patients per hour, the premium ranged from 0.18% to 3.61%, largest at the largest radius. The approach thus offers a data-driven mechanism for adapting distributional protection to demand shifts, but the results quantify the cost of this protection rather than a demonstrated operational improvement.

1. INTRODUCTION

Emergency departments (EDs) operate in highly dynamic environments characterized by unpredictable patient arrivals, varying levels of clinical urgency, and limited medical resources. Unlike elective healthcare services, ED demand cannot be scheduled in advance. Service priorities must therefore be continuously adjusted according to patient urgency, resource availability, and the evolving system state. Consequently, ED congestion and patient-flow performance are shaped by interacting factors, including demand variability, capacity constraints, prioritization decisions, and process design [1-3].

Dynamic scheduling approaches address these operational challenges by incorporating current system information into decision-making. Previous research formulated ED scheduling as a Markov decision process (MDP) and applied deep reinforcement learning to accommodate heterogeneous acuity levels and interdependent resources [4]. Other studies investigated dynamic patient scheduling, real-time resource allocation, coordinated scheduling, and related operational approaches [5-9]. Collectively, these contributions demonstrate that ED scheduling is inherently sequential and state-dependent.

Among these approaches, the MDP framework is particularly suitable because it explicitly connects system states, scheduling actions, stochastic transitions, and future operational consequences. Previous applications examined ED operational policies and multi-priority patient scheduling under limited diagnostic resources [10, 11]. However, the reliability of an MDP policy depends strongly on the transition model used during optimization. In practice, patient arrivals, urgency composition, service completions, and resource availability remain uncertain. The transition dynamics encountered during implementation may therefore differ from those assumed during policy development.

This potential mismatch created a need for decision models that remain reliable under transition-model uncertainty. Robust MDPs addressed this issue by considering sets of plausible transition probabilities [12, 13]. Distributionally robust MDPs extended this perspective by optimizing decisions over sets of probability distributions rather than relying on a single nominal model [14-16]. These frameworks allowed transition ambiguity to be incorporated directly into sequential decision-making.

Although distributionally robust MDPs provided a general framework, their practical behavior depended on how the ambiguity set was constructed. Wasserstein distance offered a natural basis for this construction because it measures the transportation effort required to transform one probability distribution into another. Wasserstein distance provides a principled metric for comparing probability distributions through optimal transport [17]. Data-driven Wasserstein ambiguity sets can further yield tractable reformulations and theoretical performance guarantees [18]. Recent developments extended this approach to computational methods for distributionally robust MDPs, robust Q-learning, and healthcare scheduling [19-21].

These contributions address distributional uncertainty in different ways. Wasserstein distributionally robust MDPs were formulated as a convex optimization problem [16], which provides a methodological foundation for handling transition-distribution uncertainty. Clement and Kroer [19] concentrated on computation and investigated first-order methods for solving Wasserstein distributionally robust MDPs. Neufeld and Sester [20] addressed robust decision-making through a learning-based formulation, namely robust Q-learning under Wasserstein uncertainty. Wasserstein distributionally robust optimization was applied by Wang et al. [21] to surgery scheduling with elective and emergency patients, showing that distributional uncertainty is relevant to healthcare scheduling. Taken together, these studies provide the methodological and computational basis for the present work. They differ from it in their decision problems and in how the ambiguity radius is used: none of them combines a finite-state, model-based Wasserstein distributionally robust MDP for dynamic ED patient prioritization with an ambiguity radius that is adjusted to estimated upward changes in patient demand.

The present study is positioned accordingly. The model is single-stage: each scheduling decision concerns only the selection of waiting patients for their first physician service, that is, a single service stage, rather than surgical scheduling or multistage clinical workflow optimization. The decision epochs within the operating horizon are repetitions of this same decision and not successive stages of care for one patient. The contribution does not lie in introducing Wasserstein ambiguity sets, distributionally robust Bellman optimization, or robust value iteration as new concepts. It lies in integrating demand-responsive ambiguity-radius calibration with a single-stage stochastic ED scheduling model, together with a numerical examination of its policy behavior, computational implications, and operational limitations.

Within a Wasserstein framework, the ambiguity radius determined the permitted deviation from the nominal transition distribution. A small radius kept the robust model close to its nominal counterpart. In contrast, a larger radius provided broader protection against transition-model deviations but also could produce more conservative decisions. Selecting the ambiguity radius was therefore central to balancing robustness and operational efficiency.

Data-driven Wasserstein ambiguity sets were usually centered on the empirical distribution of a training sample, and the radius was then chosen from a measure-concentration bound that depends on the sample size and a confidence level, or calibrated by bootstrapping or cross-validation [18]. This formulation was developed for single-stage problems. ED scheduling was formulated as an MDP by Lee and Lee [4], and a policy was also learned by deep Q-learning in a simulated ED with Poisson arrivals at 7 to 10 patients per hour, of which 7 was the normal rate. Their study did not model distributional uncertainty in the arrival process, and it did not link a level of protection to the observed deviation of arrivals from the normal rate. The present study centered the ambiguity set on the nominal transition kernel and derived its radius from the observed deviation of the arrival rate through the exact Wasserstein distance between the corresponding kernels. The radius therefore carries no finite-sample guarantee; its purpose is to scale the protection with the observed shift.

Patients were scheduled to doctors dynamically to minimize weighted tardiness, using a reoptimization heuristic, variable neighborhood search, and scenario sampling, and evaluated the effect of additional doctors; their study did not place an ambiguity set on the arrival process [5]. In the Wasserstein setting, Wang et al. [21] centered the ball on the empirical distribution of a training sample and used a fixed radius, with a radius of zero reducing to sample average approximation; Nominal Value Iteration (NVI) plays the corresponding role of a zero-radius reference in the present study, but the radius here changes with the observed arrival rate and the decisions are sequential.

A fixed ambiguity radius provides a uniform level of protection but does not reflect changes in operating conditions. Moreover, Wasserstein distance measured the magnitude of a distributional shift without identifying its operational direction. Both lower-than-nominal and higher-than-nominal demand might differ substantially from the reference distribution. However, their consequences for ED congestion were not equivalent. Upward demand shifts were generally more likely to intensify queueing pressure and resource competition. This distinction motivated an adaptive mechanism that responds specifically to congestion-relevant transition shifts.

Developing such a mechanism required an empirical basis for identifying and quantifying changes in demand. However, empirical calibration presents a methodological challenge. ED datasets commonly reported arrival times, urgency levels, waiting outcomes, and patient dispositions. Nevertheless, visit-level records did not necessarily capture the concurrent system states required to estimate every conditional transition probability of an MDP. Treating these records as a complete empirical transition trajectory could introduce assumptions that were not supported by the available data.

Given this limitation, the present study used the 2021 national hospital ambulatory medical care survey ED public-use data [22] as a calibration source. The data were used to characterize observable demand features, particularly five-level urgency composition and time-of-day variation in patient arrivals. They were not treated as a complete empirical MDP transition matrix. This distinction provided a transparent connection between the empirical data and the stochastic decision model.

A stationary nominal transition kernel was generated from the stochastic ED model. Hour-specific transition kernels were then constructed using the empirically calibrated demand profile. The Wasserstein-1 distance between each hour-specific kernel and the nominal kernel quantified the corresponding transition shift. These calibrated shifts provided the empirical basis for adapting the level of distributional protection across changing operating conditions.

Building on this calibration, the proposed Data-Calibrated Adaptive Wasserstein Robust Value Iteration (DC-AWRVI) adjusted the ambiguity radius according to congestion-related transition shifts. The baseline radius was retained when demand remained at or below the nominal level. When demand exceeded this level, the radius increased according to the calibrated Wasserstein shift, subject to an empirically determined upper bound. This directional mechanism avoided increasing conservatism when a distributional deviation does not indicate greater congestion pressure. The Bellman residual was used exclusively to assess numerical convergence and was not interpreted as a measure of environmental distribution shift.

Viewed as an automated decision system, the mechanism forms a feedback loop. The observed arrival rate is the input; the exact Wasserstein-1 distance between the corresponding transition kernel and the nominal kernel converts it into a radius; the radius selects a stored robust policy; and the policy, evaluated at the current queue state, gives the scheduling decision at each five-minute epoch, that is, how many waiting patients of each urgency class begin physician service. No parameter has to be tuned by an operator, and no robust Bellman equation is solved while the system operates, because the robust policies are computed offline at a set of radii. This is the sense in which the framework addresses automated patient scheduling in the ED: the degree of protection is set by the system from observed data and not by a fixed design choice. The loop was evaluated in simulation with the arrival rate taken from the demand profile; estimation error in the arrival rate and deployment in a hospital information system were not studied.

This study made three principal contributions. First, it introduced a transparent data-to-model calibration procedure that incorporated observable ED demand patterns without requiring a complete empirical MDP transition matrix. Second, it formulated five-level triage scheduling as a Wasserstein distributionally robust MDP and examined how robustification affects patient scheduling decisions. Third, and centrally, it developed a computable procedure that converts an observed deviation of the patient-arrival distribution from the nominal model into a level of robust protection: the Wasserstein radius is the 95th percentile, over state–action pairs, of the exact Wasserstein-1 distance between the arrival-rate-shifted and the nominal transition kernels, bounded between a baseline and a maximum value, and a library of robust policies computed offline at selected radii supplies the policy for the current radius. The procedure uses only the observed arrival rates and requires no manual tuning of the radius. The resulting robustness–efficiency trade-off was evaluated under nominal, empirically calibrated time-varying, and controlled-surge demand conditions.

The proposed framework was intended as a data-calibrated decision-support model rather than a digital representation of a specific hospital. It did not assume that greater robustness necessarily reduces operational costs. Instead, it provided a systematic mechanism for adjusting distributional protection when demand conditions departed from the nominal environment. The adaptive-radius mechanism varies the degree of protection against distributional uncertainty; it does not by itself guarantee lower waiting costs or improved operational reliability.

The remainder of this paper is organized as follows. Section 2 presents the theoretical foundations, ED scheduling model, empirical calibration procedure, and the proposed DC-AWRVI framework. Section 3 reports and discusses the numerical and operational results. Section 4 concludes the study and outlines its limitations and directions for future research.

2. THEORETICAL FOUNDATIONS AND METHODOLOGICAL DEVELOPMENT OF THE DC-AWRVI FRAMEWORK

2.1 Markov decision process and value iteration

An MDP provides a mathematical framework for sequential decision-making under uncertainty [23, 24]. In the present study, the MDP was used to represent the allocation of available physicians to waiting ED patients according to the current system state. The scheduling problem was formulated as a single-stage stochastic patient scheduling problem, in which each decision determined which waiting patients were selected to begin their first physician service. The model did not optimize subsequent clinical treatment stages or changes in physician staffing levels. A discounted MDP is represented by

$$\mathcal{M} = (\mathcal{S}, \mathcal{A}, P, c, \gamma) \tag{1}$$

where, $\mathcal{S}$ and $\mathcal{A}$ denote the state and action spaces, $P(\cdot \mid s, a)$ is the transition kernel, $c(s, a)$ is the one-stage cost, and $\gamma \in (0, 1)$ is the discount factor. Under the Markov property, the next-state distribution depends only on the current state and action,

$$P(s_{t+1} \mid s_0, a_0, \ldots, s_t, a_t) = P(s_{t+1} \mid s_t, a_t) \tag{2}$$

For a stationary policy π, the expected discounted cost is

$$J^{\pi}(s) = \mathbb{E}_s^{\pi}\left[\sum_{t=0}^{\infty} \gamma^t \, c(s_t, \pi(s_t))\right] \tag{3}$$

The optimal value function satisfies the Bellman equation [23, 24],

$$V^*(s) = min_{a \in \mathcal{A}(s)}\left[\begin{matrix} c(s,a) \\ +\gamma \sum_{s'} P\,(s' \mid s, a) V^*(s') \end{matrix}\right] \tag{4}$$

Defining the Bellman operator T,

$$(TV)(s) = min_{a\in\mathcal{A}(s)}\left[c(s,a) + \gamma\mathbb{E}_{P(\cdot|s,a)}[V(s')]\right] \quad (5)$$

The optimal value function satisfies $V^* = TV^*$. For a finite discounted MDP, T is a contraction under the sup norm [24],

$$\|TV - TU\|_\infty \le \gamma\|V - U\|_\infty \quad (6)$$

which establishes convergence of Value Iteration (VI) [24]. Starting from V_0,

$$V_{k+1} = TV_k \quad (7)$$

and the iteration terminates when the Bellman residual satisfies

$$e_k = \|V_{k+1} - V_k\|_\infty \le \varepsilon \quad (8)$$

This formulation is used as the NVI benchmark in the present study.

2.2 Distributionally robust Markov decision process with Wasserstein ambiguity

Conventional MDPs assume that the transition kernel is sufficiently well specified. Robust dynamic programming relaxes this assumption by allowing transition probabilities to vary within prescribed uncertainty sets [12, 13]. Distributionally robust MDPs extend this concept by considering ambiguity over probability distributions governing transition dynamics [14-16].

A generic distributionally robust optimization problem is expressed as [25]

$$min_{x\in X} sup_{P\in\mathcal{P}}\mathbb{E}_P[f(x,\xi)] \quad (9)$$

where, $\mathcal{P}$ is an ambiguity set of plausible probability distributions.

In this study, the ambiguity set is constructed using the Wasserstein-1 distance, which measures distributional discrepancy through optimal transport [17]. For distributions P and Q,

$$W_1(P,Q) = \inf_{\Gamma\in\Pi(P,Q)}\int d(x,y)d\Gamma(x,y) \quad (10)$$

where, $\Pi(P, Q)$ denotes the set of couplings between $P$and Q.

A Wasserstein ambiguity set centered at the nominal transition kernel is defined as [18]

$$\mathcal{P}_\rho(s,a) = \{P: W_1(P, P_0(\cdot|\, s,a)) \le \rho\} \quad (11)$$

where, ρ controls the degree of distributional protection.

Under a state-action rectangular ambiguity structure, the robust Bellman operator becomes [12, 13, 16]

$$\left(T_\rho V\right)(s) = min_{a\in\mathcal{A}(s)}\begin{bmatrix} c(s,a) + \\ \gamma \sup\limits_{P\in\mathcal{P}_\rho(s,a)} \mathbb{E}_P[V(s')] \end{bmatrix} \quad (12)$$

Since the next-state distribution is discrete, the nominal transition kernel for a given state-action pair can be represented as

$$P_0(\cdot|\, s,a) = \sum_q p_{0,q}\,(s,a)\delta_{s^{(q)}} \quad (13)$$

where, $\delta_{s^{(q)}}$ denotes the Dirac measure concentrated at next state $s^{(q)}$, and $p_{0,q}(s, a)$ is its associated transition probability. Using Wasserstein duality [18, 26], the inner worst-case expectation can be specialized to the finite-state setting as

$$\sup_{P:W_1(P,P_0)\le\rho}\mathbb{E}_P[V] = \inf_{\eta\ge0}\left\{\eta\rho + \sum_q p_{0,q}\max_r\left[V(s^{(r)}) - \eta d(s^{(r)}, s^{(q)})\right]\right\} \quad (14)$$

This dual representation enables the robust Bellman equation to be solved computationally without explicitly enumerating all probability distributions within the ambiguity set [18, 19, 26].

For a fixed ambiguity radius, Robust Value Iteration is defined by [15, 19]

$$V_{k+1} = T_\rho V_k \quad (15)$$

and is referred to hereafter as Fixed-Radius Robust Value Iteration (FRVI).

2.3 Emergency department scheduling model

The theoretical framework is applied to dynamic patient scheduling in an ED with five urgency classes. Dynamic and priority-based patient scheduling has previously been investigated using MDP, dynamic scheduling, and sequential resource-allocation models [4, 5, 7, 11, 27]. At decision epoch t, the state is defined as

$$s_t = (n_{1,t}, n_{2,t}, n_{3,t}, n_{4,t}, n_{5,t}) \quad (16)$$

where, $n_{i,t}$ denotes the number of waiting patients in urgency class i, and d_t denotes the number of available physicians.

Patient scheduling as a stochastic scheduling problem. The model can be read as a dynamic stochastic job scheduling problem on parallel servers. The patients are the jobs: they arrive at random times and belong to one of five urgency classes. The physicians are identical servers whose number is fixed within an experiment; the urgency weights α_i play the role of job weights, and the cost is the weighted waiting burden of the jobs that have not started. At each five-minute epoch, the decision is which waiting jobs start on the servers that are free, without preemption, and the decision is repeated for the jobs still waiting. The problem therefore remains one of patient scheduling: the decision is the start of service of waiting patients by urgency class, and physician capacity is a constraint and not a decision. Two features separate it from a deterministic job scheduling problem: arrivals and service completions are random, and the arrival law itself is uncertain, which is the reason for the ambiguity set introduced in Section 2.2.

The scheduling action is

$$a_t = (a_{1,t}, a_{2,t}, a_{3,t}, a_{4,t}, a_{5,t}) \quad (17)$$

subject to

$$0 \le a_{i,t} \le n_{i,t}, \sum_{i=1}^{5} a_{i,t} \le d_t \tag{18}$$

where, $a_{i,t}$ denotes the number of patients from urgency class i selected to begin service at decision epoch t. Thus, the action vector represents a class-level scheduling decision, namely the number of waiting patients of each urgency class that begin service, and not the selection of individual patients.

The immediate cost represents the urgency-weighted waiting burden remaining after allocation,

$$c(s_t, a_t) = \Delta \sum_{i=1}^{5} \alpha_i \, (n_{i,t} - a_{i,t}) \tag{19}$$

with

$$\alpha = (5,4,3,2,1), \tag{20}$$

where, α_i denotes the priority weight assigned to urgency class i. The vector $\alpha = (5, 4, 3, 2, 1)$ assigns progressively larger waiting penalties to higher-urgency patients, ensuring that delays in more critical classes contribute more strongly to the objective function. This weighting is of the same form as the acuity weights 5, 4, 3, 2, 1 used by the study [4].

Patient arrivals were modeled as class-specific Poisson processes, consistent with standard queueing representations used in healthcare operations [28],

$$A_{i,t} \sim Poisson(\lambda p_i \Delta_h) \tag{21}$$

where, p_i denotes the urgency-class probability. After the allocation decision, the remaining number of available physicians is $d_t^+ = d_t - \sum_{i=1}^{5} a_{i,t}$. Accordingly, the number of physicians engaged in service is $B_t = D - d_t^+$, which is equivalently expressed as

$$B_t = (D - d_t) + \sum_{i=1}^{5} a_{i,t} \tag{22}$$

Assuming exponentially distributed service times with rate μ [28], the probability of completion during one decision interval is

$$p_c = 1 - e^{(-\mu \Delta_h)} \tag{23}$$

where, p_c denotes the probability that a physician who is currently engaged in service completes the ongoing treatment within one decision interval. Under the exponential service-time assumption with rate μ, this probability is given by $1 - \exp(-\mu \Delta_h)$.

Given B_t physicians engaged in service and assuming independent completion events with probability p_c, the number of service completions during epoch t, denoted by C_t, follows a binomial distribution. Each completed service releases one physician back to the available physician pool for the subsequent decision epoch, so that

$$C_t \sim Binomial(B_t, p_c) \tag{24}$$

While C_t captures service completions, the patient queues evolve according to the balance between allocated patients and new arrivals. Thus, the queue length of urgency class i at the next decision epoch is

$$n_{i,t+1} = min\{K, n_{i,t} - a_{i,t} + A_{i,t}\} \tag{25}$$

and number of physicians available at the beginning of the next decision epoch is

$$d_{t+1} = d_t - \sum_{i=1}^{5} a_{i,t} + C_t \tag{26}$$

The transition kernel Eq. (27) characterizes the probability distribution of the next system state conditional on the current state and allocation decision. In the present model, this distribution was induced by stochastic patient arrivals and service completions and was used to evaluate the expected future cost in the Bellman equation.

$$P_0(\cdot \mid s, a) \tag{27}$$

Remark on the distributional assumptions. The class-specific Poisson arrivals and exponentially distributed service times are adopted as a tractable representation of ED operations and not as an exact description of any particular department. The Poisson assumption gives a compact distribution for the number of arrivals within a decision interval. In the nominal transition model the arrival intensity is constant at λ_0. Under the NHAMCS-calibrated time-varying scenario, arrivals follow a nonhomogeneous Poisson process (NHPP) with piecewise-constant hourly intensity $\lambda_h = \lambda_0 r_h$ (Eq. (29), Section 2.4), so that the number of class-i arrivals in a five-minute interval is Poisson($\lambda_{h(t)} p_i \Delta$) with Δ = 5/60 hour, where, p_i is the calibrated triage-class probability. This accommodates hourly variation in demand while keeping the Poisson form within each decision interval.

The exponential service-time assumption supports a Markovian representation of service completions and physician availability. By the memoryless property, the completion probability $p_c = 1 - \exp(-\mu\Delta_h)$ in Eq. (23) does not depend on the elapsed service time, so the elapsed time need not be added to the state, which keeps the finite-state MDP tractable. The assumed service rate, μ = 2.5 patients per hour per physician (a mean service duration of 24 minutes), is an experimental assumption and not a parameter estimated from NHAMCS.

These assumptions have limitations. Actual arrivals may show clustering, overdispersion, or temporal dependence beyond that represented by the NHPP, and service durations may depend on patient complexity, clinical urgency, physician workload, and operating conditions, so they may deviate from the exponential distribution. Because the transition kernels, and therefore the Wasserstein shifts, radii, and policies, are derived from these assumptions, the numerical findings should be interpreted within the specified stochastic modeling framework and not as direct predictions for every ED.

2.4 Empirical demand calibration and transition metric

The 2021 NHAMCS ED public-use data were used to calibrate observable demand characteristics [22]. Because the dataset contains individual visits rather than continuous

system-state trajectories, it was used to estimate urgency composition and relative hourly arrival patterns, rather than the completed MDP transition matrix.

The calibrated urgency distribution was

$$p = (0.042868, 0.150370, 0.504397, 0.266113, 0.036251) \tag{28}$$

With a nominal arrival rate of

$$\lambda_0 = 7$$

patients per hour, the hour-specific rate is

$$\lambda_h = \lambda_0 r_h \tag{29}$$

where, r_h denotes the empirically derived relative hourly demand factor. Each λ_h induces an alternative transition kernel $P_h(\cdot\,|s,a)$

Following norm-based ground-metric formulations commonly used in Wasserstein distributionally robust optimization [17, 18], a normalized metric was defined for the six-dimensional ED state space as

$$d_K\left(s^{(r)}, s^{(q)}\right) = \frac{1}{6}\left[\sum_{i=1}^{5} \frac{|n_i^{(r)} - n_i^{(q)}|}{K} + \frac{|d^{(r)} - d^{(q)}|}{D}\right] \tag{30}$$

The normalization by K and D, together with the factor $1/6$, was introduced in this study to place all state components on a comparable scale and to obtain a bounded metric Eq. (31).

$$0 \le d_K(s^{(r)}, s^{(q)}) \le 1 \tag{31}$$

2.5 Proposed data-calibrated adaptive Wasserstein robust value iteration

Previous studies have shown that Wasserstein ambiguity can be incorporated into multistage dynamic programming and data-driven stochastic control [29, 30]. Building on this foundation, DC-AWRVI extends FRVI by linking the ambiguity radius to demand-induced transition shifts. Unlike FRVI, which applies a single predetermined radius to all operating conditions, DC-AWRVI adjusts the level of distributional protection according to the magnitude and direction of the observed demand shift.

For each demand condition h, the state-action-specific transition shift is measured by

$$\omega_h(s,a) = W_1(P_h(\cdot|\,s,a), P_0(\cdot|\,s,a)) \tag{32}$$

where, $P_0(\cdot|\,s,a)$ denotes the nominal transition kernel and $P_h(\cdot|\,s,a)$ denotes the transition kernel induced by the hourly demand condition. Because the shift magnitude varies across feasible state-action pairs, an hourly summary is obtained using the 95th percentile,

$$z_h = Q_{0.95}(\{\omega_h(s,a): s \in \mathcal{S}, a \in \mathcal{A}(s)\}) \tag{33}$$

Thus, $\omega_h(s,a)$ quantifies the transition shift for an individual state-action pair, where, z_h summarizes the upper range of transition shifts for demand condition h.

The 95th percentile in Eq. (33) is a risk-oriented calibration convention. It is intended to reflect relatively large upward deviations in arrival demand while avoiding the most extreme observed deviations as the default ambiguity radius. It was not selected by a formal optimization procedure, and it is not claimed to be optimal. It should also be distinguished from the confidence levels of the conditional value-at-risk (CVaR) used in Section 3.3 (90%, 95%, and 99%): these quantify the upper-tail risk of realized scheduling cost, are evaluation parameters only, and play no role in defining the Wasserstein ambiguity set. The sensitivity of the radii and of the performance to this choice was examined with the 90th and 99th percentiles and is reported in Section 3.4.

Empirical calibration yielded the lower and upper ambiguity-radius levels

$$\rho_{\text{base}} = 0.00126754, \rho_{max} = 0.01528771 \tag{34}$$

The adaptive radius is then defined as

$$\rho_h = \begin{cases} \rho_{\text{base}}, & \lambda_h \le \lambda_0, \\ \min\{\rho_{\max}, \max\{\rho_{\text{base}}, z_h\}\}, & \lambda_h > \lambda_0. \end{cases} \tag{35}$$

This rule distinguishes the magnitude of distributional displacement, represented by z_h, from its operational direction, represented by the comparison between λ_h and the nominal arrival rate λ_0. Consequently, demand conditions at or below the nominal level retain the baseline protection, whereas upward-demand shifts may activate a larger Wasserstein radius.

To avoid repeatedly solving a new robust Bellman problem for every value of ρ_h, a finite robust-policy library was constructed offline. The calibrated radius nodes were

$$\mathcal{R} = \{0.00126754, 0.00636465, 0.00878870, 0.01259272, 0.01528771\} \tag{36}$$

For each $\rho^{(m)} \in \mathcal{R}$, the robust value function is obtained from

$$V_{k+1}^{(m)} = T_{\rho^{(m)}} V_k^{(m)} \tag{37}$$

with iteration continued until

$$\left\|V_{k+1}^{(m)} - V_k^{(m)}\right\|_\infty \le \varepsilon \tag{38}$$

The resulting robust policy $\pi^*_{\rho^{(m)}}$ is stored in the policy library. For each demand condition h, the calibrated adaptive radius is mapped to the nearest available policy node,

$$\tilde{\rho}_h = argmin_{\rho^{(m)} \in \mathcal{R}} \left|\rho_h - \rho^{(m)}\right| \tag{39}$$

and the scheduling action is selected as

$$a_t = \pi^*_{\tilde{\rho}_h}(s_t) \tag{40}$$

Accordingly, the complete DC-AWRVI mechanism can be summarized as

$$\lambda_h \to P_h \to \omega_h(s,a) \to z_h \to \rho_h \to \tilde{\rho}_h \to \pi^*_{\tilde{\rho}_h}$$

The algorithmic contribution therefore lies in coupling the

Wasserstein robust Bellman solution with a data-calibrated radius-selection mechanism and an offline library of robust policies. The Bellman residual is used only as a numerical convergence criterion and does not determine the adaptive ambiguity radius. The framework of this study is described in Figure 1.

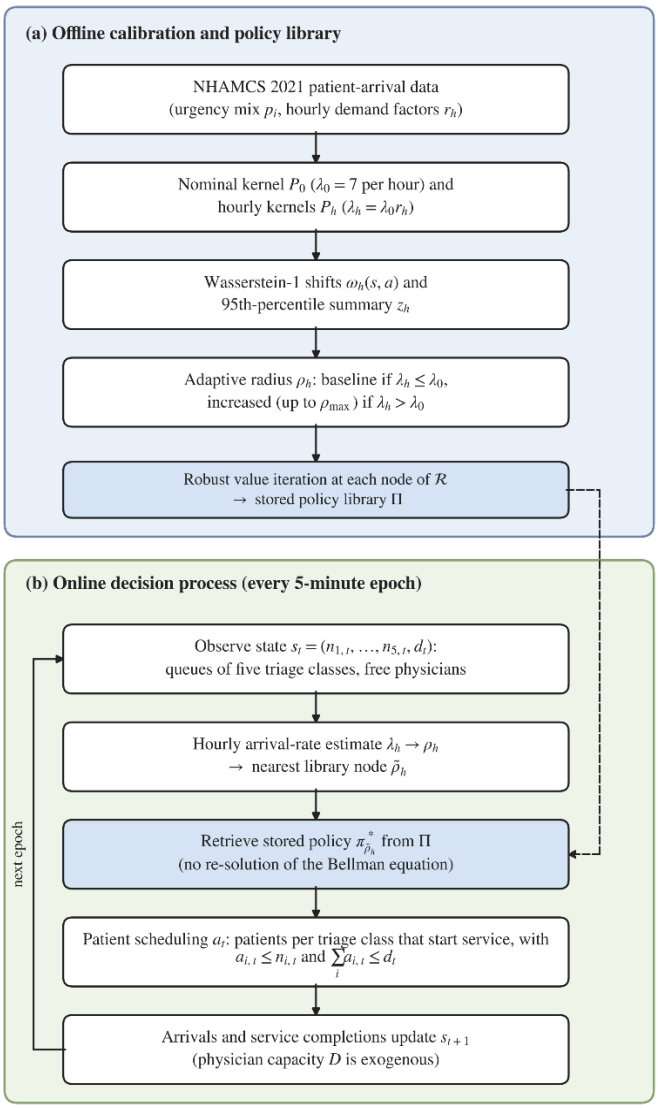


Figure 1. Framework of the Data-Calibrated Adaptive Wasserstein Robust Value Iteration (DC-AWRVI) patient-scheduling process: (a) offline calibration and policy-library construction; (b) online policy selection and patient scheduling at each decision epoch

Figure 1 shows how these components form an automated decision process. The framework operates over five-minute decision epochs within an eight-hour scheduling horizon. Offline (Figure 1(a)), the NHAMCS-derived urgency mix and hourly demand factors define the nominal and hour-specific transition kernels; the Wasserstein-1 shifts between them are summarized by z_h and mapped to an adaptive radius ρ_h; and robust value iteration is solved at each radius node to build the policy library. Online (Figure 1(b)), at each decision epoch the system observes the queue length of each of the five triage classes and the number of available physicians, uses the current hourly arrival-rate estimate to obtain the adaptive radius and its nearest library node, and retrieves the corresponding stored policy. When physician capacity is available, the policy evaluated at the current state selects the triage classes from which waiting patients begin first physician service, subject to $a_{i,t} \le n_{i,t}$ and $\Sigma a_{i,t} \le d_t$. Arrivals and service completions then update the state, and the process continues over successive decision epochs. Physician capacity D is treated as an exogenous operational constraint and not as a decision variable, so the framework does not change staffing levels or optimize multistage clinical treatment. Its automated decision component is the selection of waiting patients for physician service under uncertain demand. The service is single-stage: a patient is selected once for first physician service, and the decision is repeated at every epoch of the horizon for the patients still waiting. The study does not describe a deployed or clinically validated hospital scheduling system.

The robust policies are computed offline for a finite set of radii, so a change in the estimated arrival rate changes the selected policy without solving the robust Bellman equation again. During online operation, the current arrival-rate estimate gives the adaptive radius ρ_h (Eq. (35)), and the nearest library node (Eq. (39)) determines which stored policy is evaluated at the current state. For example, with the library $\mathcal{R}$ = {0.00126754, 0.00636465, 0.00878870, 0.01259272, 0.01528771}, an arrival rate of 8.6291 patients per hour gives an adaptive radius of 0.00565423, and the nearest available policy corresponds to the node 0.00636465. When the arrival rate rises to 9.1987 patients per hour, the adaptive radius rises to 0.00763075, and the selected policy switches to the node 0.00878870. In either case, the selected policy is evaluated at the current state (the queue lengths of the five triage classes and the number of available physicians), and the resulting action specifies how many waiting patients of each class begin first physician service, within the available physician capacity and the constraints in Eq. (18).

The library gives a computationally efficient response to changes in estimated demand, but it does not guarantee that the nearest stored policy is optimal for an intermediate radius. In the time-varying experiment, the arrival-rate profile is updated hourly, whereas queue states and scheduling decisions are evaluated every five minutes. An unexpected within-hour change in demand is therefore reflected immediately in the observed queue state, but it does not change the radius-selection input until the arrival-rate estimate is updated.

From an observed arrival deviation to a level of protection. The link between the observed arrival distribution and the robust protection consists of five computable steps. (i) The hourly arrival rate λ_h is estimated from the data. (ii) It defines the hour-specific kernel $P_h(\cdot|s, a)$, which differs from the nominal kernel $P_0(\cdot|s, a)$ only through the Poisson arrival means of the five triage classes. (iii) For every state–action pair the distance $\omega_h(s, a) = W_1(P_h, P_0)$ is computed exactly: because the kernels are products over the classes and the ground metric is a normalized L1 metric, it equals the sum of the class-wise one-dimensional Wasserstein distances between the capped Poisson laws, scaled by the normalization constant of the ground metric. (iv) The radius is the 95th percentile of this distance over the state–action pairs, clipped to the interval $[\rho_{base}, \rho_{max}]$. (v) The radius is mapped to the nearest node of a library of robust policies solved offline, and the stored policy is evaluated at the current state. As a result, the level of protection increases with the deviation of the observed arrival rate from the nominal rate (from 0.00127 at the nominal rate to 0.01529 at 12 patients per hour, Section 3.4), it is obtained from the observed rates alone without

simulation or manual tuning of the radius, and it is reproducible. This link is what the framework contributes; Section 3 evaluates the resulting protection costs and what it does not achieve.

In the data-driven Wasserstein formulation in this study [18], the radius is derived from a measure-concentration bound that depends on the number of observations and on a confidence level, and it shrinks as data accumulate; because this a priori bound is conservative, the authors recommend calibrating the radius by bootstrapping or cross-validation. The radius used here has a different purpose and a different status. It is not a confidence set for an unknown distribution estimated from a sample; it measures how far the arrival process currently observed lies from the nominal model, and it scales the protection with that distance. It therefore carries no finite-sample guarantee, and its 95th-percentile and clipping choices are conventions that were not calibrated against out-of-sample performance.

Steps (i)–(v) and the stored library close a loop between observation and decision: the arrival-rate estimate is the measured input, the radius is computed from it by a fixed rule, the nearest library node selects the policy, and the policy evaluated at the observed state is the control action. The loop is automatic in that none of its elements requires operator intervention, and it is light online because the demanding computation, robust value iteration at each radius node, is done offline. In the experiments the arrival-rate input is the profile of the scenario and is updated hourly; the effect of estimation error in this input, of a faster update, and of a delay between estimate and decision were not examined.

Algorithm 1. Data-Calibrated Adaptive Wasserstein Robust Value Iteration

Input: Nominal arrival rate λ_0; hourly demand rates $\{\lambda_h\}$; state space $\mathcal{S}$; feasible action sets $\mathcal{A}(s)$; nominal transition kernel P_0; discount factor γ; convergence tolerance ε.
Parameters: percentile level q (0.95 in the experiments); clipping bounds ρ_{base} and $\rho_{\max}$.
Output: Robust policy library Π and adaptive scheduling policy $\pi^*_{\tilde{\rho}_h}$.

Phase 1 (offline or whenever the arrival-rate estimate is updated): radius calibration from the observed arrival rate

1: for each demand condition h do
2: Construct $P_h(\cdot|s, a)$ using the estimated arrival rate λ_h
3: for each feasible state-action pair (s, a) do
4: $\omega_h(s, a) \leftarrow \mathrm{W}_1(P_h(\cdot|s, a), P_0(\cdot|s, a))$
5: end for
6: $z_h \leftarrow Q_q(\{\omega_h(s,a): s \in S, a \in A(s)\})$
7: if $\lambda_h \le \lambda_0$ then
8: $\rho_{\text{h}} \leftarrow \rho_{\text{base}}$
9: else
10: $\rho_{\text{h}} \leftarrow \min\{\rho_{\max}, \max\{\rho\text{base}, \text{zh}\}\}$
11: end if
12: end for

Phase 2 (offline): robust policy library

13: Set $R \leftarrow$ {0.00126754, 0.00636465, 0.00878870, 0.01259272, 0.01528771}
14: for each $\rho^{(m)} \in R$ do
15: Initialize $V_0^{(m)}$
16: $k \leftarrow 0$
17: repeat
18: $V_{k+1}^{(m)} \leftarrow \mathrm{T}\rho^{(m)}\ \mathrm{Vk(m)}$
19: $e_k \leftarrow \|V_{k+1}^{(m)} - V_k^{(\mathrm{m})}\|_\infty$
20: $k \leftarrow k + 1$
21: until $e_k \le \varepsilon$
22: Extract $\pi^*\rho^{(m)}$
23: Store $\pi^*\rho^{(m)}$ in Π
24: end for

Phase 3 (online): policy selection and patient scheduling at each decision epoch

25: for each demand condition h do
26: $\tilde{\rho}_h \leftarrow \arg\min \rho^{(m)} \in \mathrm{R}\ |\rho_h - \rho^{(m)}|$
27: for each decision epoch t under condition h do
28: Obtain current state s_t
29: $a_t \leftarrow \pi^*\tilde{\rho}_h(s_t)$
30: Apply a_t (the number of waiting patients of each triage class that start physician service, subject to Eq. (18)) and update the ED state
31: end for
32: end for
33: return Π and $\pi^*\tilde{\rho}_h$.

2.6 Computational design

The computational model considered five urgency classes and three physicians. Decisions were made at five-minute intervals over an eight-hour operating horizon, corresponding to 96 decision epochs. The physician service rate was set to $\mu = 2.5$ patients per hour, and the discount factor was $\gamma = 0.95$.

These settings define a controlled, resource-constrained scheduling environment and not the staffing structure of a particular hospital. The five-minute decision interval is short enough to capture changes in queue lengths and physician availability, and the eight-hour horizon represents one operating shift, over which the same single-stage decision is repeated at each of 96 decision epochs. Table 1 lists the principal parameters and separates empirical inputs (the urgency mix and the relative hourly demand, both from NHAMCS) from modeling assumptions (the nominal arrival rate, number of physicians, service rate, priority weights, discount factor, decision interval, and horizon). In particular, the service rate of 2.5 patients per hour per physician, equivalent to a mean service duration of 24 minutes, is an experimental assumption and not an NHAMCS estimate. The configuration was not calibrated to a formally defined ED size category; applying the model to a specific hospital would require site-specific calibration of staffing, patient arrivals, service times, and operational constraints.

To maintain computational tractability, policy construction used a queue bound of $K_{\text{DP}} = 4$, yielding

$$|\mathcal{S}| = (K_{\text{DP}} + 1)^5(D + 1) = 12500$$

states and 161,000 feasible state-action pairs. Operational simulation used a larger queue bound, $K_{\text{sim}} = 30$, to reduce artificial truncation under high-demand conditions. Simulation states exceeding the policy-state bound were mapped to the finite dynamic programming state space using componentwise clipping,

$$\phi(n_i) = \min\{n_i, K_{\text{DP}}\}$$

The state-space sizes quoted above (4,096 states and 45,536 state–action pairs) correspond to the bound used in the numerical verification of Sections 3.1 and 3.2. The operational comparisons of Sections 3.3 and 3.4 used the bound $K_{DP} = 4$,

that is, 12,500 policy states and 161,000 feasible state–action pairs. Componentwise clipping maps every queue length above K_{DP} to the boundary value, so distinct high-congestion states can become indistinguishable to the scheduling policy and may affect how physician service is allocated among triage classes. The larger simulation limit K_{sim} reduces the truncation of the simulated queues but does not remove this information loss in the policy state. The sensitivity of the decisions to K_{DP} is examined in Section 3.5.

Table 1. Summary of model parameters, their values, and their basis

Parameter	Value	Basis
Urgency classes	5 (class 1 most urgent)	NHAMCS triage levels (empirical)
Urgency mix p	(0.042868, 0.150370, 0.504397, 0.266113, 0.036251)	NHAMCS 2021 (empirical)
Relative hourly demand r_h	24 hourly factors	NHAMCS 2021 (empirical)
Nominal arrival rate λ_0	7 patients/h	Assumed
Physicians D	3 (3–7 in the capacity sensitivity analysis)	Assumed
Service rate μ	2.5 patients/h per physician (mean 24 min)	Assumed; not estimated from NHAMCS
Decision interval Δ	5 min (1/12 h)	Design choice
Operating horizon	8 h = 96 epochs (08:00–16:00)	Design choice; high-arrival window of the NHAMCS profile
Priority weights α	(5, 4, 3, 2, 1)	Assumed
Discount factor γ	0.95	Assumed
Policy-state queue bound K_{DP}	4 (12,500 states) in the operational comparison; 3 (4,096 states) in the verification experiments (Sections 3.1–3.2)	Computational choice
Simulation queue bound K_{sim}	30	Computational choice
Radius calibration quantile	0.95	Calibration convention (Section 2.5)
Dual-grid resolution	513 points	Verified in Section 3.1
Arrival-rate sensitivity	7 to 12 patients/h, constant over the horizon	Sensitivity analysis (Section 3.4)
Monte Carlo replications	2,000 per policy–scenario (main); 2,000 per arrival rate and per surge profile (Section 3.4); 1,000 per configuration (Section 3.5)	Computational choice

Three scheduling policies were evaluated: NVI, Fixed-Radius Robust Value Iteration (FRVI), and the proposed DC-AWRVI. The fixed-radius benchmark used

$$\rho_{\mathrm{FRVI}} = 0.00878870$$

Performance was evaluated under three demand conditions: a nominal scenario with an arrival rate of 7 patients per hour, an NHAMCS-calibrated time-varying demand profile over the 08:00–16:00 operational window and a controlled surge scenario with 10 patients per hour. Each scenario was simulated over 2,000 replications using common random numbers across the competing policies to improve the precision of paired comparisons.

Two-time spans are used and should not be confused. The complete 24-hour NHAMCS-derived arrival profile is the *calibration profile*: it characterizes hourly variation in demand and is used to calibrate the Wasserstein ambiguity radii. The eight-hour *operational simulation window*, 08:00–16:00, was selected from the hourly arrival rates as a period of relatively high patient arrivals, so that the policies are compared in a demand-intensive setting. It comprises 96 consecutive five-minute decision epochs. In the time-varying scenario, the hourly arrival rates for 08:00–15:00 are applied in sequence, each held for 12 decision epochs; the nominal and controlled-surge scenarios are simulated over the same eight-hour horizon to keep the comparison consistent.

The evaluation metrics included cumulative priority-weighted waiting cost, mean queue length, peak queue length, final queue length, physician utilization, service completions, and tail-risk performance measured by $\mathrm{CVaR}_{0.95}$. Numerical verification of the Bellman solvers and Wasserstein robustification, followed by the comparative operational results, is presented in Section 3. CVaR is evaluated at the 90%, 95%, and 99% levels; the 95% level is used in the tables and the other levels are reported in Section 3.3.

A separate sensitivity experiment varied the number of physicians from three to seven and the per-class queue bound from three to seven, using 1,000 stochastic replications per configuration, and evaluated waiting cost, queue length, overflow, and service completions. It is distinct from the main comparison of NVI, FRVI, and DC-AWRVI, which uses 2,000 replications per policy–scenario combination. The response of the policies to the arrival rate and to demand surges is examined in Section 3.4.

3. RESULT AND DISCUSSION

3.1 Numerical verification of the nominal and robust Bellman solvers

The numerical implementation was first verified to ensure consistency of the nominal Bellman solver, the Wasserstein ground metric, and the robust continuation operator. The finite decision model comprised 4,096 states, 56 global patient scheduling actions, and 45,536 feasible state-action pairs. NVI, using a discount factor of $\gamma = 0.95$, converged after 424 iterations with a final Bellman residual of approximately 9.999×10^{-9}, confirming numerical convergence of the nominal solution.

The normalized Wasserstein ground metric was subsequently verified over the six-dimensional state space consisting of five urgency-specific queues and physician availability. Numerical tests confirmed the required metric properties, including identity, symmetry, and the triangle inequality. Across 10,000 randomly generated state triplets, no meaningful violation was observed. The inner Wasserstein problem was then evaluated through its dual representation. At $\rho = 0$, the robust continuation value recovered the nominal

expectation, while increasing ρ produced a monotonic increase in the worst-case continuation value, as expected from the robust formulation.

Because the dual problem was solved using a discretized dual-variable grid, grid resolutions of 129, 257, and 513 points were compared. The corresponding mean/maximum relative errors were 0.07496%/0.40737%, 0.04695%/0.13875%, and 0.01576%/0.09033%, respectively. The 513-point grid was therefore retained because its maximum approximation error was below the predefined 0.1% verification threshold.

Figure 2 illustrates the progressive reduction in approximation error as the dual-grid resolution increased and confirms the suitability of the 513-point configuration for subsequent robust-policy computation.

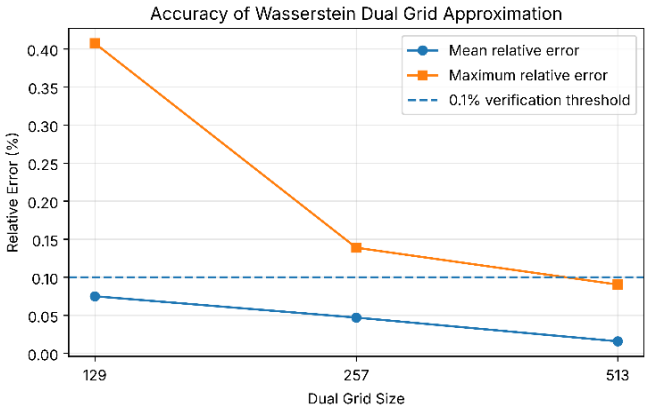


Figure 2. Accuracy of the Wasserstein dual-grid approximation

3.2 Adaptive robustness and policy adjustment

Wasserstein robustification altered scheduling decisions only in a limited subset of the state space. At the verification radius $\rho = 0.05$, the nominal and robust policies differed in 173 of the 4,096 states, corresponding to 4.22% of the policy space. At the smaller radius $\rho = 0.01$, only 37 states, or 0.90%, differed from NVI. Thus, increasing the ambiguity radius increased the number of modified decisions, but the effect remained selective rather than replacing the nominal scheduling policy globally.

The adaptive-radius mechanism was subsequently calibrated from demand-induced transition shifts. The nominal transition kernel associated with $\lambda_0 = 7$ patients per hour was compared with the hourly transition kernels derived from the NHAMCS-calibrated demand profile. Across 1,092,864 state-action kernel comparisons, the Wasserstein distances ranged from 0 to 0.02258. The overall distribution had values of approximately 0.00633, 0.00974, 0.01408, and 0.01875 at the 25th, 50th, 75th, and 95th percentiles, respectively.

The hourly pattern showed that Wasserstein distance captured the magnitude of transition-distribution deviation rather than its direction (see Figure 3). Consequently, the final adaptive mechanism distinguished upward-demand shifts from non-increasing demand so that additional robustness was activated only when the demand condition was operationally relevant to congestion.

The calibration yielded a baseline ambiguity radius of approximately

$$\rho_{\text{base}} = 0.00126754$$

and an upper bound of

$$\rho_{\max} = 0.01528771$$

Across the 24-hour calibration profile (used only for radius calibration), 11 hours remained at the baseline radius, while 13 hours activated a higher level of robustness. Four demand periods reached the calibrated upper bound. The mean adaptive radius was approximately 0.006855, and the median was 0.004387.

As shown in Figure 4, the ambiguity radius remained near the baseline level during non-increasing demand and increased selectively when the hourly arrival rate exceeded the nominal reference. This behavior demonstrates the central mechanism of DC-AWRVI: distributional protection was adapted according to the calibrated transition environment rather than maintained at a single fixed level.

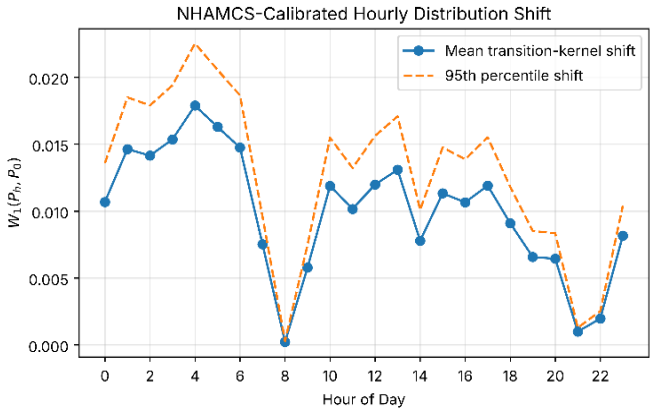


Figure 3. NHAMCS-calibrated hourly Wasserstein distribution shift over the 24-hour calibration profile

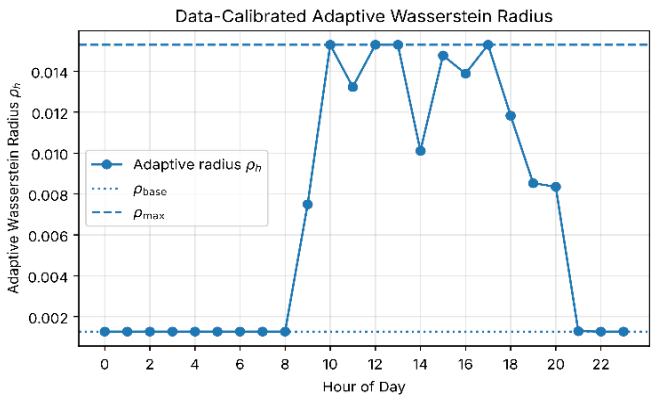


Figure 4. Data-calibrated adaptive Wasserstein radius over the 24-hour calibration profile

3.3 Operational performance and robustness–efficiency trade-off

The operational performance of NVI, FRVI, and DC-AWRVI was evaluated under three demand conditions: nominal demand, the NHAMCS-calibrated 08:00–16:00 profile, and a controlled surge of 10 patients per hour. Each scenario was simulated over 2,000 replications using common random numbers across the three policies.

Two senses of robustness need to be distinguished. Distributional robustness refers to the treatment of uncertainty in the transition distribution: the Wasserstein ambiguity set

defines a neighborhood of plausible transition distributions around the nominal model, and the robust Bellman operator evaluates the worst-case expected continuation cost over this set. A larger radius widens the set of distributions considered in policy optimization, which is a model-based way of accounting for ambiguity but does not by itself guarantee better operational performance. Operational robustness concerns the observed stability and reliability of scheduling outcomes under different demand conditions and has to be assessed empirically through waiting cost, queue length, variability across replications, and upper-tail waiting-cost risk. The results below are interpreted according to this distinction; a larger Wasserstein radius is not treated as evidence of more reliable scheduling.

Under nominal demand, the mean priority-weighted waiting costs were 4,709.34, 4,776.44, and 4,718.07 for NVI, FRVI, and DC-AWRVI, respectively. Under the NHAMCS-calibrated profile, the mean costs were 13,652.78, 13,828.32, and 13,842.99 and under controlled surge they were 16,100.02, 16,301.47, and 16,301.47. DC-AWRVI therefore did not achieve the lowest mean waiting cost in any of the three scenarios. Under controlled surge, FRVI and DC-AWRVI selected the same library node (0.00878870) and produced identical results.

Aggregate queue measures differed only marginally among the three policies (Table 2, which gives means and standard deviations across the 2,000 replications). Under the NHAMCS profile, the mean queue was 10.99 under NVI and 11.02 under DC-AWRVI (paired difference 0.028, 95% confidence interval 0.021 to 0.035), the peak queue was 25.26 and 25.31, and the final queue was 23.85 and 23.90; DC-AWRVI was 0.26%, 0.18%, and 0.22% higher than NVI on these three measures. Under controlled surge and nominal demand, the three measures were practically identical (differences below 0.05% under controlled surge and none under nominal demand). DC-AWRVI therefore did not reduce congestion relative to NVI. Statistical and practical importance are distinct: even when a small difference can be estimated precisely, its operational relevance depends on its magnitude, consistency, and consequences for patient waiting and physician utilization.

The corresponding weighted waiting-cost comparison is illustrated in Figure 5.

Table 2. Operational performance across demand scenarios (mean, with the standard deviation across 2,000 replications in parentheses)

Scenario	Method	Mean Cost	Mean Queue	Peak Queue	Final Queue
Nominal	NVI	4,709 (4,004)	4.19 (3.0)	10.7 (5.3)	7.3 (6.3)
	FRVI	4,776 (4,047)	4.19 (3.0)	10.7 (5.3)	7.3 (6.3)
	DC-AWRVI	4,718 (4,005)	4.19 (3.0)	10.7 (5.3)	7.3 (6.3)
NHAMCS 08–16	NVI	13,653 (6,139)	10.99 (4.5)	25.3 (8.5)	23.9 (9.0)
	FRVI	13,828 (6,196)	10.99 (4.5)	25.3 (8.5)	23.8 (9.0)
	DC-AWRVI	13,843 (6,181)	11.02 (4.5)	25.3 (8.5)	23.9 (9.0)
Controlled Surge	NVI	16,100 (6,910)	12.80 (5.1)	26.7 (8.6)	25.1 (9.0)
	FRVI	16,301 (6,978)	12.80 (5.1)	26.7 (8.6)	25.1 (9.0)
	DC-AWRVI	16,301 (6,978)	12.80 (5.1)	26.7 (8.6)	25.1 (9.0)

Note: NVI: Nominal Value Iteration, FRVI: Fixed-Radius Robust Value Iteration, DC-AWRVI: Data-Calibrated Adaptive Wasserstein Robust Value Iteration.

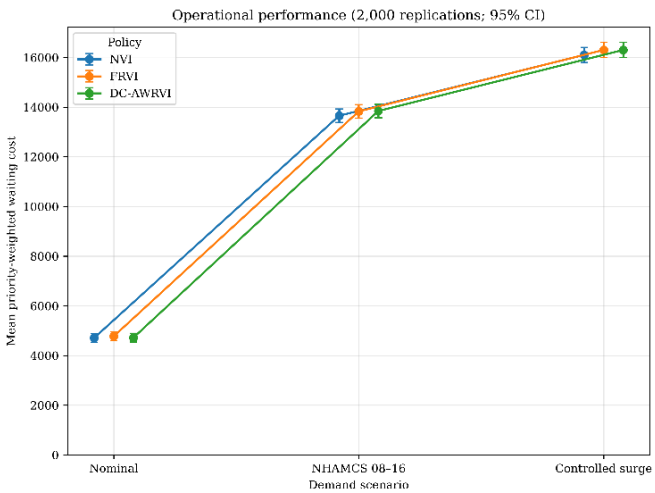


Figure 5. Mean priority-weighted waiting cost of NVI, FRVI, and DC-AWRVI under the three demand scenarios (error bars: 95% confidence intervals)

Note: NVI: Nominal Value Iteration, FRVI: Fixed-Radius Robust Value Iteration, DC-AWRVI: Data-Calibrated Adaptive Wasserstein Robust Value Iteration.

Figure 5 shows that the policies produced almost indistinguishable performance under nominal demand, whereas differences became more visible under the NHAMCS-calibrated and controlled-surge scenarios.

Tail-risk performance showed a similar pattern. $CVaR_{0.95}$ is reported as the mean cost over the replications at or above the empirical 95th percentile. Under nominal demand, $CVaR_{0.95}$ was 15,733.40 for NVI, 15,902.60 for FRVI, and 15,733.40 for DC-AWRVI. Under the NHAMCS-calibrated profile it was 27,072.25, 27,374.40, and 27,366.50, and under controlled surge 30,476.40, 30,828.60, and 30,828.60. Thus, adaptive robustification did not improve tail waiting-cost risk under the tested demand conditions: DC-AWRVI was 1.09% above NVI under the NHAMCS profile and 1.16% above NVI under controlled surge. The CVaR confidence level is an evaluation parameter and is unrelated to the percentile used to calibrate the ambiguity radius. The same pattern held at the other evaluation levels. Relative to NVI, the CVaR of DC-AWRVI was 1.02%, 1.09%, and 1.09% higher at the 90%, 95%, and 99% levels under the NHAMCS-calibrated profile, 1.07%, 1.16%, and 1.33% higher under controlled surge, and 0.02%, 0.00%, and 0.00% higher under nominal demand (CVaR at level a is the mean cost over the replications at or above the empirical a-quantile). In the arrival-rate and transient-surge experiments of Section 3.4 the CVaR of DC-AWRVI was never below that of NVI at any of the three levels (differences between 0.00% and at most 1.62%, 1.66%, and 1.53% at the 90%, 95%, and 99% levels).

Paired comparisons using common random numbers confirmed that DC-AWRVI incurred higher weighted waiting costs than NVI. Relative to NVI, the mean paired cost difference was 8.73 under nominal demand (95% confidence

interval 4.91 to 12.54), 190.20 under the NHAMCS-calibrated scenario (172.88 to 207.52; $p < 0.001$), and 201.44 under controlled surge (182.54 to 220.35; $p < 0.001$). Relative to FRVI, the paired difference was -58.37 under nominal demand (DC-AWRVI lower; -67.21 to -49.52), 14.67 under the NHAMCS-calibrated scenario (DC-AWRVI higher; 6.86 to 22.47), and zero under controlled surge, where the two policies coincide.

The results show a cost of distributional protection and not a demonstrated operational benefit. Relative to NVI, DC-AWRVI increased the mean weighted waiting cost by 0.19% under nominal demand, 1.39% under the NHAMCS-calibrated scenario (190.20 on a base of 13,652.78) and by 1.25% under controlled surge, and it did not improve tail waiting-cost risk or aggregate queue lengths. The method is therefore assessed on its robustness–efficiency trade-off and not on an assumed improvement in every performance indicator.

The two measures capture different aspects of scheduling performance. Mean queue length counts all waiting patients equally, whereas the priority-weighted cost also accounts for clinical urgency through the class weights $\alpha = (5, 4, 3, 2, 1)$. A policy can therefore have a similar aggregate queue while leaving a different mix of waiting patients: additional waiting among high-weight (more urgent) patients raises the weighted cost even when the aggregate queue is unchanged. Under changing demand, DC-AWRVI selects policies from the precomputed Wasserstein-robust library, and a change of the selected policy can shift which urgency classes start service first. The aggregate measures in Table 2 do not, however, identify which scheduling changes produced the observed differences; doing so would require class-level waiting and service-start statistics, which are not reported here. The link between queue length and weighted cost is therefore offered as a possible explanation and not as a result established by the simulations.

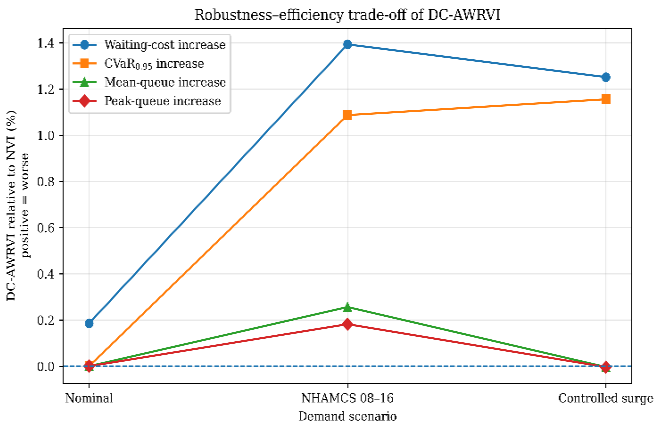


Figure 6. Robustness–efficiency trade-off of DC-AWRVI relative to NVI (positive values indicate a worse outcome for DC-AWRVI)

Note: NVI: Nominal Value Iteration, DC-AWRVI: Data-Calibrated Adaptive Wasserstein Robust Value Iteration.

As shown in Figure 6, the results demonstrate that increasing distributional protection was not cost-free. DC-AWRVI stayed close to nominal-policy performance under stable demand (premium 0.19%) and increased the level of distributional protection under congestion-relevant demand shifts, but this protection was accompanied by a higher mean weighted waiting cost and no improvement in tail risk, so it is not evidence of more reliable operational performance. Accordingly, the contribution of DC-AWRVI is best interpreted as a data-calibrated mechanism for adapting Wasserstein robustness to transition-distribution shifts, rather than as a policy that uniformly dominates nominal scheduling, that is, one that outperforms it across all performance measures (mean waiting cost, queue measures, and tail risk) and across all demand conditions (arrival rates and surge scenarios) examined.

Overall, the proposed DC-AWRVI framework stays close to nominal behavior under stable demand and increases distributional protection when congestion-relevant shifts occur. The resulting performance shows a cost premium for this protection, indicating that the contribution of the method lies in the data-driven adaptation of the Wasserstein ambiguity radius within the robust MDP formulation and not in a claim of uniformly improved operational outcomes.

3.4 Sensitivity to the arrival rate

The preceding results are for one nominal rate, one calibrated hourly profile, and one controlled surge. To examine how the adaptive mechanism responds to the arrival rate itself, the arrival rate was held constant at $\lambda = 7$, 8, 9, 10, 11, and 12 patients per hour (1.0 to 1.7 times the nominal rate $\lambda_0 = 7$) over the eight-hour horizon. The lower part of this range follows [4], who tested 7 to 10 patients per hour with 7 as the normal rate; the range is extended here to 12. For each rate, the adaptive radius of Eq. (35) was computed and mapped to its nearest library node (Eq. (39)), and NVI, FRVI, and DC-AWRVI were simulated with 2,000 replications and common random numbers, using the policies of Section 3.3 ($K_{DP} = 4$). FRVI used the fixed node 0.00878870. Because independent random-number streams were used, the NVI costs at $\lambda = 7$ and 10 (4,843 and 16,480) differ from those of the nominal and controlled-surge scenarios of Section 3.3 (4,709 and 16,100) by amounts within sampling error (standardized differences 1.0 and 1.7). Because three physicians with $\mu = 2.5$ patients per hour can serve at most 7.5 patients per hour, every rate above 7 overloads the system and the queue grows during the shift under any policy. The analysis therefore concerns how the policies differ under overload and not whether they can prevent it. $CVaR_{0.95}$ is the mean cost over the replications at or above the empirical 95th percentile.

Table 3 and Figure 7(a) show how the adaptive radius and the selected policy respond to the arrival rate. At the nominal rate the radius is at the baseline value (0.00126754). It then rises to 0.00347 at $\lambda = 8$, 0.00693 at $\lambda = 9$, 0.01039 at $\lambda = 10$, 0.01384 at $\lambda = 11$, and reaches the upper bound of 0.01528771 at $\lambda = 12$. The selected node follows in steps because of the nearest-node rule, and the nodes selected for $\lambda = 7$ and 8 coincide. The policy stored at each node differs from the NVI policy in 0.12%, 0.40%, 0.40%, 0.57%, and 0.62% of the 12,500 policy states for the five nodes in increasing order of radius, so even the largest radius changes the action in fewer than 1% of states.

Relative to NVI, the mean cost of DC-AWRVI was higher at every rate: by +0.28%, +0.18%, +1.22%, +1.26%, +1.37%, and +3.61% at $\lambda = 7$ to 12, with confidence intervals that exclude zero in each case (Table 3). The premium was small at 7 and 8 patients per hour, where the selected node is the baseline node, rose to about 1.2–1.4% at 9 to 11, and was largest, 3.61%, at $\lambda = 12$, where the radius is at its upper bound. FRVI, which uses a fixed intermediate radius, had a premium of 1.15% to 1.45% across the rates. DC-AWRVI therefore had

a lower cost than FRVI only at the two lowest rates (by 57 and 79 cost units at λ = 7 and 8), where the adaptive radius is smaller than the fixed one; the two policies coincided at 9 and 10; and at λ = 12 DC-AWRVI cost 560 units (95% confidence interval 514 to 605) more than FRVI. No tested arrival rate showed a cost advantage of DC-AWRVI over NVI. At λ = 12, the upper-tail cost also deteriorated ($CVaR_{0.95}$ 38,666.50 for higher (19.85 against 19.49 patients), the mean number of overflow patients was higher (6.30 against 5.49), and physician utilization was lower (94.5% against 96.3%).

Table 3. Adaptive radius, selected node, and cost relative to NVI (paired differences, percent of the NVI cost) as a function of the arrival rate λ (patients per hour); 2,000 replications. "States" is the percentage of the 12,500 policy states in which the action at the selected node differs from the NVI action

λ	Radius	Node	NVI Cost	FRVI vs. NVI (%) [95% CI]	DC-AWRVI vs. NVI (%) [95% CI]	States (%)
7	0.00127	0.00127	4,843	+1.45 [1.24, 1.67]	+0.28 [0.16, 0.40]	0.12
8	0.00346	0.00127	8,094	+1.15 [1.02, 1.29]	+0.18 [0.12, 0.23]	0.12
9	0.00693	0.00636	12,124	+1.22 [1.09, 1.36]	+1.22 [1.09, 1.36]	0.40
10	0.01038	0.00879	16,480	+1.26 [1.15, 1.38]	+1.26 [1.15, 1.38]	0.40
11	0.01384	0.01259	20,769	+1.23 [1.14, 1.33]	+1.37 [1.26, 1.48]	0.57
12	0.01529	0.01529	24,771	+1.35 [1.26, 1.43]	+3.61 [3.42, 3.80]	0.62

Note: NVI: Nominal Value Iteration, FRVI: Fixed-Radius Robust Value Iteration, DC-AWRVI: Data-Calibrated Adaptive Wasserstein Robust Value Iteration.

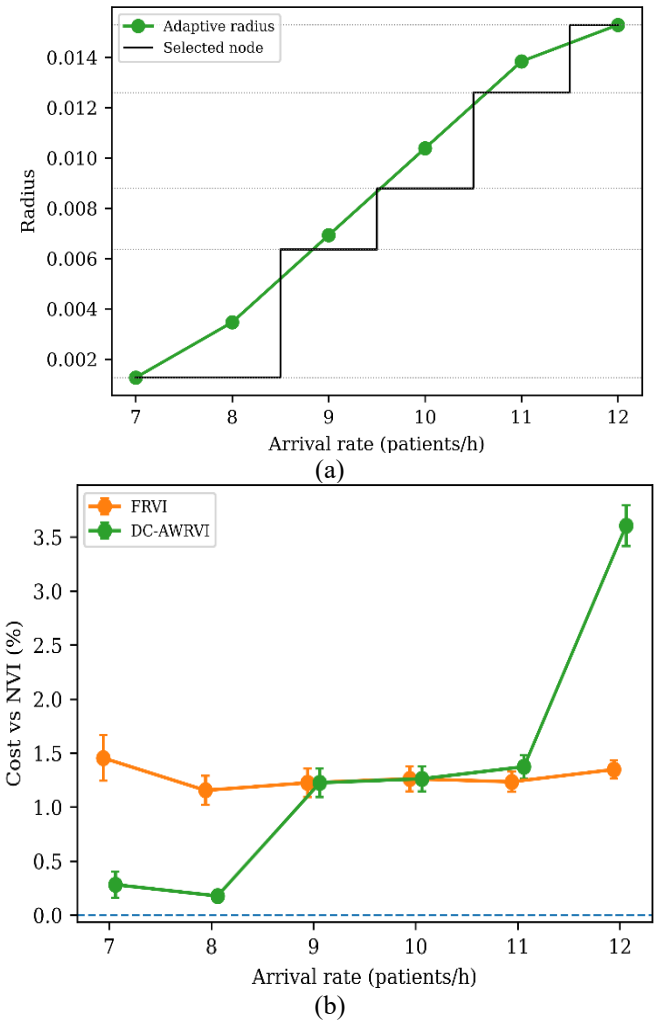


Figure 7. Response to the arrival rate: (a) adaptive radius and selected library node (dotted lines are the library nodes); (b) cost of FRVI and DC-AWRVI relative to NVI with 95% confidence intervals

Note: NVI: Nominal Value Iteration, FRVI: Fixed-Radius Robust Value Iteration, DC-AWRVI: Data-Calibrated Adaptive Wasserstein Robust Value Iteration.

Transient surges were examined with three profiles in which the rate is 7 patients per hour except for a pulse that starts at the beginning of the third hour: a one-hour pulse to 12 patients per hour, a two-hour pulse to 12, and a two-hour pulse to 10 (2,000 replications, same seeds for all policies). DC-AWRVI uses the node selected by the current rate, so it switches to a larger-radius node during the pulse and returns to the baseline node afterwards, whereas FRVI keeps its fixed node. Table 4 and Figure 8 give the results. The pulse of the same size and duration as the sweep rate at 12 per hour raised the cost of DC-AWRVI above that of NVI by 1.65% (one hour) and 2.64% (two hours), against 1.38% and 1.32% for FRVI; the two-hour pulse to 10 per hour, whose selected node is the one used by FRVI, cost 0.55% more than NVI under DC-AWRVI and 1.21% more under FRVI. The peak and final queues did not differ between the policies by more than 0.27 patients, so the transient surges did not show a benefit of switching to a larger radius for congestion either; the switch raised the cost and, for the pulses to 12 per hour, the tail of the cost distribution ($CVaR_{0.95}$ 1.20% and 1.11% above NVI).

In summary, the adaptive mechanism responds to the arrival rate as designed: the radius is held at the baseline at or below the nominal rate, increases with the rate, and reaches its upper bound at 12 patients per hour, and the selected policy changes accordingly. The operational effect, however, is a cost: the larger the radius, the larger the premium relative to NVI, and the largest premium occurs where the radius is largest. The advantage of the adaptive radius over a fixed radius appears only at low demand, where it avoids unnecessary protection. Because overload is set by the capacity of three physicians, the growth of the queue under high arrival rates is dominated by capacity and not by the scheduling policy. Transient surges gave the same picture: switching to a larger radius for a short pulse did not reduce cost or congestion. The effect of physician capacity is examined in Section 3.5.

The radius of Eq. (35) uses the 95th percentile of the Wasserstein-1 shifts over the state–action pairs. To examine this choice, the radius was recomputed with the 90th and the 99th percentile for constant arrival rates of 7 to 12 patients per hour and for the NHAMCS-calibrated hourly rates, with the same baseline and maximum radius and the same five library nodes, and the policies were simulated with 2,000 replications and the seeds of Section 3.4. The percentile has little effect on the radius: at arrival rates of 8 to 11 patients per hour, the radius from the 99th percentile exceeded that from the 90th by 0.35% to 0.52%, and by 0.38% to 0.50% for the NHAMCS hourly rates; at 12 patients per hour all three percentiles exceed the maximum radius and are clipped to it. In every case the three percentiles selected the same library node, so the policies, and therefore the costs, were identical: the cost of

DC-AWRVI relative to NVI was 0.28%, 0.18%, 1.22%, 1.26%, 1.37%, and 3.61% at 7 to 12 patients per hour and 1.41% under the NHAMCS profile, for each of the three percentiles. This insensitivity reflects two features: the shifts are concentrated across the state–action pairs, so that the percentile moves the radius very little, and the library has only five nodes, so that small differences in radius do not change the selected policy. The result shows that the selected policies do not depend on the choice among these three percentiles; it does not show that 95 is the best choice, and a finer library or different bounds could make the choice matter more.

Table 4. Transient surges: baseline rate 7 patients per hour with a pulse from the beginning of the third hour (2,000 replications; percentages relative to NVI)

Surge profile	NVI Cost	FRVI vs. NVI (%) [95% CI]	DC-AWRVI vs. NVI (%) [95% CI]	NVI Peak Queue	DC Peak Queue	CVaR DC vs. NVI (%)
Pulse to 12/h, 1 h	7,688	+1.38 [1.20, 1.55]	+1.65 [1.43, 1.87]	14.67	14.78	+1.20
Pulse to 12/h, 2 h	10,296	+1.32 [1.18, 1.45]	+2.64 [2.41, 2.86]	18.82	19.09	+1.11
Pulse to 10/h, 2 h	7,946	+1.21 [1.06, 1.35]	+0.55 [0.44, 0.65]	15.33	15.33	+0.58

Note: NVI: Nominal Value Iteration, FRVI: Fixed-Radius Robust Value Iteration, DC-AWRVI: Data-Calibrated Adaptive Wasserstein Robust Value Iteration.

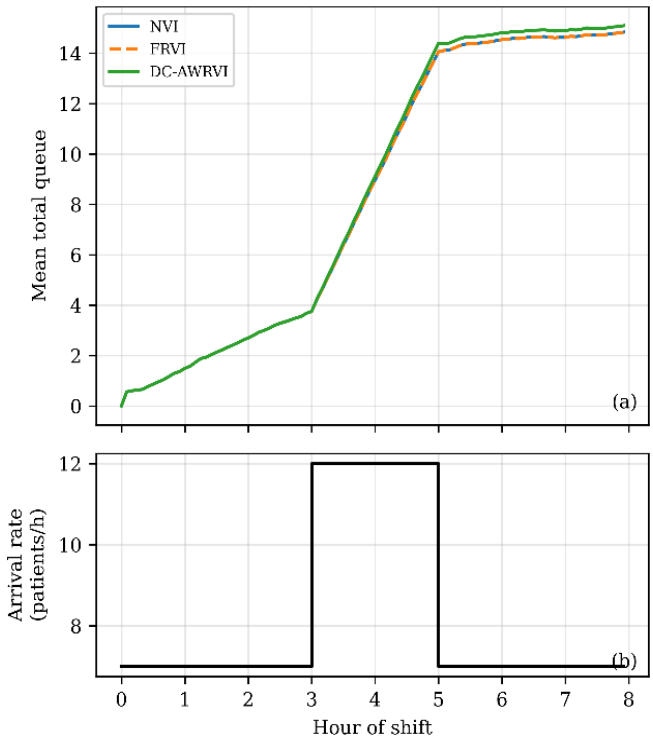


Figure 8. Mean total queue over the shift under the two-hour pulse to 12 patients per hour: (a) NVI, FRVI, and DC-AWRVI; (b) arrival rate

Note: NVI: Nominal Value Iteration, FRVI: Fixed-Radius Robust Value Iteration, DC-AWRVI: Data-Calibrated Adaptive Wasserstein Robust Value Iteration.

3.5 Sensitivity to the queue-state bound and physician staffing

The size of the policy state space depends on the queue-state bound (K_{DP}): 972 states for $K_{DP} = 2$, 4,096 for $K_{DP} = 3$, 12,500 for $K_{DP} = 4$, and 31,104 for $K_{DP} = 5$ at the baseline capacity of three physicians. The verification experiments of Sections 3.1 and 3.2 used $K_{DP} = 3$, and the operational comparison of Sections 3.3 and 3.4 used $K_{DP} = 4$. To examine the effect of the bound, the nominal policy and the five robust policies of the library were recomputed for $K_{DP} = 3$, 4, and 5 at the same radius nodes (the nodes were not recalibrated for each bound), and their decisions were compared on the 4,096 states common to the three bounds, that is, the states with at most three waiting patients in every class. The decisions of DC-AWRVI differed in 2.56% to 3.10% of these states between $K_{DP} = 3$ and 4, depending on the radius node, in 3.59% to 4.27% between $K_{DP} = 3$ and 5, and in 1.03% to 1.17% between $K_{DP} = 4$ and 5; for NVI, the figures were 2.47%, 3.34%, and 0.88%. Over all 12,500 states of the $K_{DP} = 4$ table, the $K_{DP} = 4$ and $K_{DP} = 5$ decisions differed in 1.53% to 1.71% of the states for the robust policies (1.55% for NVI). The decisions were therefore fairly stable across the tested bounds, and the differences were smaller between the larger bounds, but this does not establish equivalence under severe congestion, where clipping matters most and which the common states do not include.

Componentwise clipping assigns every queue length above the bound to the boundary value, so distinct high-congestion states can become indistinguishable to the scheduling policy, which may affect how physician service is allocated among competing triage classes. The larger simulation queue limit does not remove the information lost by clipping in the policy state. The effect of truncation under more extreme congestion therefore remains a limitation of the finite-state approximation.

A separate experiment varied the number of physicians from three to seven and the per-class queue bound from three to seven, with 1,000 replications per configuration, under a strict triage-priority policy and the hourly NHAMCS-based demand profile. It describes how the simulated system responds to physician capacity and to the queue bound; it does not itself evaluate NVI, FRVI, or DC-AWRVI. This is done below for four and five physicians. Table 5 summarizes three of the five queue bounds. Increasing the number of physicians lowered cost, queue length, and overflow substantially, with diminishing returns: at $K_{DP} = 7$ the mean paired cost reduction was 2,416.08 from three to four physicians (95% confidence interval 2,313.38 to 2,518.78), 820.19 from four to five, 256.48 from five to six, and 67.35 from six to seven (57.89 to 76.81). Increasing the queue bound raised the recorded cost and mean queue and lowered overflow, because fewer patients are lost to truncation and more are retained and counted; at three physicians, the mean number of overflow patients fell from 5.19 at $K = 3$ to 1.43 at $K = 7$. Utilization was not recorded in this experiment. The behavior of the adaptive policy with more physicians was examined for four and five physicians with $K_{DP} = 4$: the policy library was recomputed at the same five radius nodes, the radius rule and its mapping to library nodes were kept as calibrated for three physicians, and NVI, FRVI, and DC-AWRVI were simulated with 2,000 replications at constant arrival rates of 7, 10, and 12 patients per hour and under the NHAMCS-calibrated profile. With four physicians (capacity 10 patients per hour at a service rate of 2.5) the cost of DC-AWRVI relative to NVI was 0.00% [0.00,

0.00] at 7, 1.19% [1.06, 1.31] at 10, 6.31% [6.02, 6.59] at 12 patients per hour, and 4.58% [4.27, 4.88] under the NHAMCS profile, with $CVaR_{0.95}$ 3.46% and 4.18% higher at 12 patients per hour and under the NHAMCS profile. With five physicians (capacity 12.5), the three policies coincided or nearly so: the difference between DC-AWRVI and NVI was 0.04% [0.01, 0.07] at 12 patients per hour and nil in the other scenarios. Thus, DC-AWRVI had a higher cost than NVI whenever the policies differed; the premium was larger with four than with three physicians at the radii that the three-physician calibration assigns, and with ample capacity the robust policies reduced to the nominal policy. Because the radius rule was not recalibrated for four or five physicians, these results describe the library under other staffing levels and not a recalibrated system. The effects of the service rate and of overdispersed arrivals are examined next.

Because the policies were computed for a service rate of μ = 2.5 and Poisson arrivals, they were also evaluated when the true system deviated from these assumptions, with the policies themselves unchanged: service rates μ = 1.5, 2.0, 2.5, and 3.0, and overdispersed arrivals (negative binomial with the same mean and a variance of two and three times the mean), using 2,000 replications and the same random-number seeds as in Section 3.3 (Table 6). The service rate has a large effect on congestion: under the NHAMCS profile, the mean queue under NVI was 18.30, 14.44, 10.99, and 8.09 patients at μ = 1.5, 2.0, 2.5, and 3.0, with physician utilization of 94%, 93%, 92%, and 89%. Overdispersed arrivals raised the nominal-demand cost of NVI from 4,709 to 6,204 and 7,361. Despite these deviations, DC-AWRVI had a higher mean cost than NVI in every setting: by 0.02% to 0.22% under nominal demand and 1.20% to 1.55% under the NHAMCS profile, and its $CVaR_{0.95}$ was equal or higher (Table 6). The robust policies thus did not protect against the deviations tested here. This result concerns only these deviations and fixed policies; it does not exclude benefits under other forms of model error, and the policies were not recomputed for each service rate.

Table 5. Effect of the number of physicians *D* and the queue bound *K* on cumulative cost, mean queue, overflow patients, and completions (strict triage-priority policy; 1,000 replications; K = 4 and 6 omitted)

D	K	Mean Cost (SD)	Mean Queue	Overflow	Completions
3	3	2,161 (1,139)	2.45	5.19	45.3
3	5	3,092 (1,872)	3.33	2.72	46.1
3	7	3,590 (2,436)	3.79	1.43	46.2
4	3	842 (692)	1.29	1.57	50.3
4	5	1,082 (1,050)	1.52	0.50	50.9
4	7	1,174 (1,251)	1.61	0.17	51.0
5	3	291 (352)	0.82	0.43	52.1
5	5	341 (477)	0.87	0.09	52.3
5	7	354 (529)	0.88	0.02	52.4
6	3	86 (148)	0.65	0.11	52.5
6	5	96 (180)	0.66	0.01	52.6
6	7	97 (196)	0.66	0.00	52.6
7	3	28 (69)	0.60	0.05	52.6
7	5	30 (82)	0.60	0.00	52.6
7	7	30 (85)	0.60	0.00	52.6

Table 6. Policies evaluated under deviations from the model assumptions (service rate; overdispersed arrivals); 2,000 replications; percentages relative to NVI under the same deviation

Deviation	Scenario	NVI Cost	NVI Mean Queue	NVI Utiliz. (%)	DC-AWRVI vs. NVI (%) [95% CI]	CVaR vs. NVI (%)
μ = 1.5	Nominal	13,301	10.50	93.5	+0.08 [0.05, 0.11]	+0.07
μ = 1.5	NHAMCS 08-16	23,215	18.30	94.1	+1.20 [1.13, 1.26]	+0.82
μ = 2.0	Nominal	8,309	6.83	90.6	+0.21 [0.14, 0.28]	+0.06
μ = 2.0	NHAMCS 08-16	18,254	14.44	93.0	+1.29 [1.20, 1.38]	+0.99
μ = 2.5	Nominal	4,709	4.19	85.3	+0.19 [0.10, 0.27]	+0.00
μ = 2.5	NHAMCS 08-16	13,653	10.99	91.5	+1.39 [1.27, 1.52]	+1.09
μ = 3.0	Nominal	2,549	2.57	78.1	+0.22 [0.14, 0.31]	+0.09
μ = 3.0	NHAMCS 08-16	9,720	8.09	89.4	+1.55 [1.38, 1.71]	+1.14
Variance = 2 × mean	Nominal	6,204	5.26	82.8	+0.02 [0.01, 0.04]	+0.01
Variance = 2 × mean	NHAMCS 08-16	14,521	11.67	89.8	+1.48 [1.34, 1.62]	+1.00
Variance = 3 × mean	Nominal	7,361	6.13	80.8	+0.02 [0.01, 0.03]	+0.00
Variance = 3 × mean	NHAMCS 08-16	15,407	12.35	88.1	+1.36 [1.22, 1.50]	+0.57

Note: NVI: Nominal Value Iteration, DC-AWRVI: Data-Calibrated Adaptive Wasserstein Robust Value Iteration.

3.6 Summary of the results by demand scenario

This subsection collects the results of Sections 3.3 to 3.5 for each demand scenario. The three scenarios differ in the arrival process and in how close the demand is to physician capacity: with three physicians and μ = 2.5 patients per hour, capacity is 7.5 patients per hour. Table 7 gives the main comparison with NVI (2,000 replications), and the paragraphs that follow add the evidence from the arrival-rate sweep and from the deviations of Table 6.

Nominal demand. The arrival rate equals the nominal rate (7 patients per hour, 93% of capacity), so the adaptive radius stays at the baseline value and DC-AWRVI selects the node 0.00126754. Its cost was 4,718.07 against 4,709.34 for NVI (0.19%; 95% confidence interval 4.91 to 12.54), with identical $CVaR_{0.95}$ and identical aggregate queue measures. The policy

is therefore practically the nominal policy, which is the intended behavior under stable demand. The premium stayed between 0.08% and 0.22% when the service rate was varied from 1.5 to 3.0 and was 0.02% with overdispersed arrivals.

NHAMCS-calibrated demand. The hourly rates run from 6.94 to 10.73 patients per hour, so the demand is above capacity for most of the eight hours and the adaptive radius changes hour by hour, selecting the nodes 0.00126754, 0.00636465, 0.00878870, and 0.01259272. The cost was 13,842.99 against 13,652.78 for NVI (1.39%; 172.88 to 207.52), $CVaR_{0.95}$ was 1.09% higher, and the mean queue was 0.26% higher. This is the scenario with the clearest premium and the one in which the policy switches most often. The premium was 1.20% to 1.55% across service rates of 1.5 to 3.0 and 1.36% to 1.48% with overdispersed arrivals, so the deviations tested did not make DC-AWRVI preferable.

Controlled surge. A constant rate of 10 patients per hour (1.43 times the nominal rate, 133% of capacity) selects the node 0.00878870 throughout, which is also the node used by FRVI, so the two policies coincide. The cost was 16,301.47 against 16,100.02 for NVI (1.25%; 182.54 to 220.35) and $CVaR_{0.95}$ was 1.16% higher, with no difference in aggregate queue measures. The arrival-rate sweep of Section 3.4 shows that the premium grows if the surge is stronger (3.61% at 12 patients per hour, where the radius reaches its upper bound), and the deviations of Table 6 gave premiums of 1.18% to 1.36% (service rate) and 1.18% to 1.25% (overdispersed arrivals).

Table 7. Summary of the DC-AWRVI results by demand scenario (2,000 replications; percentages relative to NVI)

Scenario	NVI Cost (SD)	DC-AWRVI vs. NVI (%) [95% CI]	CVaR vs. NVI (%)	Mean Queue NVI/DC	Utilization (%) NVI/DC
Nominal (7 patients/h, constant)	4,709 (4,004)	+0.19 [0.10, 0.27]	+0.00	4.19 / 4.19	85.3 / 85.3
NHAMCS 08–16 (hourly rates 6.94–10.73)	13,653 (6,139)	+1.39 [1.27, 1.52]	+1.09	10.99 / 11.02	91.5 / 91.4
Controlled surge (10 patients/h, constant)	16,100 (6,910)	+1.25 [1.13, 1.37]	+1.16	12.80 / 12.80	94.7 / 94.7

Note: NVI: Nominal Value Iteration, DC-AWRVI: Data-Calibrated Adaptive Wasserstein Robust Value Iteration.

Across the three scenarios, the cost premium of DC-AWRVI relative to NVI increased with the demand level and with the radius selected, from about 0.2% under nominal demand to 1.3–1.4% under the calibrated and surge scenarios, and no scenario showed an advantage of DC-AWRVI in cost, tail risk, or aggregate queue length.

4. CONCLUSIONS

This study developed a data-calibrated adaptive Wasserstein robust decision framework for dynamic ED scheduling under time-varying demand. The model integrates a finite-state MDP, Wasserstein distributional robustness, and an adaptive ambiguity-radius mechanism calibrated from empirically derived transition-kernel shifts. The final formulation separates the role of the Bellman residual as a convergence criterion from the role of the Wasserstein distance as an indicator of congestion-relevant distribution shift. In the sense of automated systems, the framework closes a loop from the observed arrival rate to the robust protection level, the stored policy, and the patient scheduling decision, and the system sets the level of protection from data instead of relying on a manually chosen radius; its evaluation, however, was by simulation only, with a known arrival-rate input.

The numerical experiments confirmed that the nominal and robust Bellman solvers were implemented consistently and that Wasserstein robustification altered scheduling decisions only in a limited subset of states. The data-calibrated radius ranged from $\rho_{\text{base}} = 0.0012675$to $\rho_{\max} = 0.0152877$, with higher values activated when the hourly arrival rate exceeded the nominal demand level. This allowed the robustness level to respond directly to empirically observed upward-demand shifts rather than relying on a manually fixed ambiguity radius.

Operationally, DC-AWRVI reproduced nominal-policy performance under stable demand, while under the NHAMCS-calibrated and controlled-surge scenarios, it did not achieve the lowest mean waiting cost: based on 2,000 replications per policy–scenario combination, its mean waiting cost exceeded that of NVI by 1.39% under the NHAMCS-calibrated profile and by 1.25% under the controlled surge, and its aggregate queue measures were equal or marginally higher. Tail-risk analysis did not show an improvement in $CVaR_{0.95}$. Therefore, the proposed method should not be interpreted as uniformly outperforming nominal scheduling in expected or tail waiting cost. Instead, the results show the cost of the additional distributional protection; whether it translates into operational robustness depends on the demand scenario and the performance criterion, and was not established here. The arrival-rate sensitivity analysis (Section 3.4) showed that the adaptive radius followed the arrival rate and reached its upper bound at 12 patients per hour, that the cost premium relative to NVI increased with the radius to 3.61% at that rate, and that DC-AWRVI had a lower cost than FRVI only at the lowest arrival rates. At the nominal rate the policies were practically identical; [4] likewise found little difference between deep reinforcement learning and the best dispatching rule at 7 patients per hour, although their methods, performance measure, and horizon differ from those used here.

The study has several limitations. The Poisson arrival and exponential service-time assumptions give a tractable Markovian model but do not capture clustering, overdispersion, or workload-dependent service. The service rate and the number of physicians are experimental assumptions, and the model was not calibrated to a specific hospital or ED size. The 95th-percentile radius calibration is a convention and, unlike radii derived from measure-concentration bounds [18], it carries no finite-sample guarantee, and the comparison with the 90th and 99th percentiles (Section 3.4) is limited by the coarse library of five radius nodes, which maps the three percentiles to the same policies. Componentwise clipping of the policy state means that distinct high-congestion states may be treated alike, and the decisions were compared across queue bounds of three to five only at fixed radius nodes, with the radius rule not recalibrated for each bound or for other staffing levels. The policy library returns only the nearest stored policy for an intermediate radius, and the radius input is updated hourly

while decisions are made every five minutes. Finally, the framework is a simulation-based decision-support model that has not been deployed or clinically validated. Because the model is single-stage, only first physician service is represented, and later diagnostic, treatment, and disposition stages are not modeled.

The main contribution of the study is the explicit linkage between empirically calibrated demand shifts, Wasserstein-based transition uncertainty, adaptive ambiguity-radius selection, and robust scheduling policy selection. This linkage provides a reproducible mechanism for adjusting the level of distributional protection according to changes in the operating environment. Future work may compare alternative calibration percentiles with a finer policy library, calibrate the radius by bootstrapping or cross-validation as recommended for data-driven Wasserstein ambiguity sets [18], and extend the framework to multi-hospital data, richer state representations, non-stationary service dynamics, and online estimation of transition distributions.

REFERENCES

[1] Hoot, N.R., Aronsky, D. (2008). Systematic review of emergency department crowding: Causes, effects, and solutions. Annals of Emergency Medicine, 52(2): 126-136. https://doi.org/10.1016/j.annemergmed.2008.03.014

[2] Saghafian, S., Austin, G., Traub, S.J. (2015). Operations research/management contributions to emergency department patient flow optimization: Review and research prospects. IIE Transactions on Healthcare Systems Engineering, 5(2): 101-123. https://doi.org/10.1080/19488300.2015.1017676

[3] Elalouf, A., Wachtel, G. (2021). Queueing problems in emergency departments: A review of practical approaches and research methodologies. Operations Research Forum, 3: 2. https://doi.org/10.1007/s43069-021-00114-8

[4] Lee, S., Lee, Y.H. (2020). Improving emergency department efficiency by patient scheduling using deep reinforcement learning. Healthcare, 8(2): 77. https://doi.org/10.3390/healthcare8020077

[5] De Queiroz, T.A., Iori, M., Kramer, A., Kuo, Y.H. (2023). Dynamic scheduling of patients in emergency departments. European Journal of Operational Research, 310(1): 100-116. https://doi.org/10.1016/j.ejor.2023.03.004

[6] Daknou, A., Zgaya, H., Hammadi, S., Hubert, H. (2010). A dynamic patient scheduling at the emergency department in hospitals. In 2010 IEEE Workshop on Health Care Management (WHCM), Venice, Italy, pp. 1-6. https://doi.org/10.1109/WHCM.2010.5441267

[7] Luscombe, R., Kozan, E. (2016). Dynamic resource allocation to improve emergency department efficiency in real time. European Journal of Operational Research, 255(2): 593-603. https://doi.org/10.1016/j.ejor.2016.05.039

[8] Ajmi, F., Ajmi, F., Othman, S.B., et al. (2024). Real time patient scheduling orchestration for improving key performance indicators in a hospital emergency department. Journal of Computational Science, 82: 102422. https://doi.org/10.1016/j.jocs.2024.102422

[9] Daldoul, D., Nouaouri, I., Bouchriha, H., Allaoui, H. (2017). Scheduling patients in emergency department: A case study. In 2017 IEEE International Conference on Industrial Engineering and Engineering Management (IEEM), Singapore, pp. 870-874. https://doi.org/10.1109/IEEM.2017.8290016

[10] Hamasha, M.M., Rumbe, G. (2017). Determining optimal policy for emergency department using Markov decision process. World Journal of Engineering, 14(5): 467-472. https://doi.org/10.1108/WJE-12-2016-0148

[11] Patrick, J., Puterman, M.L., Queyranne, M. (2008). Dynamic multipriority patient scheduling for a diagnostic resource. Operations Research, 56(6): 1507-1525. https://doi.org/10.1287/opre.1080.0590

[12] Iyengar, G.N. (2005). Robust dynamic programming. Mathematics of Operations Research, 30(2): 257-280. https://doi.org/10.1287/moor.1040.0129

[13] Nilim, A., El Ghaoui, L. (2005). Robust control of Markov decision processes with uncertain transition matrices. Operations Research, 53(5): 780-798. https://doi.org/10.1287/opre.1050.0216

[14] Xu, H., Mannor, S. (2012). Distributionally robust Markov decision processes. Mathematics of Operations Research, 37(2): 288-300. https://doi.org/10.1287/moor.1120.0540

[15] Bäuerle, N., Glauner, A. (2022). Distributionally robust Markov decision processes and their connection to risk measures. Mathematics of Operations Research, 47(3): 1757-1780. https://doi.org/10.1287/moor.2021.1187

[16] Yang, I. (2017). A convex optimization approach to distributionally robust Markov decision processes with Wasserstein distance. IEEE Control Systems Letters, 1(1): 164-169. https://doi.org/10.1109/LCSYS.2017.2711553

[17] Villani, C. (2009). Optimal Transport: Old and New. Springer, Berlin. https://doi.org/10.1007/978-3-540-71050-9

[18] Mohajerin Esfahani, P., Kuhn, D. (2018). Data-driven distributionally robust optimization using the Wasserstein metric: Performance guarantees and tractable reformulations. Mathematical Programming, 171: 115-166. https://doi.org/10.1007/s10107-017-1172-1

[19] Clement, J.G., Kroer, C. (2021). First-order methods for Wasserstein distributionally robust MDP. In International Conference on Machine Learning, pp. 2010-2019.

[20] Neufeld, A., Sester, J. (2024). Robust Q-learning algorithm for Markov decision processes under Wasserstein uncertainty. Automatica, 168: 111825. https://doi.org/10.1016/j.automatica.2024.111825

[21] Wang, Y., Zhang, Y., Tang, J. (2024). Wasserstein distributionally robust surgery scheduling with elective and emergency patients. European Journal of Operational Research, 314(2): 509-522. https://doi.org/10.1016/j.ejor.2023.10.026

[22] National Center for Health Statistics. (2024). 2021 NHAMCS Questionnaires, Datasets, and Documentation. Centers for Disease Control and Prevention. https://www.cdc.gov/nchs/nhamcs/documentation/about-the-data-2021.html, accessed on Aug. 28, 2026.

[23] Bellman, R.E. (1957). Dynamic Programming. Princeton University Press, Princeton, NJ.

[24] Puterman, M.L. (1994). Markov Decision Processes:

Discrete Stochastic Dynamic Programming. John Wiley & Sons. https://doi.org/10.1002/9780470316887
[25] Delage, E., Ye, Y. (2010). Distributionally robust optimization under moment uncertainty with application to data-driven problems. Operations Research, 58(3): 595-612. https://doi.org/10.1287/opre.1090.0741
[26] Gao, R., Kleywegt, A. (2023). Distributionally robust stochastic optimization with Wasserstein distance. Mathematics of Operations Research, 48(2): 603-655. https://doi.org/10.1287/moor.2022.1275
[27] Kırış, Ş., Yüzügüllü, N., Ergün, N., Çevik, A.A. (2010). A knowledge-based scheduling system for emergency departments. Knowledge-Based Systems, 23(8): 890-900. https://doi.org/10.1016/j.knosys.2010.06.005
[28] Green, L. (2006). Queueing analysis in healthcare. In Patient Flow: Reducing Delay in Healthcare Delivery, Springer, New York, pp. 281-307. https://doi.org/10.1007/978-0-387-33636-7_10
[29] Duque, D., Morton, D.P. (2020). Distributionally robust stochastic dual dynamic programming. SIAM Journal on Optimization, 30(4): 2841-2865. https://doi.org/10.1137/19M1309602
[30] Yang, I. (2020). Wasserstein distributionally robust stochastic control: A data-driven approach. IEEE Transactions on Automatic Control, 66(8): 3863-3870. https://doi.org/10.1109/TAC.2020.3030884

NOMENCLATURE

$\mathcal{A}(s)$	feasible action set in state s
$A_{i,t}$	number of arrivals in urgency class $i$during epoch t
a_t	patient-scheduling action vector at epoch t
B_t	number of physicians engaged in service after allocation
$c(s, a)$	one-stage priority-weighted waiting cost
C_t	number of service completions during epoch t
d_t	number of available physicians at epoch t
d_t^+	number of physicians available immediately after allocation
$d_K(s^{(r)}, s^{(q)})$	normalized ground metric between states $s^{(r)}$and $s^{(q)}$
D	total number of physicians
e_k	Bellman residual at iteration k
h	hourly demand-condition index
i	urgency-class index
$J^\pi(s)$	expected discounted cumulative cost under policy π
k	value-iteration index
K_{DP}	queue bound used for policy construction
K_{sim}	queue bound used in simulation
$n_{i,t}$	number of patients waiting in urgency class $i$at epoch t
$P(\cdot \mid s, a)$	transition probability kernel
$P_0(\cdot \mid s, a)$	nominal transition kernel
$P_h(\cdot \mid s, a)$	transition kernel under hourly demand condition h
p_i	probability of an arrival belonging to urgency class i
p_c	probability of service completion within one decision interval
$\mathcal{P}_\rho(s, a)$	Wasserstein ambiguity set for state-action pair (s, a)
$Q_{0.95}$	95th-percentile quantile operator
r_h	relative hourly demand factor
$\mathcal{R}$	set of calibrated Wasserstein-radius nodes
s_t	ED system state at decision epoch t
$\mathcal{S}$	finite state space
T	nominal Bellman operator
T_ρ	Wasserstein robust Bellman operator
$V(s)$	value function at state s
$V_k(s)$	value-function estimate at iteration k
$V^*(s)$	optimal value function
$W_1(P, Q)$	Wasserstein-1 distance between distributions $P$and Q
z_h	95th-percentile summary of transition shifts for demand condition h

Greek Symbol

α_i	priority weight for urgency class i
Δ	decision interval used in waiting-cost calculation
Δ_h	decision interval expressed in hours
$\delta_{s^{(q)}}$	Dirac measure concentrated at state $s^{(q)}$
ε	convergence tolerance
η	dual variable in the Wasserstein worst-case expectation
Γ	transport coupling between probability distributions
γ	discount factor
λ	patient arrival rate
λ_0	nominal patient arrival rate
λ_h	hourly patient arrival rate under condition h
μ	physician service rate
$\omega_h(s, a)$	Wasserstein transition shift for (s, a)under condition h
$\Pi(P, Q)$	set of couplings between distributions $P$and Q
π	stationary scheduling policy
π_ρ^*	optimal robust policy for ambiguity radius ρ
ρ	Wasserstein ambiguity radius
ρ_{base}	baseline ambiguity radius
$\rho_{\max}$	maximum calibrated ambiguity radius
ρ_h	adaptive ambiguity radius for condition h
$\tilde{\rho}_h$	nearest calibrated radius selected for condition h
$\phi(s)$	mapping from simulation state to bounded policy state

Subscript

0	nominal operating condition or nominal transition kernel
0.95	95th-percentile or 95% risk level
base	baseline robustness level
DP	dynamic-programming policy space
h	hourly demand condition
i	urgency class, $i = 1, \ldots, 5$
k	value-iteration step
max	maximum calibrated robustness level
ρ	quantity associated with Wasserstein ambiguity radius ρ

Sim	operational simulation space	(m)	m-th calibrated Wasserstein-radius node
t	decision epoch	(q)	q-th discrete next-state support point
$*$	optimal value function or optimal policy	(r)	r-th candidate state in the Wasserstein dual formulation
$+$	post-allocation quantity		