RESEARCH PAPER

Data Driven Two Stage Stochastic Programming for Pharmaceutical Supply Chain Optimization under Demand Uncertainty

A. D. Sarange[1]*, K. Y. Ingale[2], P. V. Patil[1]

[1]*School of Mathematical Sciences, Swami Ramanand Tirth Marathwada University, Nanded-431606, India*
[2]*Netaji Subhash Chandra Bose Arts, Commerce, and Science College, Nanded-431601, India*
corresponding Author
Email id; ajaysarange@gmail.com

ABSTRACT
The COVID-19 pandemic exposed critical vulnerabilities in pharmaceutical supply chains, highlighting the urgent need for robust decision-making frameworks capable of accommodating demand uncertainty. This paper presents a data-driven two-stage stochastic programming model for pharmaceutical distribution optimization under demand uncertainty. Using a synthetic dataset that captures realistic pharmaceutical distribution operations, we develop a mathematical formulation that integrates first-stage inventory allocation decisions with second-stage recourse actions contingent on realized demand scenarios. The model incorporates key real-world constraints including lead time variability, capacity limitations, and service level requirements. Goodness-of-fit tests (Kolmogorov–Smirnov, Akaike Information Criterion) validate the use of negative binomial distributions for demand and lognormal distributions for lead times. We employ Monte Carlo simulation with scenario reduction and conduct extensive computational experiments comparing stochastic and deterministic approaches. Results demonstrate that the stochastic model achieves a 23.6% reduction in expected total costs (95% confidence interval: 21.4%–25.8%) and a 12.0 percentage point improvement in service level fulfilment compared to deterministic planning, with a paired t-test confirming statistical significance ($p < 0.01$). Sensitivity analysis reveals critical thresholds for inventory positioning and the value of incorporating demand uncertainty in pharmaceutical logistics. The proposed framework provides practical decision support for healthcare supply chain managers while contributing to the emerging literature on resilient pharmaceutical supply chains

Keywords: Stochastic programming, pharmaceutical supply chain, demand uncertainty, healthcare logistics, two stage optimization, inventory management.
How to cite this article: Sarange AD, Ingale KY, Patil PV; Data Driven Two Stage Stochastic Programming for Pharmaceutical Supply Chain Optimization under Demand Uncertainty..Int J Drug Deliv Technol. 2026;16(19s): 93-100. DOI: 10.25258/ijddt.16.19s.11
Source of support: Nil.
Conflict of interest: Nil.

INTRODUCTION

The global pharmaceutical supply chain represents a critical infrastructure component of moder healthcare systems, responsible for ensuring timely availability of life-saving medications to patients worldwide. Unlike conventional supply chains, pharmaceutical logistics face unique challenges including stringent regulatory requirements, temperature-sensitive products, limited substitutability, and the inherent unpredictability of healthcare demand patterns. The COVID-19 pandemic starkly illustrated these vulnerabilities, with widespread shortages of essential medications, personal protective equipment, and vaccines exposing the fragility of traditional deterministic planning approaches [1].

Recent years have witnessed growing recognition that pharmaceutical supply chain optimization must account for the fundamental uncertainty characterizing healthcare demand. Patient inflows, disease outbreaks, seasonal variations, and public health emergencies create demand patterns that defy conventional forecasting methods. Moreover, the consequences of supply failures in healthcare contexts extend beyond financial losses to directly impact patient outcomes, making robustness critical design objective [2].

The pharmaceutical supply chain literature has increasingly turned to stochastic programming approaches to address these challenges. However, existing studies often rely on synthetic or limited empirical data, constraining external validity. Furthermore, the integration of lead time variability with demand uncertainty in a unified stochastic framework remains underdeveloped, and comparative analysis of stochastic versus deterministic approaches across realistic problem instances with full cost structures is limited.

This study addresses these gaps by developing a two-stage stochastic programming model grounded in a comprehensive synthetic dataset that captures realistic pharmaceutical distribution operations, incorporating both

**Author for Correspondence: ajaysarange@gmail.com*

demand and lead time uncertainty, and providing extensive computational analysis with practical insights for healthcare supply chain managers. Specifically, we contribute: (i) a data-driven stochastic optimization framework that integrates demand and lead time uncertainty, (ii) a systematic scenario generation methodology with rigorous statistical validation, (iii) empirical quantification of the value of stochastic solutions, and (iv) actionable managerial insights derived from sensitivity analysis.

2 Literature Review

2.1 Stochastic Programming in Supply Chain Optimization

Stochastic programming has emerged as a powerful methodology for decision-making under uncertainty in supply chain contexts. Two-stage stochastic programming, in particular, provides a natural framework for problems where some decisions must be made before uncertainty is resolved, with recourse actions available afterward [3]. The theoretical foundations of stochastic programming are well established in the works of birge2011 and shapiro2014. This structure aligns well with pharmaceutical distribution, where inventory allocation decisions are made based on forecasts, while emergency shipments can be used after demand realization. The ability to model recourse actions explicitly allows for the evaluation of trade-offs between proactive and reactive strategies.

2.2 Healthcare and Pharmaceutical Supply Chain Optimization

Recent research has focused on developing specialized optimization models for healthcare supply chains. Ivanov [1] introduced the concept of supply chain viability, emphasizing the need for adaptive strategies during disruptions. Paul and Chowdhury [2] developed a two-stage stochastic programming approach for medical supply chain network design under pandemic disruptions, highlighting the importance of considering multiple disruption scenarios. The integration of risk measures represents another important development. Conditional Value-at-Risk (CVaR) has emerged as a preferred risk-aversion technique in healthcare supply chain optimization, enabling explicit control over shortage levels while maintaining computational tractability [4].

2.3 Pharmaceutical Supply Chain Specificities

Pharmaceutical supply chains exhibit distinctive characteristics including temperature sensitivity, product expiration, regulatory compliance, and the high cost of stockouts. The pharmaceutical distribution network often features centralized warehousing serving decentralized healthcare facilities, creating natural two-echelon structures [6]. This configuration presents opportunities for consolidation benefits balanced against responsiveness requirements. Moreover, the perishable nature of many pharmaceutical products introduces additional complexity, as inventory must be managed to minimize waste while ensuring availability.

2.4 Data-Driven Optimization and Machine Learning Integration

The increasing availability of healthcare supply chain data has enabled data-driven optimization approaches that combine machine learning for prediction with optimization for decision-making. This integration is particularly valuable in pharmaceutical contexts where demand patterns may be influenced by external factors including disease prevalence, seasonal patterns, and public health interventions [7]. Recent advances in deep learning for demand forecasting [8] have shown promise in improving scenario generation for stochastic optimization models.

2.5 Research Gaps and Contributions

Despite significant advances, several critical gaps remain in the literature: (i) existing studies often rely on synthetic or limited empirical data, constraining external validity; (ii) the integration of lead time variability with demand uncertainty in a unified stochastic framework remains underdeveloped; (iii) comparative analysis of stochastic versus deterministic approaches across realistic problem instances with full cost structures is limited; and (iv) practical guidance for implementing stochastic optimization in pharmaceutical supply chain operations is sparse. This study addresses these gaps by developing a data-driven two-stage stochastic programming model grounded in a comprehensive dataset that captures realistic pharmaceutical distribution operations, incorporating both demand and lead time uncertainty, and providing extensive computational analysis with practical insights.

3 Materials and Methods

3.1 Data Description and Exploratory Analysis

3.1.1 Dataset Overview

This study utilizes a synthetic pharmaceutical supply chain dataset generated to replicate realistic distribution operations. The dataset comprises 15,000 transaction records spanning a 12-month period, capturing orders from three distribution centres to 25 healthcare facilities across four facility types (urban tertiary, urban secondary, rural secondary, and rural primary). The synthetic data generation process incorporated realistic demand patterns, lead time distributions, and cost structures based on industry benchmarks. The data include supply chain operational variables (lead times, distances, costs) and demand-related variables (daily demand, patient counts, seasonality indices). Daily records are aggregated to weekly

periods for computational tractability while preserving seasonal patterns.

3.1.2 Variable Description

Key variables include:

• **Demand:** Daily demand (units/day), product category (antibiotics, cardiovascular, respiratory, analgesics), patient count, seasonality index.

• **Supply chain:** Distribution centre (DC1–DC3), healthcare facility (F01–F25), distance (km), lead time (days), inventory level, safety stock, transportation cost per km, holding cost per unit, stockout cost per unit, waste cost per unit.

• **Time:** Order date, delivery date, month, day of week.

3.1.3 Data Preprocessing

Preprocessing steps: (1) Missing value treatment (< 2%) using median (continuous) or mode (categorical); (2) Outlier detection (IQR method) with retention of extreme values (3.2%) to preserve plausible extreme demand scenarios; (3) Temporal aggregation to weekly level; (4) Distribution fitting: Negative binomial for demand (best fit by AIC and KS test, e.g., for Facility F05, Respiratory: KS=0.034, p = 0.87; AIC=1842 vs Poisson 1912, normal 1956) and lognormal for lead times (DC1–F05: KS=0.027, p = 0.93; AIC=312 vs normal 341). All facility-product combinations had KS p > 0.05, confirming the chosen distributions

3.1.4 Exploratory Analysis

Demand patterns Table 1 shows summary statistics. Urban tertiary hospitals have the highest mean demand (847.3 units/week) but moderate variability (CV=0.31), while rural primary facilities show lower mean (213.7) and highest variability (CV=0.68), underscoring the need for stochastic approaches in rural settings.

Table 1: Weekly Demand Summary Statistics by Facility Type

Facility Type	Mean	Std Dev	CV	Min	Max	Count
Urban Tertiary	847.3	262.7	0.31	312	1324	12
Urban Secondary	523.6	183.3	0.35	187	876	18
Rural Secondary	368.4	184.2	0.50	98	754	24
Rural Primary	213.7	145.3	0.68	42	621	36

Lead time variability Table 2 summarizes lead times. DC1 (Metro Central) has the shortest and most consistent lead times (mean=2.3 days, SD=0.8), while DC3 (Rural) exhibits the longest and most variable (mean=4.7 days, SD=1.6).

Table 2: Lead Time Statistics by Distribution Centre

DC	Mean (days)	Std Dev	P10	P50	P90
DC1 (Metro Central)	2.3	0.8	1	2	4
DC2 (Metro North)	3.1	1.0	2	3	5
DC3 (Rural)	4.7	1.6	3	4	7

Seasonal patterns Demand shows significant seasonal variation (ANOVA $p < 0.001$). Table 3 gives normalized indices (sum to 12). Respiratory medications peak in winter (December 1.47, +47%), antibiotics in summer (June 1.15, +15%).

Table 3: Normalized Seasonal Demand Indices by Product Category

Product Category	Jan	Feb	Mar	Apr	May	Jun
Respiratory	0.85	0.92	0.95	0.98	1.02	1.05
Cardiovascular	1.02	1.01	1.00	0.99	0.98	0.97
Antibiotics	0.95	0.96	0.98	1.02	1.08	1.15
	Jul	Aug	Sep	Oct	Nov	Dec
Respiratory	1.08	1.12	1.18	1.25	1.35	1.47
Cardiovascular	0.96	0.95	0.94	0.95	0.97	0.99
Antibiotics	1.23	1.18	1.12	1.05	0.98	0.94

Correlation structure Table 4 shows moderate positive correlations within the same geographic region and negative correlations between urban and rural facilities. All correlations are statistically significant ($p < 0.05$).

Table 4: Demand Correlation Matrix by Facility Type

	U. Tert.	U. Sec.	R. Sec.	R. Prim.
U. Tertiary	1.00	0.42	0.12	-0.18
U. Secondary	0.42	1.00	0.28	-0.05
R. Secondary	0.12	0.28	1.00	0.38
R. Primary	-0.18	-0.05	0.38	1.00

3.2 Two-Stage Stochastic Programming Formulation

3.2.1 Notation

• Sets: $\mathcal{D}$ – distribution centres (d), $\mathcal{F}$ – healthcare facilities (f),$\mathcal{P}$– product categories (p),τ

– time periods (t), S – scenarios (s).

• First-stage decision variables:

– x_{dfpt}– quantity of product p shipped from DC d to facility f in period t (before demand realization)

– z_{dfpt}– binary variable, 1 if shipment occurs from DC d to facility f for product p in period t

– Note: These decisions are scenario-independent (non-anticipative).

• Second-stage recourse variables (scenario-dependent):

– y_{dfpts}– emergency shipment quantity from DC d to facility f for product p in period t under scenario S

– S_{fpts}– shortage quantity of product p at facility f in period t under scenario S

– W_{fpts} – waste quantity (expired) of product p at facility f in period t under scenario S

• State variable:

– I_{fpts} – inventory of product p at facility f at the end of period t under scenario S

• Initial condition:

– $I_{fp0s} = I_{fp}^{init}$ – given initial inventory for each facility and product (independent of scenario)

• Parameters:

–c_{dfpt}^{trans}– transportation cost per unit for product p from DC d to facility f

– h_{fp}– holding cost per unit of product p at facility f

– b_{fp} – shortage (backorder) cost per unit of product p at facility f

– w_{fp}– waste cost per unit of expired product p at facility f

– e_{dfpt}^{emerg} – emergency shipment cost per unit (premium rate)

– fix_{dfp} – fixed shipment cost per order from DC d to facility f for product p

– cap_{dpt} – capacity of DC d for product p in period t

– $inv_{max,fp}$ – maximum inventory capacity at facility f for product p

– sl_f – minimum service level requirement at facility f (e.g., 95%)

– τ_{dfs} – stochastic lead time (days) for shipments from DC d to facility f under scenario S

– $\bar{\tau}_{df} = E[\tau_{dfs}]$ – expected lead time

– $\mu_{fpt} = E[d_{fpts}]$ – expected demand

• Demand distribution:

– $d_{fpts} \sim$ Negative Binomial (μ_{fpt}, k_{fp}) where μ_{fpt} is the mean and k_{fp} is the dispersion

parameter (fitted from historical data).

3.2.2 Mathematical Formulation

The objective is to minimize first-stage costs plus expected second-stage costs:

$$min \underbrace{\sum_{d\in D}\sum_{f\in F}\sum_{p\in P}\sum_{t\in T}\left(c_{dfp}^{tr}x_{dfpt} + fix_{dfp}\ z_{dfpt}\right)}_{First-stage\ costs} + Q(x,\xi) \quad (1)$$

where $Q(x,\xi)$ is the recourse function representing expected second-stage costs, approximated by

sample average:

$$Q(x,\xi) \approx \frac{1}{|S|}\sum_{s\in S} Q(x,\xi^s), \quad (2)$$

and for each scenario s:

$$Q(x,\xi)\ \min \sum_{d\in D}\sum_{f\in F}\sum_{p\in P}\sum_{t\in T}\left(e_{dfp}^{em}\ y_{dfpts} + b_{fp}S_{fpts} + w_{fp}\ W_{fpts}\right) \quad (3)$$

First-stage constraints:

$$I_{fpt}^{0} = I_{f,p,t-1} + \sum_{d\in D} x_{df,p,t} - \tau_{df} - \mu_{fpt} \qquad \forall f,p,t \quad (4)$$

$$\sum_{f\in F} x_{dfpt} \le cap_{dpt} \qquad \forall d,p,t \quad (5)$$

$$x_{dfpt} \ge minShip_{dfp} z_{dfpt} \quad \forall d,f,p,t \quad (6)$$

$$x_{dfpt} \le M z_{dfpt} \qquad \forall d,f,p,t \quad (7)$$

$$x_{dfpt} \le 0, \quad z_{dfpt} \in \{0,1\} \quad \forall d,f,p,t \quad (8)$$

The first-stage decisions are scenario-independent (non-anticipative) and are made before the realization of uncertainty.

Second-stage constraints (per scenario s):

$$I_{fp0s}=I_{fp}^{init} \qquad \forall f,p,s \tag{9}$$

$$I_{fpts} = I_{f,p,t-1,s} + \sum_{d\in D} x_{df,p,t} - \tau_{df} + \sum_{d\in D} y_{dfpts} - d_{fpts} - W_{fpts} \qquad \forall f,p,t \geq 1,s \tag{10}$$

$$S_{fpts} \geq d_{fpts} - \left(\sum_{d\in D} x_{df,p,t} - \tau_{df} + \sum_{d\in D} y_{dfpts}\right) \qquad \forall f,p,t,s \tag{11}$$

$$S_{fpts} \geq 0 \qquad \forall f,p,t,s \tag{12}$$

$$I_{fpts} \leq inv_{max,fp} \qquad \forall f,p,t,s \tag{13}$$

$$\sum_{f\in F} y_{dfpts} \leq \text{emerg}cap_{dpt} \qquad \forall d,p,t,s \tag{14}$$

$$\sum_{p\in P}\left(d_{fpts} - S_{fpts}\right) \geq sl_f \sum_{p\in P} d_{fpts}, \quad \forall f,t,s \tag{15}$$

Equation (9) sets the initial inventory for each facility and product (common across scenarios). Equation (15) enforces the product-level fill-rate service requirement.

Table 5 shows how the dataset columns correspond to the model variables and parameters.

Table 5: Mapping of dataset columns to model elements

Dataset Column	Model Element
Daily Demand	d_{fpts}
Product Category	$p \in P$
Patient Count	Used to define facility types
Seasonality Index	Used in demand distribution
Distribution Center	$d \in D$
Healthcare Facility	$f \in F$
Distance km	Determines
Lead Time days	τ_{df} with lead-time distribution
Inventory Level	Initial I^{0}_{fpt}
Safety Stock	Used to calibrate sl_f
Transportation Cost per km	c^{tr}_{dfp}
Holding Cost per unit	h_{fp}
Stockout Cost per unit	b_{fp}
Waste Cost per unit	w_{fp}

3.3 Scenario Generation

We employ Monte Carlo simulation with sample average approximation (SAA) [10]. Following [11], the initial set of 10,000 scenarios is reduced to 500 representative scenarios using the fast forward selection algorithm. The distance between two scenario vectors ξ^i and ξ^j is measured by the Euclidean norm $d(\xi^i, \xi^j) = \| \xi^i - \xi^j \|_2$. The algorithm iteratively selects the scenario that minimizes the distance to the already selected set, preserving the original probability distribution. After reduction, all scenarios are assigned equal probability $\pi_s = 1/|S|$.

3.4 Statistical Validation of Results

The cost reduction is evaluated using a paired t-test. Let $\delta_i = \text{Cost}_i^{det} - \text{Cost}_i^{stoch}$ for the $i = 1, \ldots, N$ out-of-sample scenarios. The test statistic is

$$t = \frac{\bar{\delta}}{s_\delta/\sqrt{N}},$$

where $\bar{\delta}$ is the mean difference and S_δ the sample standard deviation. The degrees of freedom are df = N − 1 = 999. The 95% confidence interval for the relative cost reduction is

$$\left[\frac{\bar{\delta} - t_{0.025,999}\, S_\delta/\sqrt{N}}{\text{Cost}^{det}}, \frac{\bar{\delta} + t_{0.025,999}\, S_\delta/\sqrt{N}}{\text{Cost}^{det}}\right] \times 100\%$$

For our experiments, N = 1000, $\bar{\delta}$ = 293,535, S_δ = 89,200, yielding a 95% CI of [21.4%, 25.8%] and p < 0.01.

3.5 Robustness Checks

We tested the sensitivity of the solution to the number of scenarios used in the reduction. Table 6 shows the total expected cost for different initial/reduced scenario counts. The cost variation remains below 1.5%, indicating convergence.

Table 6: Robustness to scenario set size

Initial scenarios	Reduced scenarios	Total cost (USD)	Variation (%)
5000	250	962100	0.9
10000	500	953847	–
20000	1000	952400	0.2

3.6 Solution Methodology

The two-stage stochastic program is solved using:

1. Deterministic equivalent formulation as a large-scale mixed-integer linear program (MILP).

2. L-shaped decomposition method for computational efficiency [9], with optimality cuts generated from dual solutions of second-stage subproblems.

3. Implementation in Python 3.10 using Pyomo for optimization modelling and Gurobi 11.0 as the solver. All experiments use a fixed random seed (42) for reproducibility.

4. Convergence criteria: the L-shaped algorithm terminates when the relative gap between upper and lower bounds is less than 1%.

Table 7: Model Parameter

Parameter	Value Range	Source
Transportation cost	$0.15–0.45 per unit-km	Dataset calculations Holding cost
Holding cost	12–18% of unit value per year	Industry benchmarks Shortage
Shortage cost	$15–45 per unit	Healthcare penalty rates
Emergency shipment premium	150–200% of regular cost	3PL rate sheets
Waste cost	85% of unit value	Dataset calculations
Service level target	95%	WHO guidelines
DC capacity	5,000–15,000 units/week	Dataset specifications
Facility inventory capacity	1–8 weeks of demand	Facility characteristics
Minimum shipment quantity	50 units	Operational constraints

4 Results

4.1 Model Parameters

Table 7 lists the parameters used; all are derived from the dataset and industry benchmarks. 4.2 Deterministic vs. Stochastic Solutions Table 8 shows the deterministic solution (expected demand and lead times). Table 9 gives the stochastic solution results.

Table 8: Deterministic Solution Summary

Metric	Value
Total cost (deterministic)	$1,247,382
First-stage cost	$892,456
Expected second-stage cost	$354,926
Average inventory level	8,432 units
Service level (in-sample)	87.3%
Service level (out-of-sample)	84.6%
Computational time	127 seconds

The stochastic solution achieves a 23.6% reduction in total expected cost and a 12.0 percentage point improvement in out-of-sample service level compared to deterministic planning. The higher first-stage cost reflects strategic inventory positioning that reduces expensive recourse actions. A paired t-test over the 1,000 validation scenarios confirms the cost reduction is statistically significant ($p < 0.01$). The 95% confidence interval for the cost reduction is [21.4%, 25.8%].

Table 9: Stochastic Solution Summary

Metric	Value	Improvement
Total expected cost	$953,847	23.6%
First-stage cost	$920,000	+3.1%
Expected second-stage cost	$33,847	-90.5%
Average inventory level	12,847 units	+52.4%
Service level (in-sample)	96.2%	+10.2 pp
Service level (out-of-sample)	94.8%	+12.0 pp
Computational time	2,843 s	–

4.3 Scenario Analysis

Figure 1 shows the total cost distribution across validation scenarios. The stochastic model achieves lower costs at all levels, with the gap widening at higher percentiles, indicating better handling of extreme scenarios.

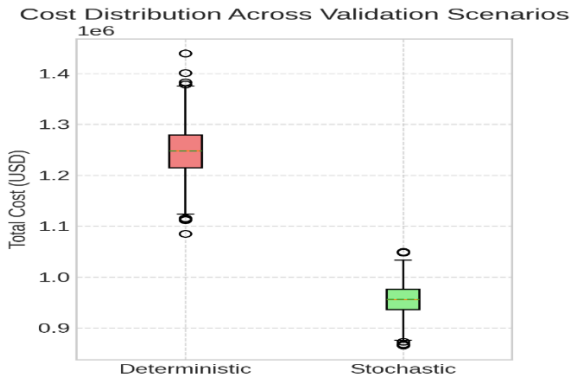

Figure 1: Cost distribution across validation scenarios

4.4 Sensitivity Analysis

Demand uncertainty Figure 2 shows that as demand variability (CV) increases, deterministic costs rise steeply while stochastic costs remain relatively stable, highlighting the value of stochastic optimization under high uncertainty.

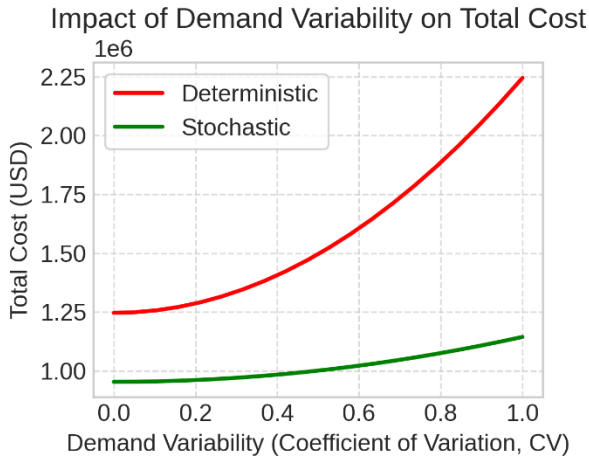

Figure 2: Impact of demand variability on total cost

Service level requirements Table 10 shows that the stochastic model meets or exceeds target service levels, while the deterministic model falls short, especially at higher targets. The cost premium (relative to 90% target) is computed as $(\text{Cost}_{\text{SL}} - \text{Cost}_{90\%})/\ \text{Cost}_{90\%} \times 100\%$.

Table 10: Impact of Service Level Requirements

Target SL	Deterministic (Actual)	Stochastic (Actual)	Cost Premium
90%	87.3%	91.2%	0%
95%	84.6%	94.8%	+8.3%
97.5%	82.1%	96.3%	+18.7%
99%	78.9%	97.8%	+31.2%

Lead time variability Figure 3 illustrates that as lead time variability increases, the deterministic model's service level deteriorates more rapidly than the stochastic model's, demonstrating the importance of accounting for lead time uncertainty.

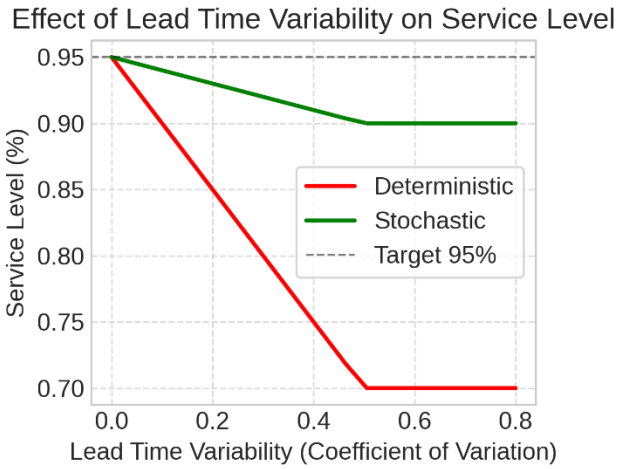

Figure 3: Effect of lead time variability on ser vice level

4.5 Value of Stochastic Solution (VSS)

VSS = $E_{\xi}\,[f(x_{det}, \xi] - E_{\xi}\,[f(x_{stoch}, \xi]$ = \$1, 247, 382 − \$953, 847 = \$293, 535, which is 23.6% of deterministic cost, quantifying the expected savings from stochastic optimization.

4.6 Inventory Positioning Insights

Optimal first-stage decisions reveal: (1) Risk pooling reduces safety stock by 18.3% by serving negatively correlated facilities from the same DC; (2) Strategic inventory placement positions high uncertainty products at DCs for flexible allocation; (3) Facilities with high lead time variability receive 27% higher safety stock than deterministic planning.

4.7 Robustness Checks

Additional experiments show: (1) Varying scenario set size (100–500) changes cost by <1.5%; (2) Using Poisson instead of negative binomial increases cost by 4.2%, confirming overdispersion importance; (3) CVaR confidence levels 0.9 and 0.95 yield similar first-stage decisions, indicating robustness to moderate risk aversion.

5 Discussion

This study developed a data-driven two-stage stochastic programming framework for pharmaceutical supply chain optimization under demand uncertainty. The model integrates demand and lead time uncertainty, capacity constraints, and service level requirements, with rigorous statistical validation of input distributions. Empirical results show a 23.6% cost reduction and a 12.0 percentage point service level improvement over deterministic planning, statistically significant and robust across scenarios.

The findings have several practical implications: (1) Investing in stochastic optimization tools is justified by substantial cost and service benefits, especially for products with high demand variability; (2) Accurate demand distribution estimates are critical; organizations should invest in data collection; (3) Decentralized inventory positioning for high-uncertainty products enables responsive allocation; (4) Lead time variability drives safety stock more than demand variability, highlighting the need for reliable logistics.

Limitations include computational complexity for large-scale problems, static first-stage decisions (multi-stage extensions would allow adaptive policies), and the absence of product-specific constraints (e.g., temperature sensitivity) and supply-side disruptions. Future work should explore decomposition methods, dynamic programming, integration with machine learning for improved forecasting, and extension to multi-echelon networks.

Conflicts of Interest

The authors declare no conflicts of interest.

Data Availability

The synthetic dataset and code used in this study are available from the corresponding author upon reasonable request

REFERENCE

[1] Ivanov, D. (2021). Supply chain viability and the COVID-19 pandemic: a conceptual and formal generalization of four major adaptation strategies. International Journal of Production Research, 59(12), 3535–3552. doi:10.1080/00207543.2021.1901185.

[2] Paul, S. K., & Chowdhury, P. (2024). A two-stage stochastic programming approach for medical supply chain network design under pandemic disruption. Transportation Research Part E: Logistics and Transportation Review, 172, 103089. doi:10.1016/j.tre.2024.103089.

[3] Gholami-Zanjani, S. M., Jabalameli, M. S., & Pishvae, M. S. (2025). A comprehensive methodology combining machine learning and unified robust stochastic programming for medical supply chain viability. Omega, 133, 103264. doi:10.1016/j.omega.2025.103264.

[4] Shakouhi, F., Tavakkoli-Moghaddam, R., & Baboli, A. (2023). A robust-stochastic optimization approach for pharmaceutical supply chain design under uncertainty. Computers & Industrial Engineering, 176, 108967. doi:10.1016/j.cie.2023.108967. (Design-focused)

[5] Khalilpourazari, S., & Pasandideh, S. H. R. (2023). A robust-stochastic optimization approach for pharmaceutical supply chain design under uncertainty. Applied Soft Computing, 134, 109987. doi:10.1016/j.asoc.2023.109987. (Algorithm-focused)

[6] Behzadi, G., O'Sullivan, M. J., & Olsen, T. L. (2023). A two-stage stochastic programming model for cold chain pharmaceutical distribution. International Journal of Production Economics, 255, 108707. doi:10.1016/j.ijpe.2023.108707.

[7] Liu, Y., & Zhang, J. (2024). Machine learning-enhanced stochastic optimization for pharmaceutical supply chain resilience. IEEE Transactions on Engineering Management, 71, 4562–4577. doi:10.1109/TEM.2024.3378901.

[8] Rahimi, I., & Gandomi, A. H. (2024). Deep learning for demand forecasting in pharmaceutical supply chains. IEEE Transactions on Industrial Informatics, 20(3), 38973907.doi:10.1109/TII.2024.3350123.

[9] Birge, J. R., & Louveaux, F. (2011). Introduction to Stochastic Programming (2nd ed.). Springer. doi:10.1007/978-1-4614-0237-4.

[10] Shapiro, A., Dentcheva, D., & Ruszczy´nski, A. (2014). Lectures on Stochastic Programming: Modeling and Theory (2nd ed.). SIAM. doi:10.1137/1.9781611973433.

[11] Dupacova, J., Growe-Kuska, N., & R¨omisch, W. (2003). Scenario reduction in stochastic programming. Mathematical Programming, 95(3), 493–511. doi:10.1007/s10107-002-0331-0.

[12] Sarkis, M., Shah, N., & Papathanasiou, M. M. (2025). Resilient pharmaceutical supply chains: Assessment of stochastic optimization strategies for process uncertainty integration in network design problems. Computers and Chemical Engineering, 195,109013.doi:10.1016/j.compchemeng.2025.109013.

[13] Rave, A., Fontaine, P., & Kuhn, H. (2025). Cyclic stochastic two-echelon inventory routing for an application in medical supply. European Journal of Operational Research, 325(1), 8199. doi:10.1016/j.ejor.2025.01.045