Data-driven System Simulation in NYGH ED

Presenter: Nancy Li
Supervised by: Prof. Arik Senderovich, Prof. Dmitry Krass, and Prof. Opher Baron
September 1, 2021

Agenda

01

Introduction

Preliminary Data Analysis

02

Analysis

Model Learning and Experimental Setting

03

Results

LOS Goodness of Fit, and Intervention Effect

04

Conclusion

Limitations and Possible Future Directions

01

Preliminary Data Analysis

Arrival Process

Service & Departure Process

LOS = Leaving ED Time - Arrival Time

Ambulance Arrival | **Triage** | **Registration** | **Assessment** | **Disposition** | **Leaving ED**

NYGH ED System Simulation

Arrivals

Generating patient arrivals into the ED

Capacity

Inferring resource capacity (missing resource information; approximate via time of day/day of week)

Length of Stay (LOS)

Modelling patient's LOS in the ED system (can extend to individual stations in future work)

Interventions

Cutting down consult patients' LOS and analyzing possible secondary congestion effects

Overview

Assumptions

- Patient's LOS is **state-aware**

Model Selection

- **Random Forest Regressor** as the LOS model
- LOS = RF pred + noise

Simulation Approach

- Data-driven arrivals
- Length of stay (LOS)
- Simulate

Initial Data*

- **Training**: 2016 & 2017
- **Testing**: 2018

Data Used

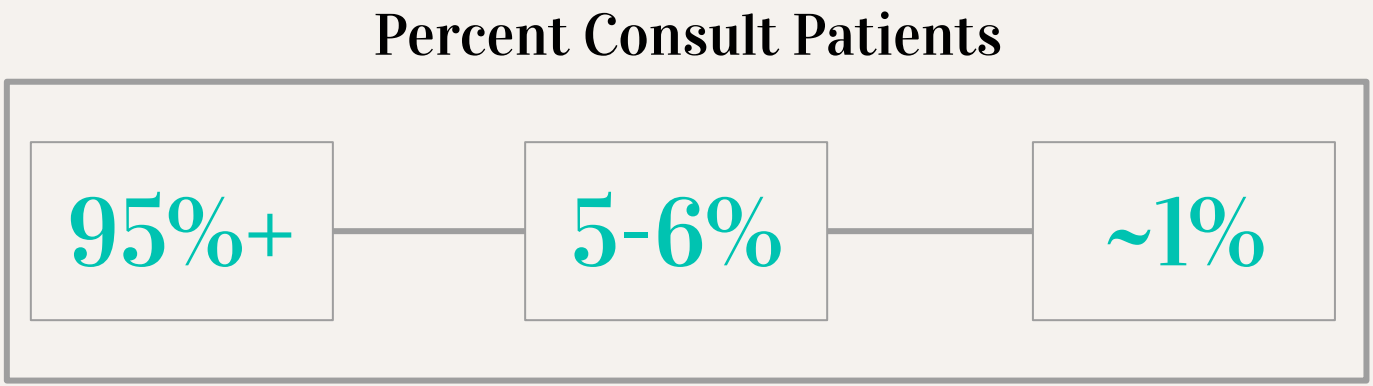
- **Training**: May, June, July 2018
- **Testing**: July & August 2018

Intervention Analysis

- **Cut down consult patients' LOS** by 10%, 20%, 30%, 40%, and 50%

*__Concept Drift__ = data distribution changed over time
(Focused on a shorter training & testing period)

Data (May-August 2018)



Patient Types:

T123 Admitted

~12% of total patients

T123 Not Admitted

66-68% of total patients

T45

20-22% of total patients

Train-Test Split Details

02

Analysis

Models, Experiments, and Evaluation

System States, Model Features, Model Evaluation

4 System States

System State 0	General NIS (1)
System State 1	NIS by Patient Type (3)
System State 2	NIS by Zone (11)
System State 3	NIS by Patient Type x Zone (33)

NIS = Number in System

Features

Features in the models:

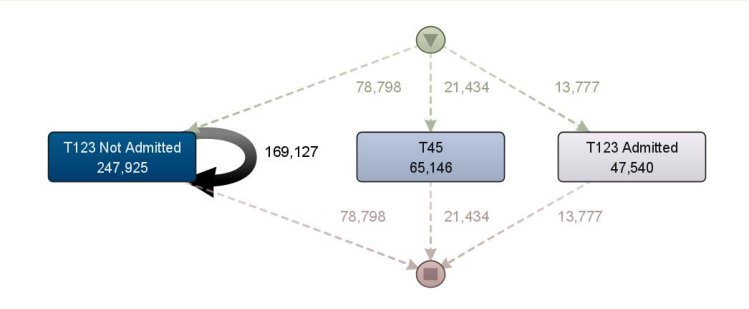
1. Patient Static Information*
2. Capacity Proxy*
3. System State X
4. Season, Trend, Holidays*

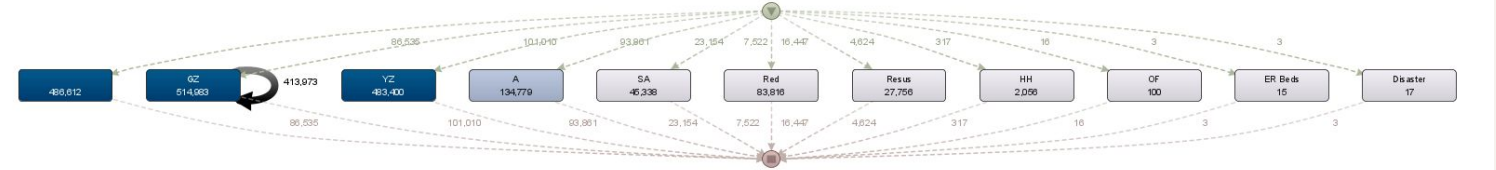
Model Evaluation

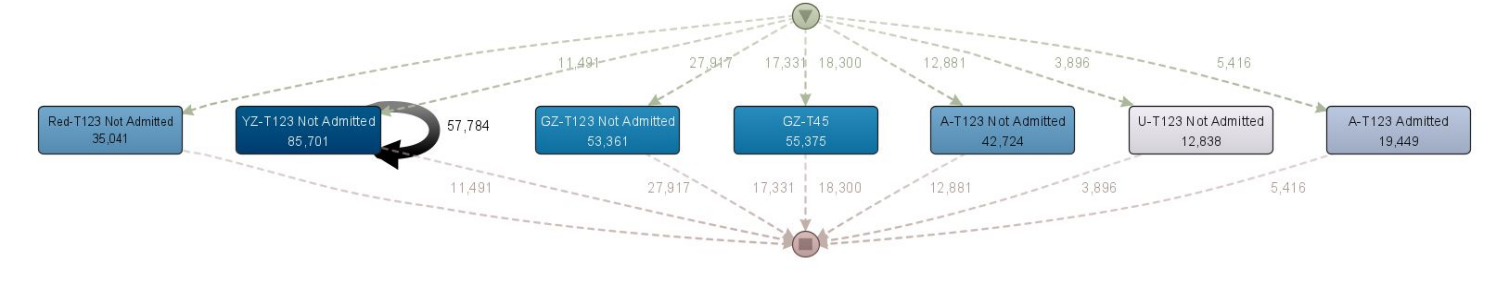
1. **Descriptive statistics** (mean & 90th percentile, histograms, Q-Q plots)
2. **Inference** (Confidence intervals, Kolmogorov–Smirnov Test)

***Patient static info:** age, gender, ambulance, consult, initial zone
***Capacity proxy:** arrival hour, arrival day of week
***Season** (arrival week number, arrival month), **Trend** (number of weeks since beginning of training data), **Holidays** (Ontario public holidays)

System States as Process Abstractions



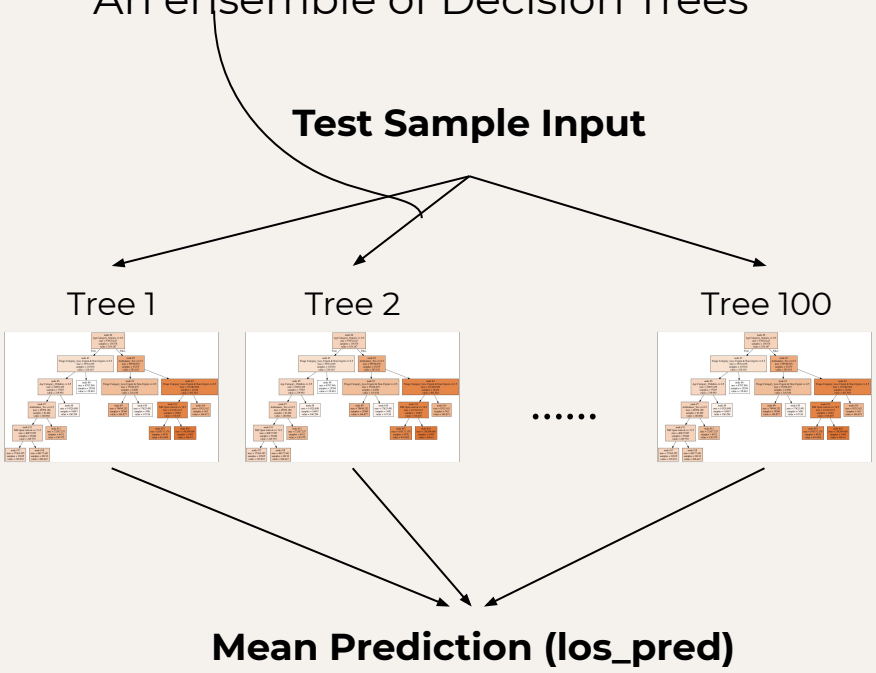




System State 1	NIS by Patient Type (3)
System State 2	NIS by Zone (11)
System State 3	NIS by Patient Type x Zone (33)

Modeling LOS: Random Forest Regression

An ensemble of Decision Trees



LOS Random Forest (RF) Regressor:

- **x** = Patient static info*, Capacity features*, System State X, Season, Trend, Holidays*
- **y** = **log** of patient's length of stay (LOS)
- **Hyperparameters**: 30 minimum samples in the leaf and 100 trees in the RF model

Sampling Method from RF Model:

- Store a bucket of training errors
- For each observation (in log minutes):
  - los_pred = RF mean prediction
  - Error = random sample from the bucket of training errors
- **Sampled LOS = exp(los_pred + error)**

Automated Data-Driven Modeling Pipeline

Tiered Process Map

Pipeline Start

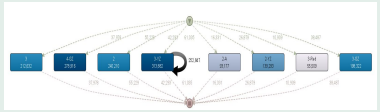
Get User Specification
- Context
- Process abstraction (system state)
- Interventions

Construct ED Model
- Automatically learned from the data
- Apply interventions

Apply Diagnostics & Analytics
- Diagnostics on goodness of fit
- Intervention impact

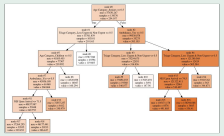
Pipeline End

Select System State



Gather All Model Training Features

Train LOS Model
- Random Forest



Simulate + Apply Interventions

03

Results

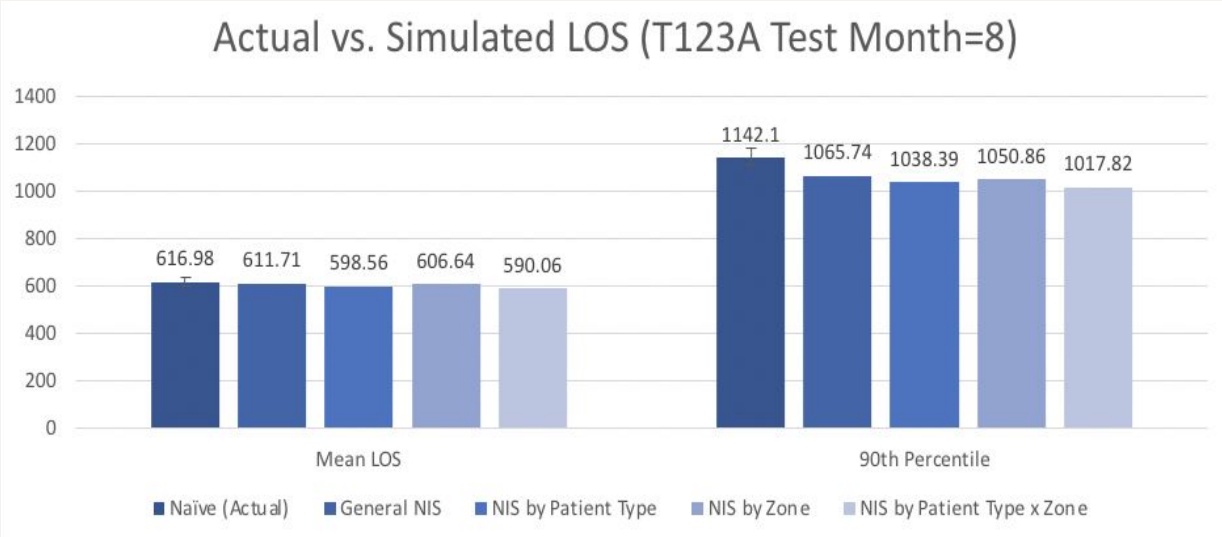
Descriptive Statistics and Inference

Mean & 90th Percentile LOS Predictions

August 2018

Actual vs. Simulated LOS, T123 Admitted

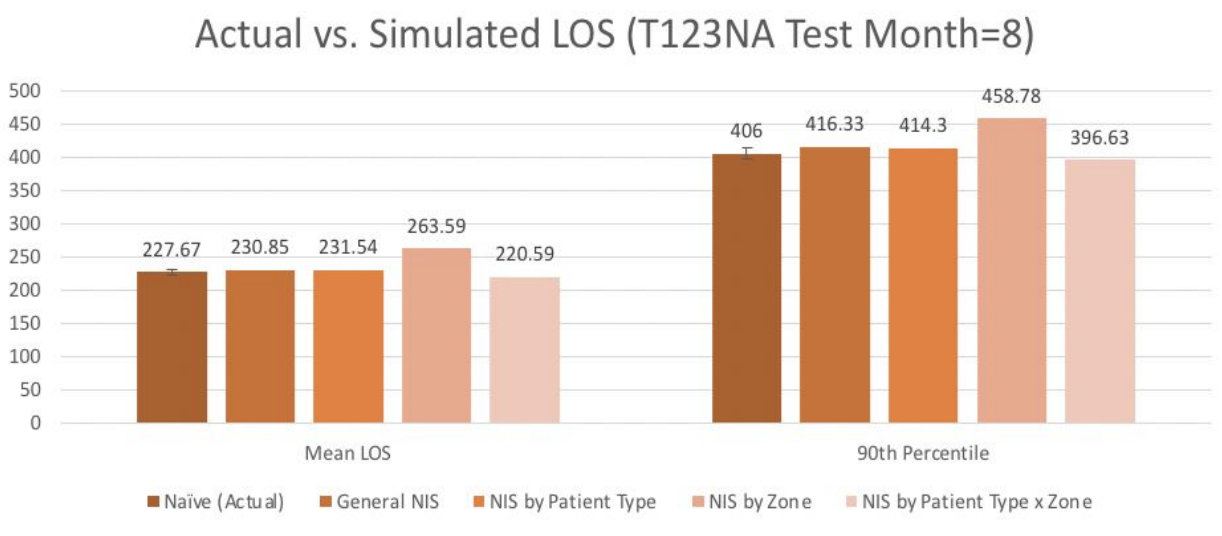
Goodness of Fit: Mean & 90th Percentile Predictions (in Minutes)



Mean 95% CI: (597.82, 636.13)
90th Percentile 95% CI: (1108, 1180)

Actual vs. Simulated LOS, T123 Not Admitted

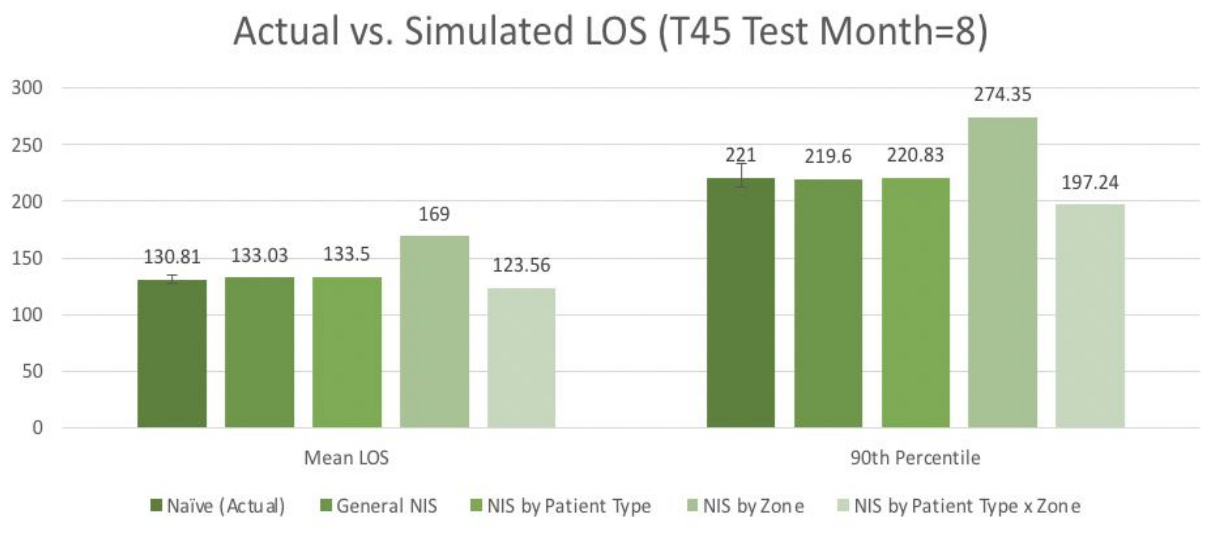
Goodness of Fit: Mean & 90th Percentile Predictions (in Minutes)



Mean 95% CI: (223.81, 231.6)
90th Percentile 95% CI: (399, 416)

Actual vs. Simulated LOS, T45

Goodness of Fit: Mean & 90th Percentile Predictions (in Minutes)



Mean 95% CI: (127.11, 134.5)
90th Percentile 95% CI: (213, 234)

Mean & 90th Percentile Predictions

X = Falling Within 95% Confidence Interval of Actual Mean & 90P

	System State 0	System State 1	System State 2	System State 3
T123A (Mean)	X	X	X	
T123NA (Mean)	X	X		
T45 (Mean)	X	X		
T123A (90P)				
T123NA (90P)		X		
T45 (90P)	X	X		

LOS Histograms & Q-Q Plots

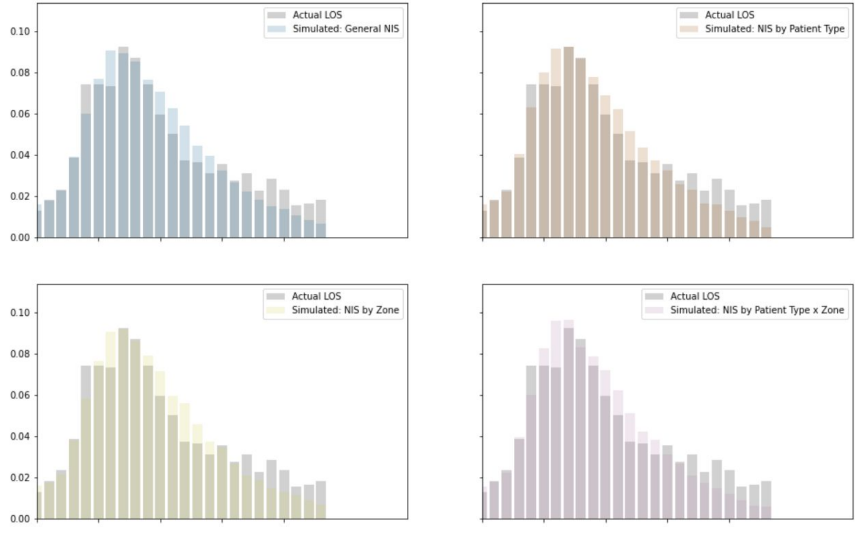
August 2018

System State 0	System State 1
System State 2	System State 3

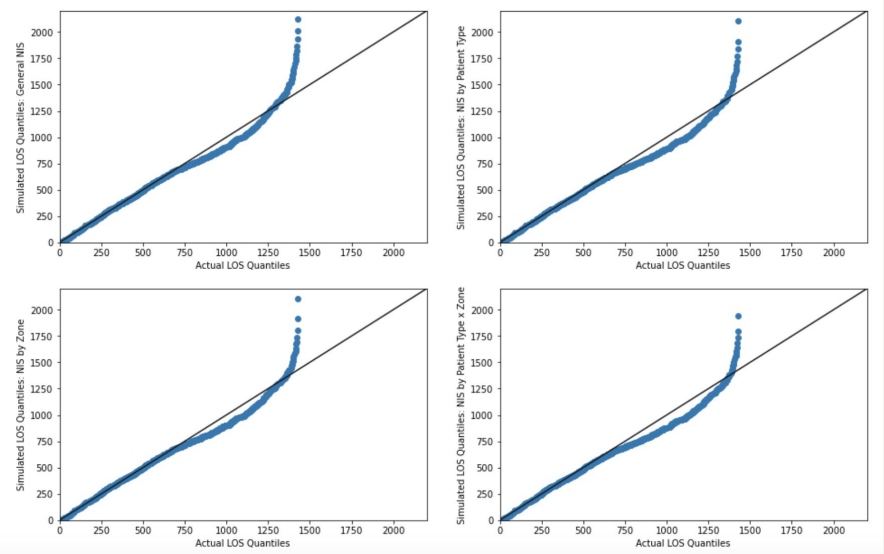
Patient Type: T123 Admitted

Goodness of Fit: LOS Distribution

Relative Frequency Histograms



Q-Q Plots

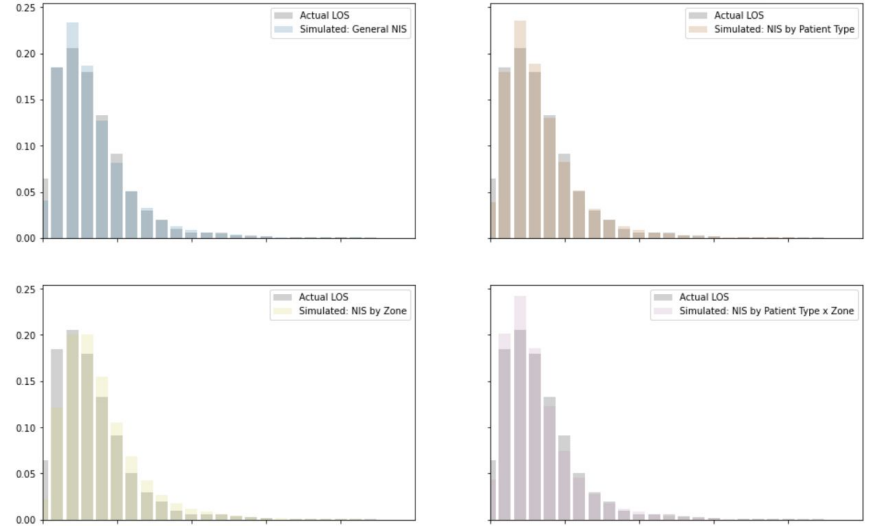


System State 0	System State 1
System State 2	System State 3

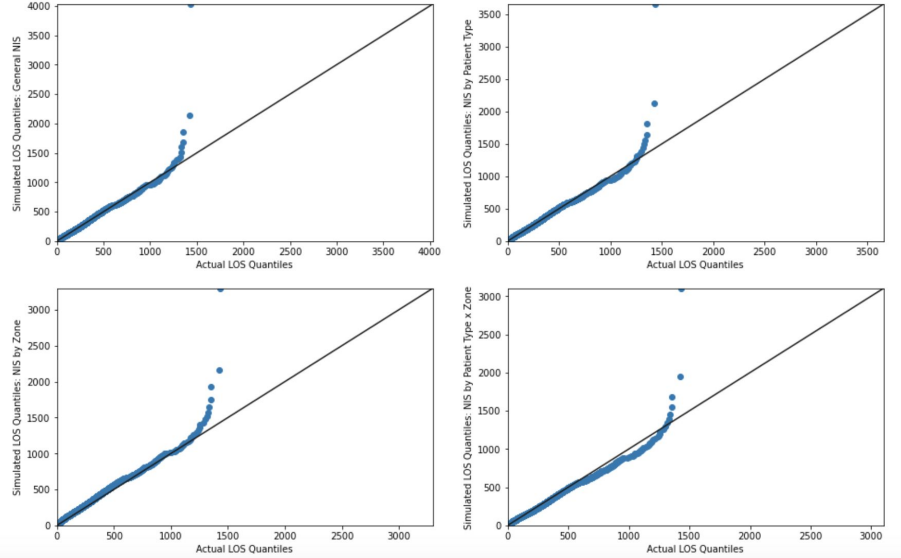
Patient Type: T123 Not Admitted

Goodness of Fit: LOS Distribution

Relative Frequency Histograms



Q-Q Plots

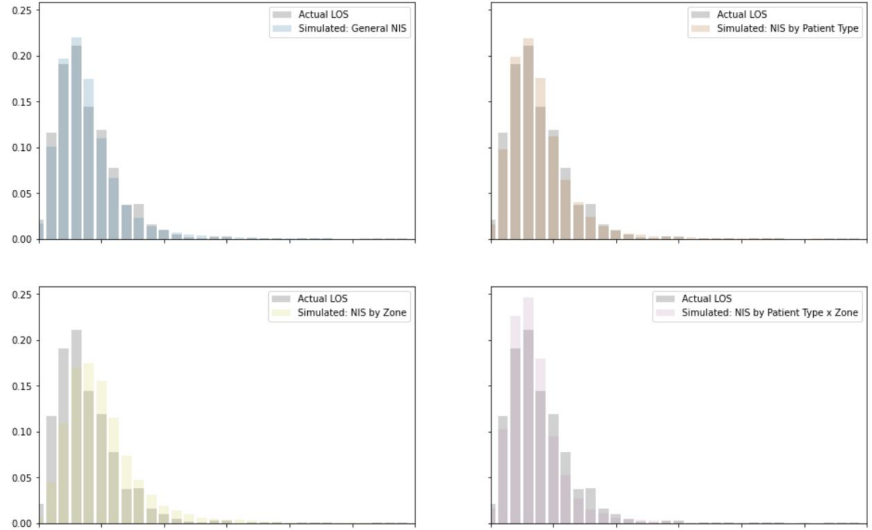


System State 0	System State 1
System State 2	System State 3

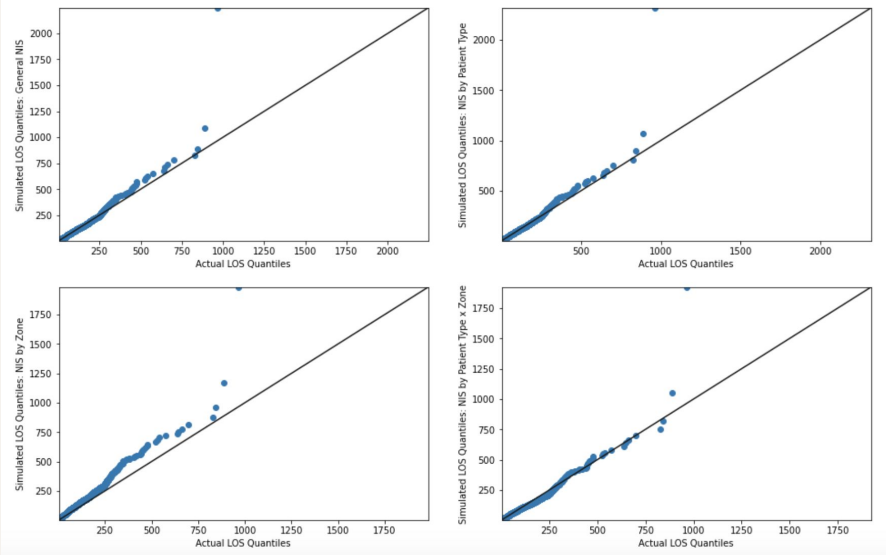
Patient Type: T45

Goodness of Fit: LOS Distribution

Relative Frequency Histograms



Q-Q Plots



Kolmogorov-Smirnov (KS) Test

August 2018

With Season, Trend, Holidays Features

KS Test Statistics

Goodness of Fit: LOS Distribution (Smaller KS test statistic suggests better fit)

	System State 0	System State 1	System State 2	System State 3
T123A	**0.0447**	0.0545	0.0496	0.0641
T123NA	**0.0379**	0.0429	0.118	0.0441
T45	**0.0315**	0.0351	0.217	0.0884

Best Models

	Mean & 90th percentile LOS predictions	Histograms and Q-Q Plots	KS Test
System State of the "Best" Model(s)	1 (NIS by Patient Type)	0 (general NIS) and 1 (NIS by Patient Type)	0 (general NIS)
"Best" in terms of	# of times predictions fall within 95% CI of actual data	Deviation from the 45-degree line	KS test statistic

Best Model: Model with System State 0 (General NIS)

Intervention Analysis

Cutting down LOS of consult patients by 10%, 20%, 30%, 40%, and 50%

August 2018

Intervention and Models

Patient Arrives

Baseline — "**Naive**" (Obtained From Data)

Sampled From Random Forest Models

1. **System State 0** (General NIS)
2. **System State 1** (NIS by Patient Type)
3. **System State 2** (NIS by Zone)
4. **System State 3** (NIS by Patient Type x Zone)

Patient's LOS

Is this a **Consult** patient? If yes...

Apply Intervention*

E.g., If intervention is 10% cutdown:
LOS = LOS x 0.9

***Interventions**: 10%, 20%, 30%, 40%, 50% Cutdown in Consult Patient's LOS

Mean & 90th Percentile LOS (Current Situation)

	Mean LOS			90th Percentile LOS		
	T123 Admitted	T123 Not Admitted	T45	T123 Admitted	T123 Not Admitted	T45
	Current Situation			Current Situation		
Baseline	10.28 hrs	3.79 hrs	2.18 hrs	19.04 hrs	6.77 hrs	3.68 hrs
General NIS Model	10.20 hrs	3.85 hrs	2.22 hrs	17.76 hrs	6.94 hrs	3.66 hrs

Mean & 90th Percentile LOS (10% Cut Down)

	Mean LOS			90th Percentile LOS		
	T123 Admitted	T123 Not Admitted	T45	T123 Admitted	T123 Not Admitted	T45
	Current Situation			**Current Situation**		
Baseline	10.28 hrs	3.79 hrs	2.18 hrs	19.04 hrs	6.77 hrs	3.68 hrs
General NIS Model	10.20 hrs	3.85 hrs	2.22 hrs	17.76 hrs	6.94 hrs	3.66 hrs
	Cut down consult patients' LOS by 10%			Cut down consult patients' LOS by 10%		
LOS Now	**9.20 hrs**	**3.76 hrs**	**2.15 hrs**	**16.00 hrs**	**6.76 hrs**	**3.55 hrs**
Total Time Reduced 1st and 2nd order	59.8 mins	5.2 mins	3.8 mins	105.8 mins	10.8 mins	6.5 mins
Est. 1st Order Reduction	59.1 mins	1.7 mins	0.6 mins	105.3 mins	3.6 mins	0.4 mins
Est. 2nd Order Reduction	0.7 mins	3.5 mins	3.2 mins	0.5 mins	7.2 mins	6.1 mins

Mean & 90th Percentile LOS (20% Cut Down)

	Mean LOS			90th Percentile LOS		
	T123 Admitted	T123 Not Admitted	T45	T123 Admitted	T123 Not Admitted	T45
	Current Situation			**Current Situation**		
Baseline	10.28 hrs	3.79 hrs	2.18 hrs	19.04 hrs	6.77 hrs	3.68 hrs
General NIS Model	10.20 hrs	3.85 hrs	2.22 hrs	17.76 hrs	6.94 hrs	3.66 hrs
	Cut down consult patients' LOS by 20%			Cut down consult patients' LOS by 20%		
LOS Now	**8.21 hrs**	**3.68 hrs**	**2.10 hrs**	**14.26 hrs**	**6.58 hrs**	**3.45 hrs**
Total Time Reduced 1st and 2nd order	119.4 mins	10.2 mins	7.2 mins	210.1 mins	21.5 mins	12.4 mins
Est. 1st Order Reduction	118.2 mins	3.4 mins	1.2 mins	208.8 mins	7.4 mins	0.8 mins
Est. 2nd Order Reduction	1.2 mins	6.8 mins	6.0 mins	1.3 mins	14.1 mins	11.6 mins

Mean & 90th Percentile LOS (30% Cut Down)

	Mean LOS			90th Percentile LOS		
	T123 Admitted	T123 Not Admitted	T45	T123 Admitted	T123 Not Admitted	T45
	Current Situation			**Current Situation**		
Baseline	10.28 hrs	3.79 hrs	2.18 hrs	19.04 hrs	6.77 hrs	3.68 hrs
General NIS Model	10.20 hrs	3.85 hrs	2.22 hrs	17.76 hrs	6.94 hrs	3.66 hrs
	Cut down consult patients' LOS by 30%			Cut down consult patients' LOS by 30%		
LOS Now	**7.21 hrs**	**3.60 hrs**	**2.05 hrs**	**12.53 hrs**	**6.41 hrs**	**3.38 hrs**
Total Time Reduced 1st and 2nd order	178.8 mins	14.9 mins	10.3 mins	313.9 mins	31.5 mins	16.6 mins
Est. 1st Order Reduction	177.3 mins	5.1 mins	1.7 mins	311.2 mins	11.0 mins	1.4 mins
Est. 2nd Order Reduction	1.5 mins	9.8 mins	8.6 mins	2.7 mins	20.5 mins	15.2 mins

Mean & 90th Percentile LOS (40% Cut Down)

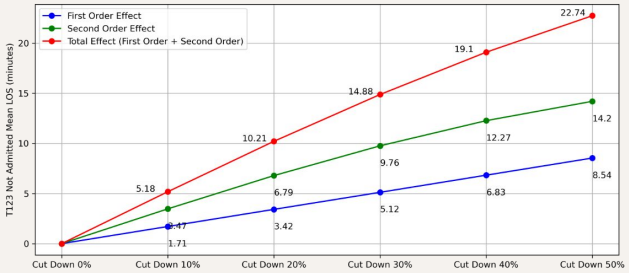
	Mean LOS			90th Percentile LOS		
	T123 Admitted	T123 Not Admitted	T45	T123 Admitted	T123 Not Admitted	T45
	Current Situation			**Current Situation**		
Baseline	10.28 hrs	3.79 hrs	2.18 hrs	19.04 hrs	6.77 hrs	3.68 hrs
General NIS Model	10.20 hrs	3.85 hrs	2.22 hrs	17.76 hrs	6.94 hrs	3.66 hrs
	Cut down consult patients' LOS by 40%			Cut down consult patients' LOS by 40%		
LOS Now	**6.22 hrs**	**3.53 hrs**	**2.00 hrs**	**10.81 hrs**	**6.26 hrs**	**3.31 hrs**
Total Time Reduced 1st and 2nd order	238.3 mins	19.1 mins	13.1 mins	417.1 mins	40.8 mins	21.1 mins
Est. 1st Order Reduction	236.3 mins	6.8 mins	2.3 mins	414.0 mins	14.3 mins	2.5 mins
Est. 2nd Order Reduction	2.0 mins	12.3 mins	10.8 mins	3.1 mins	26.5 mins	18.6 mins

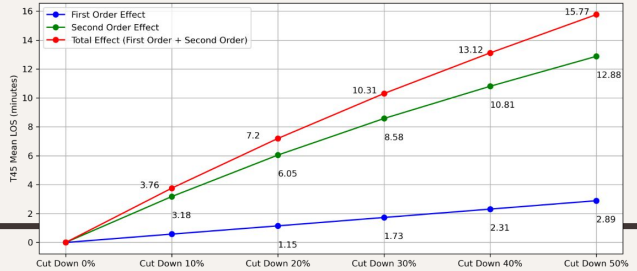
Mean & 90th Percentile LOS (50% Cut Down)

	Mean LOS			90th Percentile LOS		
	T123 Admitted	T123 Not Admitted	T45	T123 Admitted	T123 Not Admitted	T45
	Current Situation			**Current Situation**		
Baseline	10.28 hrs	3.79 hrs	2.18 hrs	19.04 hrs	6.77 hrs	3.68 hrs
General NIS Model	10.20 hrs	3.85 hrs	2.22 hrs	17.76 hrs	6.94 hrs	3.66 hrs
	Cut down consult patients' LOS by 50%			Cut down consult patients' LOS by 50%		
LOS Now	**5.24 hrs**	**3.47 hrs**	**1.95 hrs**	**9.10 hrs**	**6.13 hrs**	**3.23 hrs**
Total Time Reduced 1st and 2nd order	297.4 mins	22.7 mins	15.8 mins	519.5 mins	48.5 mins	26.1 mins
Est. 1st Order Reduction	295.4 mins	8.5 mins	2.9 mins	514.7 mins	17.3 mins	3.8 mins
Est. 2nd Order Reduction	2.0 mins	14.2 mins	12.9 mins	4.8 mins	31.2 mins	22.3 mins

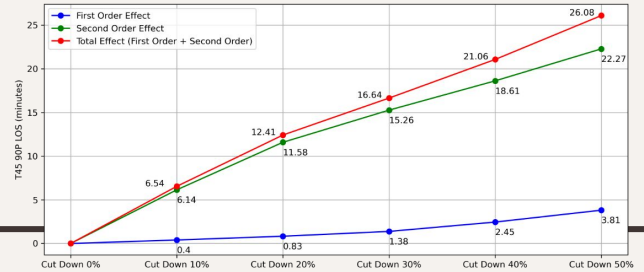
Reduction on LOS with Interventions

Mean LOS





90th Percentile LOS



04

Conclusion

Limitations and Possible Future Directions

Main Conclusions

- Our infinite-server model is able to capture **second order effects**
  - *We saw evidence of **"system state" contributing to congestion** in the system*

- Second order effects are observed in **T123 Not Admitted and T45 patients**
  - *Could be consequence of congestion alleviated from top priority patients*

- The cutdown in LOS of consult patients **will not significantly reduce** the length of stay beyond the first order effect
  - *T123 Admitted patients: **insignificant second order effect***
  - *T123 Not Admitted & T45 patients: **50% consult patients LOS cut down** results in **~12% reduction** in 90th percentile length of stay (with second order effects)*

Limitations & Future Directions

	Limitations	Future Directions
Data	• Data from **2018** (not very recent) • Training and testing data **limited** to a few months	• Consider trade-off between **recency vs. adequacy** of training and testing data
Arrival & Service Processes	• "**Black-box**" service process	• Generate **synthetic arrivals** into the ED system • Infer **resource capacity** (n servers) in simulation • Model LOS on **individual station-level**
Process Abstraction	• Representation of congestion ("system state") is **highly simplified**, by counts of patients in the ED	• Discover **alternative ways** to improve system state representation

Thank You

Do you have any questions?

Back Up

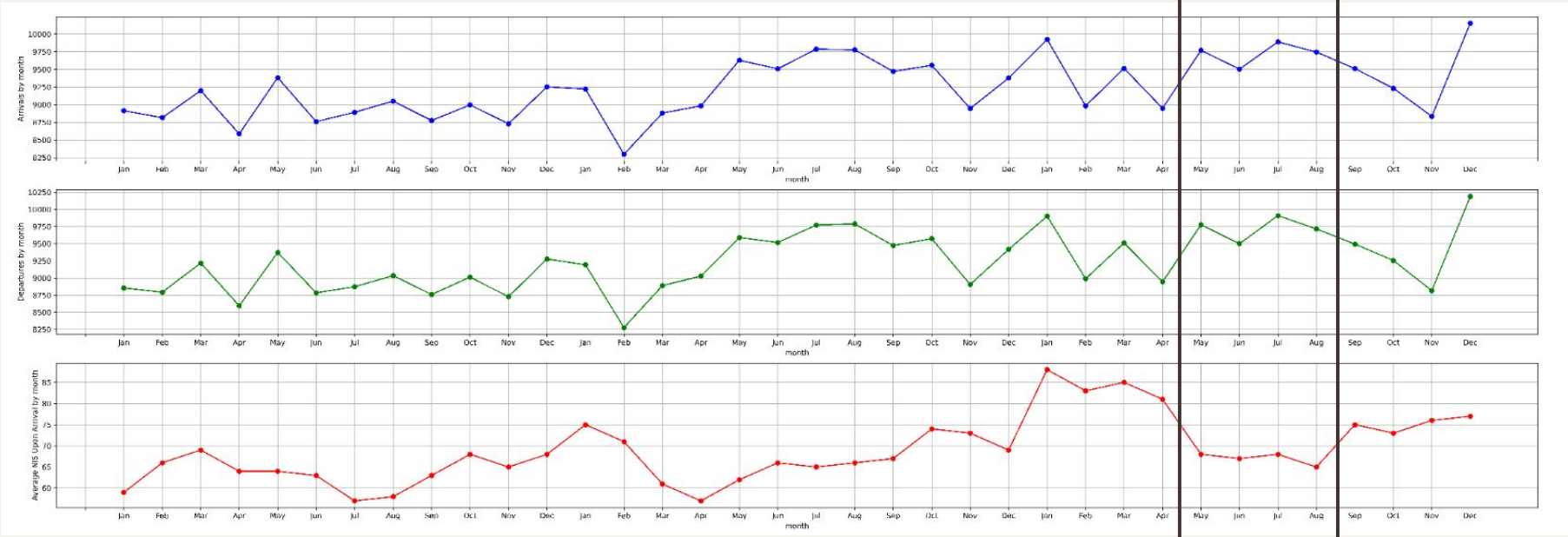
Concept Drift

The statistical properties (mean & 90th percentile) of the target variable (LOS) we are trying to model changes over time in unforeseen ways. Percent consult, a factor contributes significantly to LOS also changes over time.

	2016			2017			2018		
	T123A	T123NA	T45	T123A	T123NA	T45	T123A	T123NA	T45
Mean LOS (minutes)	629	225	123	632	226	127	641	242	147
90th Percentile LOS (minutes)	1198	406	219	1194	408	227	1225	428	248
Percent Consult (%)	96.87%	5.65%	1.74%	93.78%	5.21%	1.53%	96.36%	5.89%	1.62%

Back

Arrivals, Departures, Average NIS Upon Arrival (2016-2018 by Month) Increases Over the Years



May, Jun, Jul, Aug 2018
(Relatively Stable Counts)

Back

Data: Train-Test Split

Train-test Split

	Training (May, Jun, Jul)	In-sample Test (July)	Out-of-sample Test (August)
T123 Admitted	• 3,422 in total • 97.0% consults	• 1,171 in total • 97.8% consults	• 1,160 in total • 95.2% consults
T123 Not Admitted	• 19,281 in total • 6.17% consults	• 6,550 in total • 5.6% consults	• 6,645 in total • 5.3% consults
T45	• 6,458 in total • 1.33% consults	• 2,168 in total • 1.2% consults	• 1,939 in total • 1.2% consults

Back

Simulation Procedure

NYGH ED System Simulation Procedure

1. **Initialization**:
   a. Arrival times from test data (2018) and add to event calendar
   b. Trained LOS RF regressor model
   c. Current number of patients in the system (aka current NIS) = 0

2. **Running Simulation**:
   a. While (event calendar is not empty):
      i. Pop an event from the calendar
      ii. Event = arrival:
         1. Current NIS + 1
         2. **Assign LOS to the patient by sampling from LOS RF regressor**
         3. Departure time = arrival time + LOS
      iii. Event = departure:
         1. Current NIS - 1

If intervention: for consult patients,
LOS = LOS x (1 - % cut down)

- % cut down can be 10%, 20%, 30%, 40%, or 50%