Optimal ATM Replenishment Policies under Demand Uncertainty

Based on the work by Ekinci, Serban, and Duman (2021)

DS-502 Project by Denizcan Biçakcı – Hamza Alsatari – Metin Soydeğer

Operational Research (2021) 21:999–1029
https://doi.org/10.1007/s12351-019-00466-4

ORIGINAL PAPER

Check for updates

Optimal ATM replenishment policies under demand uncertainty

Yeliz Ekinci[1] · Nicoleta Serban[2] · Ekrem Duman[3]

Received: 18 February 2017 / Revised: 5 November 2018 / Accepted: 21 February 2019 /
Published online: 15 March 2019

Paper

Abstract
The use of Automated Teller Machines (ATMs) has become increasingly popular throughout the world due to the widespread adoption of electronic financial transactions and better access to financial services in many countries. As the network of ATMs is becoming denser while the users are accessing them at a greater rate, the current financial institutions are faced with addressing inventory and replenishment optimal policies when managing a large number of ATMs. An excessive ATM replenishment will result in a large holding cost whereas an inadequate cash inventory will increase the frequency of the replenishments and the probability of stock-outs along with customer dissatisfaction. To facilitate informed decisions in ATM cash management, in this paper, we introduce an approach for optimal replenishment amounts to minimize the total costs of money holding and customer dissatisfaction by taking the replenishment costs into account including stock-outs. An important aspect of the replenishment strategy is that the future cash demands are not available at the time of planning. To account for uncertainties in unobserved future cash demands, we use prediction intervals instead of point predictions and solve the cash replenishment-planning problem using robust optimization with linear programming. We illustrate the application of the optimal ATM replenishment policy under future demand uncertainties using data consisting of daily cash withdrawals of 98 ATMs of a bank in Istanbul. We find that the optimization approach introduced in this paper results in significant reductions in costs as compared to common practice strategies.

Keywords Automated teller machines · Replenishment policy · Demand uncertainty

✉ Yeliz Ekinci
ekinciyeliz@yahoo.com

[1] Business Administration, Istanbul Bilgi University, Istanbul, Turkey

[2] Industrial and Systems Engineering, Georgia Institute of Technology, Atlanta, USA

[3] Industrial Engineering, Ozyegin University, Istanbul, Turkey

Springer

Headline: *Optimal ATM Replenishment Policies Under Demand Uncertainty*

Authors: Yeliz Ekinci, Nicoleta Serban, Ekrem Duman

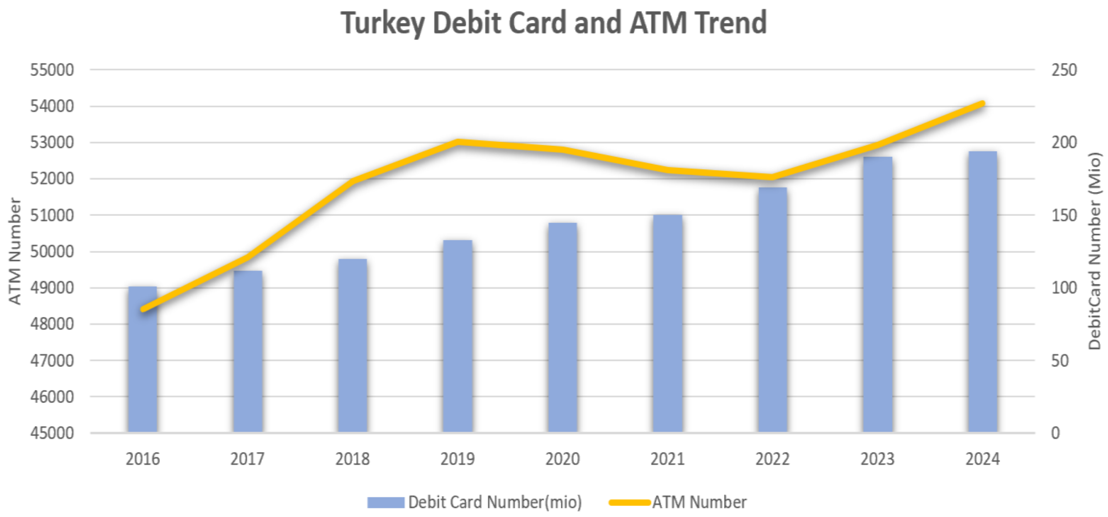
Published: *Operational Research*, Cilt 21, Sayfa 999–1029 (2021)

DOI: https://doi.org/10.1007/s12351-019-00466-4

Key idea: Join interval forecasting with robust optimization **to** minimize **worst-case regret**

Turkey ATM and DebitCard Numbers Trend (BKM)

- In 2024, 54.075 ATM in Turkey
    - %44 next to branch (on-site)
    - %56 out of branch (off-site)
- 80,3 ATM per 100.000 people
- In İstanbul, 2.100 ATM points – 10,5 ATMs per 100.000 people



Bankacılar tatile gitti ATM'de para bitti!

Problem Definiton

- **The Problem:** Banks must decide how much cash to load into each ATM, periodically, **without knowing future cash withdrawals** in advance.
- **Objective:** To **minimize the total cost** associated with:
  - **Cash holding cost** (interest on excess funds),
  - **Penalty cost** (for not meeting customer demand),
  - **Customer dissatisfaction** (non-linear, service-related impact).
- **Decision Variable** X_{it}= cash loaded into ATM i at week t.

Key Challenges

- **Why This Problem is Difficult:**
    - Future demand is stochastic, not deterministic.
    - Forecasting and optimization are usually separated; we integrate them in one loop.
    - ATM networks exhibit geographic and behavioral heterogeneity, often overlooked.
- **Research Gap:**
    - Very few models integrate **demand forecasting under uncertainty** with **robust optimization** of replenishment decisions

Proposed Approach Overview

- **Innovative Contribution of the Paper:**
  - Forecasts future ATM cash demand using **prediction intervals**, not point estimates.
  - Applies a **robust optimization model** that minimizes worst-case regret across the prediction interval.
  - Incorporates **locational clustering** and socio-economic variables into demand modeling.
  - Provide practical cost advantages in ATM cash replenishment strategies.
- Input → *Forecast (95 % PI)* → Robust Optimization → Load Decision
- **Real-World Application:**
  - Tested on real withdrawal data from 98 ATMs in Istanbul of a bank
  - Outperforms traditional fixed-safety-stock methods in cost and reliability
  - 2,5 years observation duration
  - 4-week planning prediction

İstanbul district map

- New City and Bosphorus – 31 ATMs
- Galata, Golden Horn and Sultanahmet – 5 ATMs
- Western Suburbs – 62 ATMs

Forecasting Approach – Why Regression?

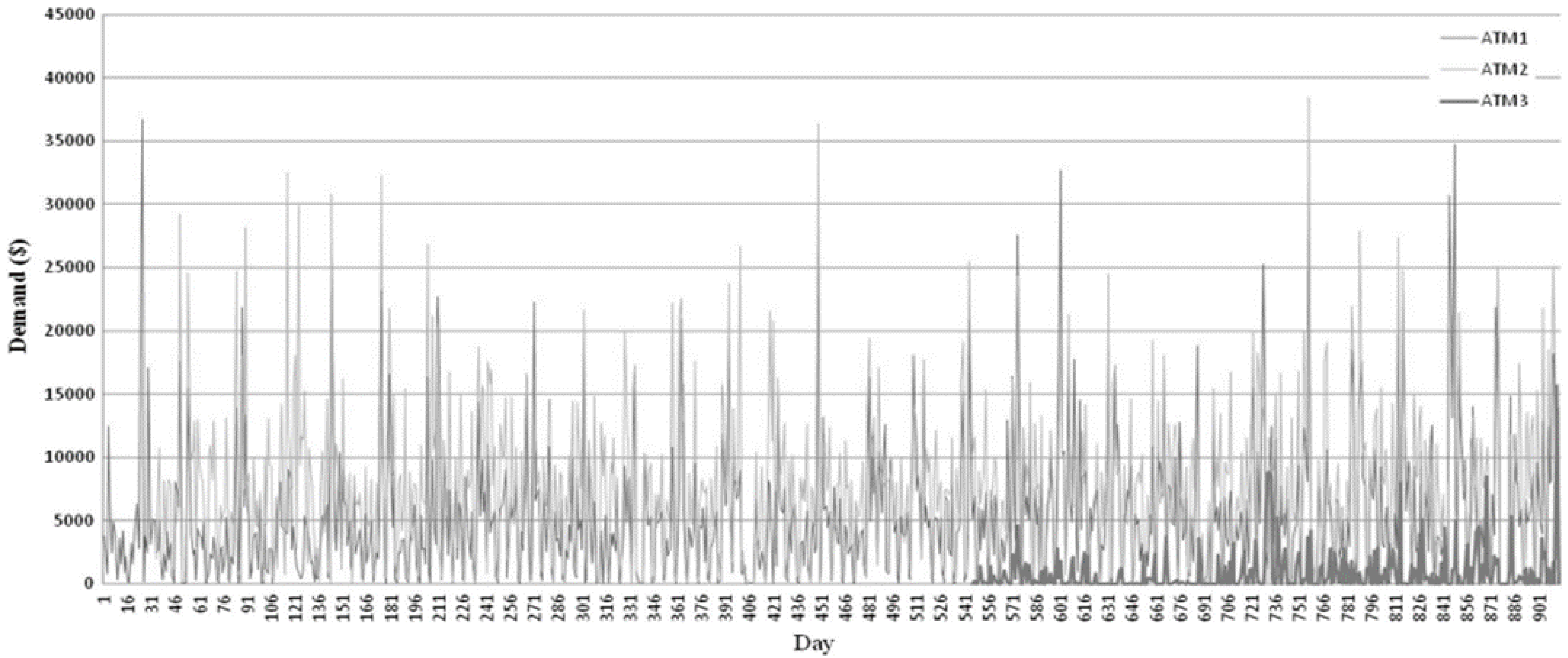
- Why not traditional methods?
  - **Time series models** require stationarity, which does not hold in ATM cash withdrawal data.
  - **Machine learning models** capture nonlinear patterns but are often **less interpretable**.
  - Most models are **univariate**, ignoring useful patterns across other ATMs
  - **Prediction intervals** are difficult to derive from non-parametric or ML models.
- Method Features:
  - **Geographically clustered ATMs** (based on proximity, urban context, usage).
  - Use of **location-related variables**: population, points of interest, nearby ATMs, demographics
  - Use of **time-related variables**: festivals, holidays, salary days, special events.

"Everything is related to everything else, but near things are more related than distant things." *— Tobler (1970)*

Independent variables in the Regression Model

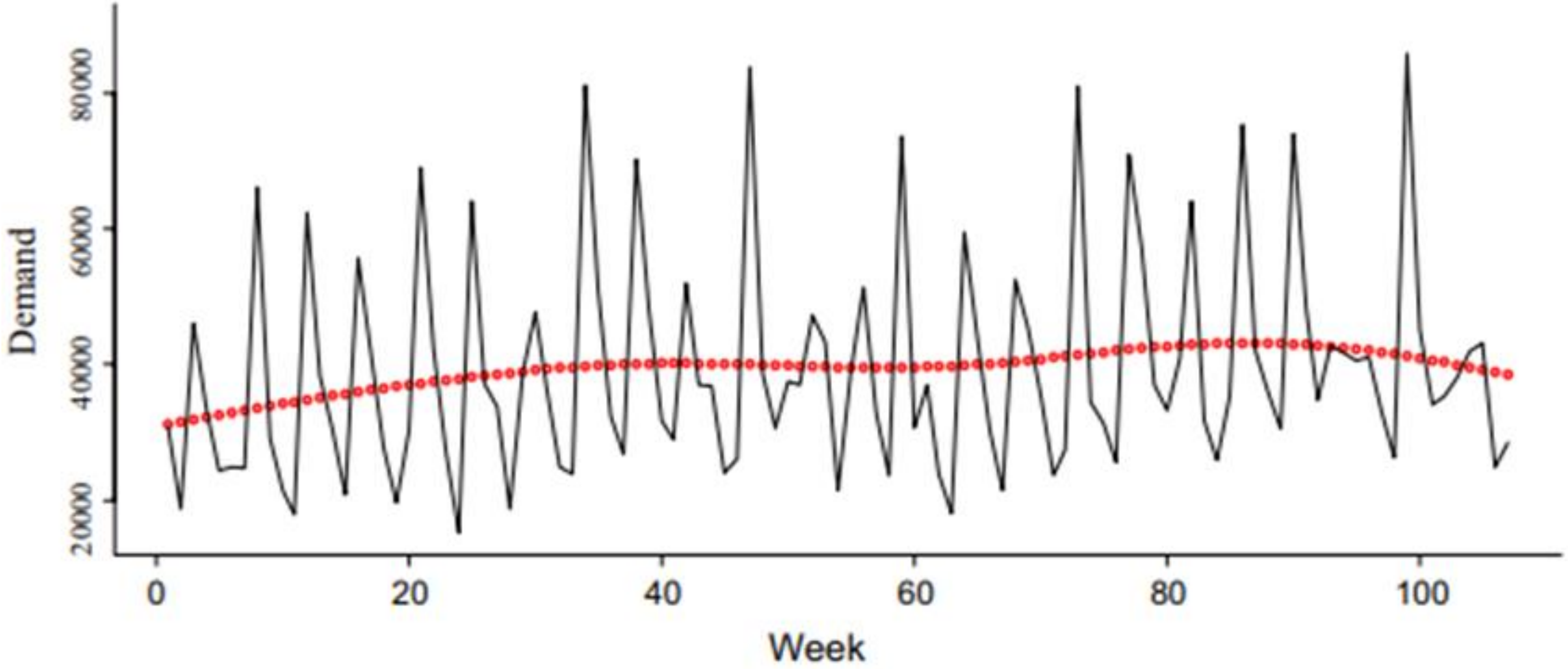
Code	Type (categorical/numerical)	Explanation of the variable
A	Categorical	Variable which denotes the location cluster
B	Categorical	Variable specifying whether the week is before a religious festival
C	Categorical	Variable specifying whether the week is before a holiday
D	Categorical	Variable specifying whether the week is before a special day, such as Valentine's day, mothers' day, fathers' day etc.
E	Categorical	Variable specifying whether the week includes a special day, such as Valentine's day, mothers' day, fathers' day etc.
F	Categorical	Variable specifying whether the week includes a religious festival or national festival
G	Categorical	Variable specifying whether the week includes the first day of the month (This is important since this day is the salary day in Turkey)
H	Categorical	Variable specifying whether the week includes the fifteenth day of the month (This is important since this day is the salary day in Turkey)
I	Numerical	Variable which denotes the week number (there are 52 weeks in a year and this variable is important to capture the seasonality)
J	Numerical	Variable which denotes the number of points of interests (school, restaurant, shopping center, airport, plaza etc.) in 500 m vicinity of the ATM
K	Numerical	Variable which denotes the number of points of interests (school, restaurant, shopping center, plaza etc.) in 1000 m vicinity of the ATM
L	Numerical	Variable which denotes the number of other ATMs in the street
M	Numerical	Variable which denotes the number of social buildings in the street (such as restaurant, shop, cafeteria etc.)
N	Numerical	Variable which denotes the number of work related buildings in the street (such as plaza, governmental building, school, etc.)
O	Numerical	Variable which denotes the number of commercial buildings in the street (such as airport, port, bus terminal)
P	Numerical	Variable which denotes the total number of points of interests in the street
Q	Numerical	Variable which denotes the population in the street
R	Numerical	Variable which denotes the number of people who are university graduate in the street

Daily cash withdrawal amounts for 3 ATMs



ATM1 and ATM2 are in the same county, ATM3 is located in a diferent county.

Weekly cash withdrawal amounts for one ATM



Time series plot of an ATM along with estimated trend in pointed line

Demand Prediction & Interval Estimation

Example Forecast Table (ATM 1)

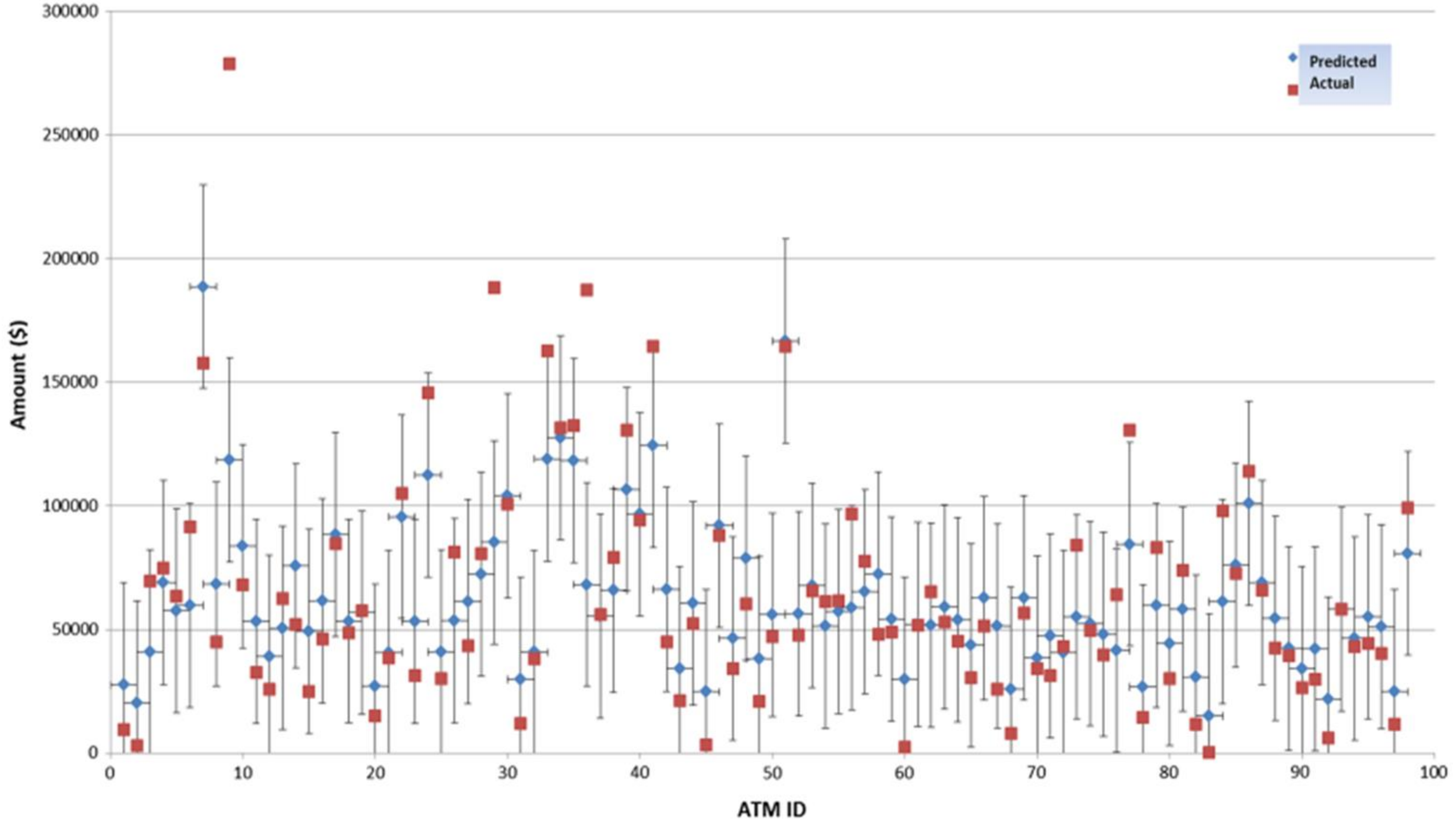
Forecast Interval for ATM 1 (₺, 95% Confidence)			
Week	Lower (₺)	Point (₺)	Upper (₺)
1	43,568	49,275	54,982
2	44,936	50,643	56,350
3	46,453	52,160	57,867
4	48,123	53,830	59,537

Example 95 % prediction interval for ATM 1 (based on 4-week forecast)

Modeling Steps:

- Step 1: **Detrending**
  - Remove long-term trends in the data via polynomial regression or spline smoothing (e.g., thin-plate splines)
  - For each ATM i-th, the smoothed weekly trend is: $\hat{r}_i(w)$
  - Detrended value: $Z_{iw} = Y_{iw} - \hat{r}_i(w)$
- Step 2: **Fit a linear regression model** to the de-trended data, Z_{iw}
  - **Categorical variables**: Cluster type, salary/holiday weeks, weekday
  - **Numerical variables**: Population density, POI density, nearby ATM count, commercial density
- Step 3**: Point Forecast and Prediction Interval**
  - Predict $\hat{z}_{it}$, then recover actual demand: $\hat{Y}_{it} = \hat{Z}_{it} + \hat{r}_i(t)$
  - Prediction interval: $\hat{Y}_{it} \pm z_{a/2} \cdot \hat{\sigma}_{\hat{Y}_{it}}$
    - $\hat{\sigma}_{\hat{Y}_{it}}$ is the estimated standard error of the predicted mean
    - $z_{a/2}$ is the normal quantile
  - Model $R^2 = 0.76$ — better than undetrended model ($R^2 = 0.32$)

Prediction intervals of the first week



Optimization Model for ATM Replenishment

Objective :

- To determine **how much cash to load into each ATM** for each week, using **robust optimization** to minimize the **worst-case cost** under demand uncertainty.

Why Robust Optimization?

- Future cash demand, Y_{it} is **unknown** but bounded within a **prediction interval**.
- Goal: Find replenishment amount X_{it} that performs well in the **worst-case scenario**.
- Avoids reliance on precise probability distributions (unlike stochastic optimization).

Key Definitions

Symbol	Description
$Y_{it} \in [L_{it}, U_{it}]$	Future demand for ATM i at time t, within lower and upper prediction bounds
$\hat{Y}_{it}$	Predicted (mean) cash demand at ATM i at time t (forecast output)
x_{it}	**Decision variable** – amount of money to load (replenish) into ATM i at time t
C_{it}	**Total cost** incurred for ATM i at time t; the optimization objective to minimize
L_{it}, U_{it}	Lower and upper bounds of the demand prediction interval
c	**Holding cost rate** – cost per dollar of excess cash (opportunity cost or interest)
h	**Fixed penalty cost** if a stockout occurs (ATM runs out of cash)
g	**Dissatisfaction cost** per dollar of unmet demand (reflects customer dissatisfaction)
$\hat{\sigma}_{Y_{it}}$	Estimated standard error of the predicted demand

Objective Function

Objective Function : The objective is to minimize the **maximum cost** over all possible outcomes under demand uncertainty.

We do not know the exact value of Y_{it} , the demand for ATM i at time t, but we know that it lies within a **prediction interval**: $Y_{it} \in [L_{it}, U_{it}]$

$$\min C_{it}$$

Where Cit is the **total cost** incurred at ATM *i*, time *t*.

$$\min_{x_{it} \in [L_{it}, U_{it}]} \max \begin{cases} c \cdot (x_{it} - L_{it}) & \text{(excess)} \\ h + g \cdot (U_{it} - x_{it}) & \text{(shortage)} \end{cases}$$

Constraints

Constraints :

- **Holding cost** :

$$C_{it} \geq c \cdot (x_{it} - L_{it})$$

When actual demand is at its lowest: Y_{it} = L_{it}. This ensures that if cash loaded exceeds the lowest possible demand, the holding cost is accounted for.

- **Shortage cost** :

$$C_{it} \geq h + g \cdot (U_{it} - x_{it})$$

When actual demand is at its highest Y_{it} = U_{it}. This accounts for the fixed penalty and customer dissatisfaction cost when the loaded cash is insufficient.

- **Feasibility Constraint:**

$$x_{it} \in [L_{it}, U_{it}]$$

The loaded cash must lie within the demand prediction interval.

Two Scenarios

Criteria	Scenario 1: Proposed Model (Robust Optimization)	Scenario 2: Classical Approach (Replenish Upper Bound)
Description	Optimizes the replenishment amount using prediction intervals and robust logic.	Replenishes the **upper bound** of the demand prediction interval.
Decision Formula	$x_{it}^{*} = \frac{h + g \cdot U_{it} + c \cdot L_{it}}{c + g}$	$x_{it} = U_{it}$
Cost Consideration	Balances **overstock (c)** and **stockout (h + g)** costs.	Primarily incurs **overstock cost**, avoids stockouts.
Service Level	Balanced; may allow occasional shortfalls.	Very high; almost no cash-outs occur.
Total Cost	**Lower overall cost**, due to smart balancing.	Generally **higher cost**, due to excess cash holding.
Realism	Reflects operational uncertainty and is **data-driven and adaptive**.	Simple to apply but **less cost-efficient**.
Main Advantage	Minimizes **maximum regret** under uncertainty.	Guarantees service reliability.
Main Disadvantage	Requires solving a formula; does not ensure 100% service.	Wastes resources due to excessive replenishment.

Comparison of Two Scenarios

Scenario 1

Cost of optimal replenishment under 0.5% dissatisfaction cost

	Week 1	Week 2	Week 3	Week 4
ATM 1	49	35	28	23
ATM 2	50	31	27	30
ATM 3	0	29	38	28
ATM 4	23	27	19	11
ATM 5	23	27	39	36
ATM 6	**23**	16	40	60
ATM 7	60	78	84	92
ATM 8	53	57	**165**	7
ATM 9	**665**	55	60	90
ATM 10	45	46	28	2
Total cost	2301			

Scenario 2

Cost of replenishing upper bound under 0.5% dissatisfaction cost

	Week 1	Week 2	Week 3	Week 4
ATM 1	59	42	34	30
ATM 2	59	37	32	36
ATM 3	12	38	46	37
ATM 4	36	39	31	23
ATM 5	36	39	50	48
ATM 6	10	28	52	72
ATM 7	72	91	96	104
ATM 8	65	69	**105**	19
ATM 9	**604**	67	72	102
ATM 10	57	58	40	14
Total cost	2562			

Total costs of this two Scenarios under different dissatisfaction costs

	0.5% Dissatisfaction cost	0.6% Dissatisfaction cost	0.7% Dissatisfaction cost	0.8% Dissatisfaction cost	0.9% Dissatisfaction cost	1% Dissatisfaction cost
Total cost—scenario 1	2301	2491	2659	2833	3000	3161
Total cost—scenario 2	2562	2700	2838	2976	3114	3252
Total cost improvement with the proposed model (i.e., scenario 1) (%)	10.19%	7.74%	6.31%	4.81%	3.66%	2.80%

Cost Structures – Upper vs Robust

Simulation-Based Comparison: Real Example from Week 1 ($\mathbf{ATM_1}$)

We simulate 1,000 possible demands uniformly sampled from the prediction interval of Week 1.
The table below shows one representative simulation, and the improvement is calculated as the **average over 1,000 runs**.

Metric	Upper Bound Policy	Robust Policy
Simulated Demand Y	\$48,328	\$48,328
Load Amount x	\$54,982	\$54,747
Cost	\$6.65	\$6.42

Cost Difference: \$0.24 saved
Improvement from 1000 runs: 0.22%

How Do the Two Policies Handle Uncertainty?

- Two possible outcomes based on actual demand Y:

If $x > Y \Rightarrow \text{Cost} = c(x - y)$ (holding/interest)

If $x < Y \Rightarrow$ Cost = $h + g(y - x)$ (penalty + dissatisfaction)

- Upper Bound Policy
  - Always loads $x = U$ (max expected demand)
  - Avoids stockouts, but increases holding costs
- Robust Policy
  - Uses a weighted formula:

$$x_{it}^{*} = \frac{h + gU + cL}{c + g}$$

  - Balances between upper and lower bounds
  - Minimizes worst-case cost instead of just risk

Python Implementation: Forecasting and Robust Decision Logic

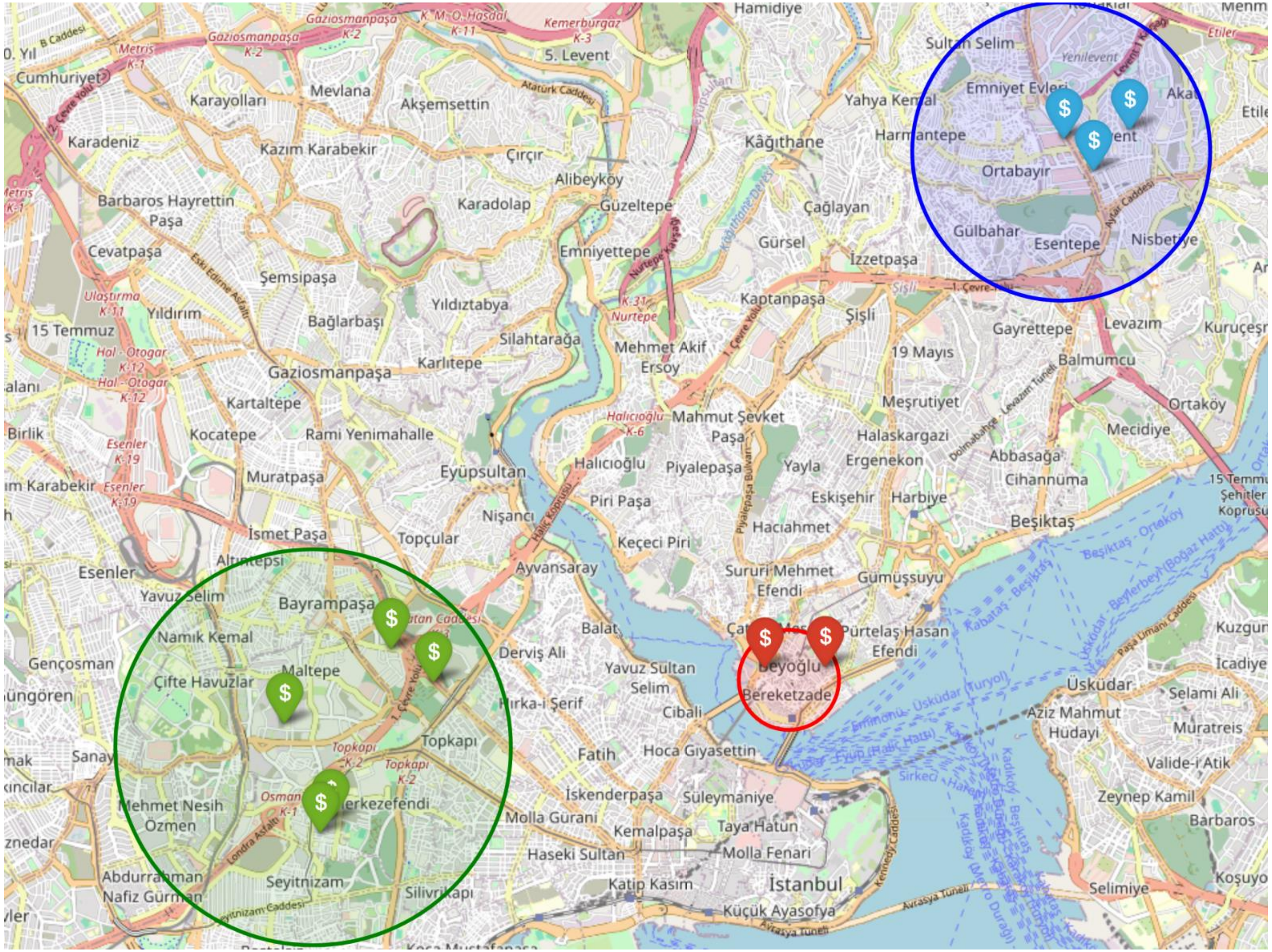
```python
def forecast_with_intervals(atm_data,
historical):
    detrended, trend = detrend_data(data)
    model = LinearRegression()
    model.fit(X, y)
    ...
    std_error = np.std(recent_errors) * 1.5
    lower = point_forecast - 1.96 * std_error
    upper = point_forecast + 1.96 * std_error
```

```python
def robust_optimization(L, U, c, h, g):
    return (h + g * U + c * L) / (c +
g)
```

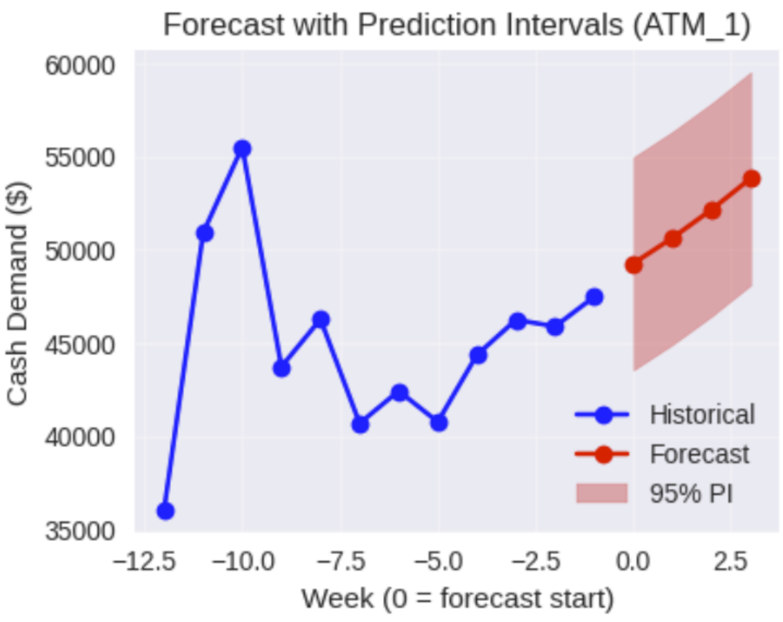
Fits linear regression on detrended data and generates 95% prediction intervals for next 4 weeks.

Implementing equation (4) from the paper to find optimal ATM cash amount minimizing worst-case cost.

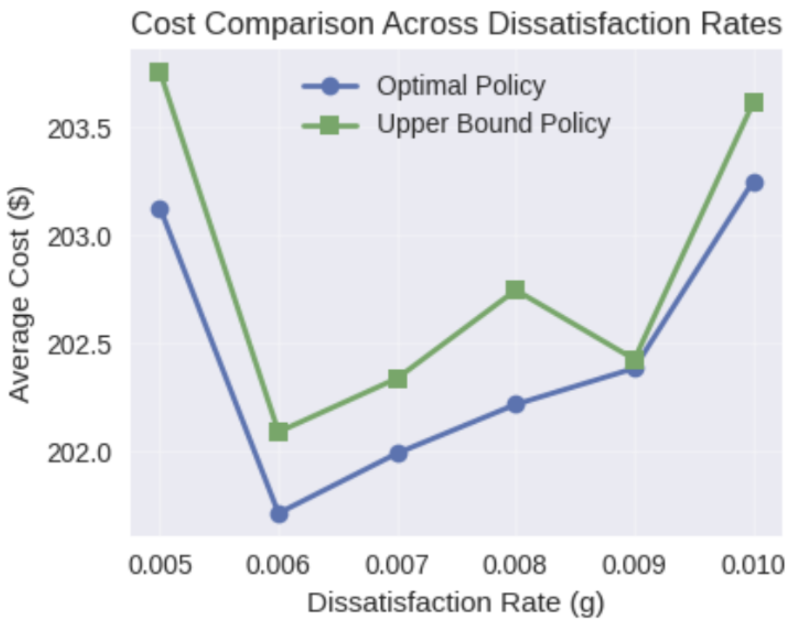
Geographic Distribution of Optimal Replenishment



Sample ATM Forecast & Decision



Historical demand: last 12 weeks (blue line)
Forecast demand: next 4 weeks (red line)
Shaded band: 95% **prediction interval**
Purpose: Quantifies uncertainty → not just one "guess," but a *range* of likely values
Why it matters: This range feeds directly into the **robust formula** to compute replenishment



Shows **expected cost** across dissatisfaction rates *(g = 0.005 to 0.010)*
Blue line: Robust Policy — consistently cheaper
Green line: Upper Bound Policy — blindly loads full max demand
Result: Robust approach reduces cost at all g values

Summary Results

g	Optimal Cost	Upper Cost	Improvement
0.005	$203.13	$203.76	0.3%
0.006	$201.71	$202.09	0.2%
0.007	$201.99	$202.34	0.2%
0.008	$202.22	$202.75	0.3%
0.009	$202.38	$202.42	0.0%
0.010	$203.25	$203.61	0.2%

For each g value, shows:
Optimal cost (our model)
Upper bound cost (baseline)
% improvement
Up to 0.3% reduction in cost
Savings are small per ATM,
but **scale** over a full network weekly

Why the Robust Policy Wins

Resilience

Performs reliably under demand uncertainty

Avoids both overfilling and stockout

Uses prediction intervals, not naive guesses

Cost-Effectiveness

Minimizes total cost by balancing risk

Achieves up to 0.3% savings

Proven better in 1000+ simulations

Business Alignment

Matches weekly operational planning banks already use

Considers clustering, demand patterns, holidays

Practical for use with real ATM networks

Model Extensions & Limitations

- Model Limitations / Assumptions
  - Assumes demand lies in forecast interval
  - No vehicle routing or delivery limits yet
  - Weekly re-optimization
  - Fixed parameters
  - No cash withdrawal pattern modeling

Extension	What It Adds	Trade-off
Inventory Carryover	Tracks leftover cash between weeks	↑ Realism, ↑ Complexity (medium)
Route Optimization (IRP)	Adds vehicle routing & delivery constraints	↑ Realism, ↑ Complexity (NP-hard)
Multi-objective Optimization	Balances cost vs. service level targets	↑ Applicability, ↑ Modeling effort
ATM Priority Classes	Differentiates ATMs by criticality	↑ Context awareness, ↓Complexity