Modeling Flight Delays and Cancellations at the National, Regional and Airport Levels in the United States

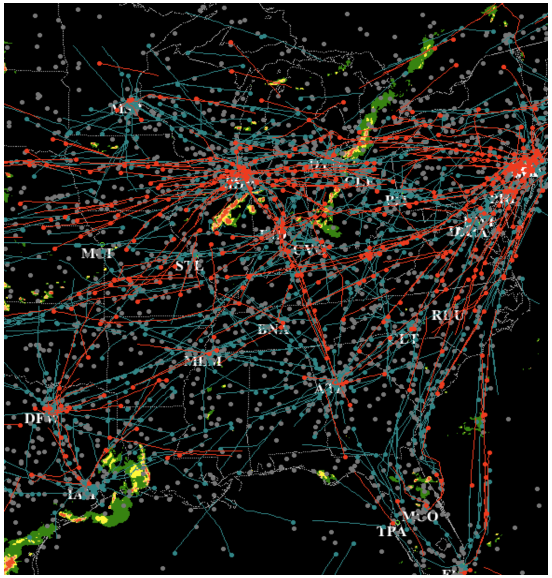
Banavar Sridhar
Yao Wang
Alex Klein
Richard Jehlen

National Airspace System Performance Workshop
Asilomar, CA
April 13-16, 2009

Motivation

- **Weather is the major cause of delay in the National Airspace System (NAS)**
- **Four possible scenarios**

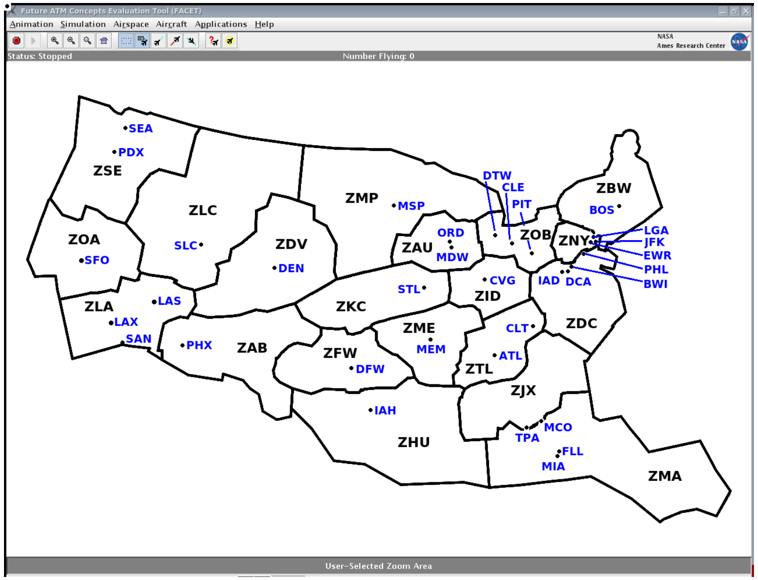
	Poor Weather Forecast	Accurate Weather Forecast
Poor operational response	X	X
Proper operational response	X	Best outcome



- **Relate delay, cancellations and other NAS performance metrics to the weather conditions to improve Traffic Flow Management**

Results

- Developed flight delay and cancellation models at the national, regional and airport levels
- Expected number of aircraft impacted by weather good proxy for delay
- Different models for summer and winter
- All metrics can be estimated to same level of accuracy
- FAA (ASPM and OPSNET) databases are complementary
- Neural Network models perform slightly better



Outline

- Objectives
- Databases
- Airspace Performance Metrics
- Modeling/Estimation of Metrics
  - Regression Models
  - Neural Network Models
- Results
- Conclusions

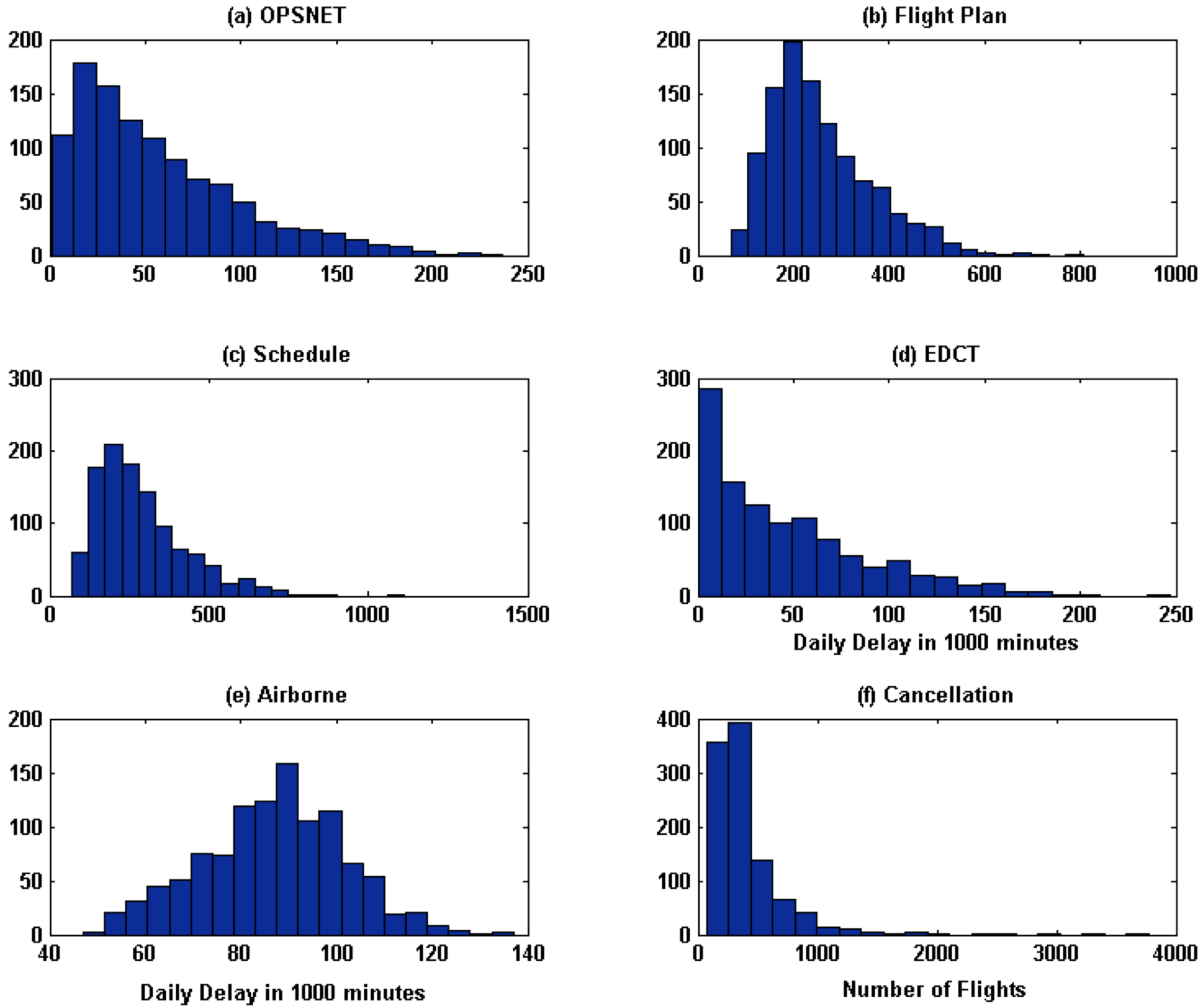
Objectives

- Develop NAS performance metric models based on FAA operational traffic databases
  - Different metrics
  - Impact of databases
  - Approach
    - Linear regression models
    - Neural networks models

Databases

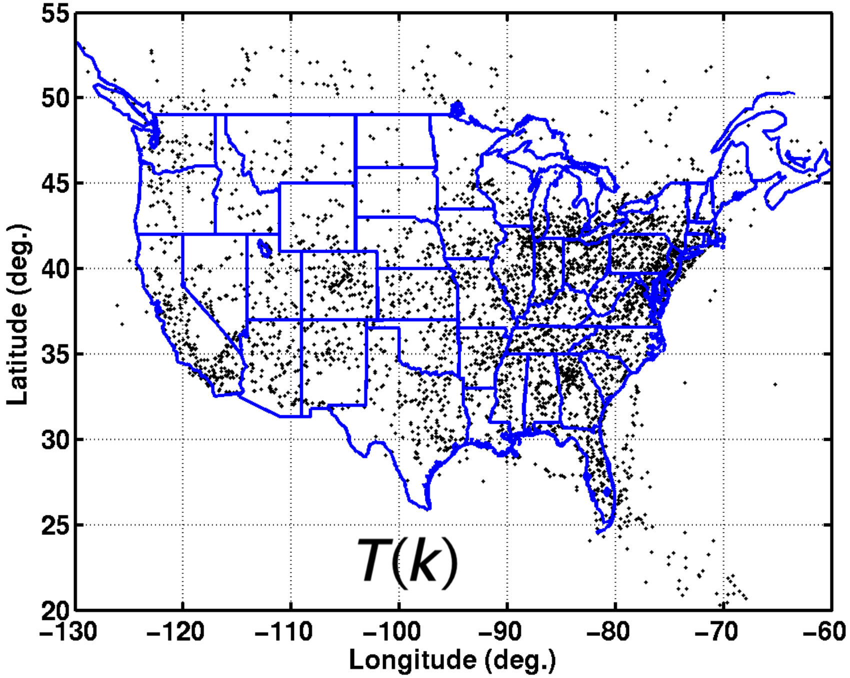
- FAA Operations Network (OPSNET)
  - Data available from 1990
  - Daily values
  - 45 airports
  - Total national delay
- Aviation System Performance Metrics (ASPM)
  - Data available from 2000
  - Every 15 minutes
  - 75 airports
  - Total OAG-based and flight-plan based arrival delays, EDCT hold minutes, airborne delay, flight cancellations
- Paper uses data from 2005-2008

NAS Performance Metrics

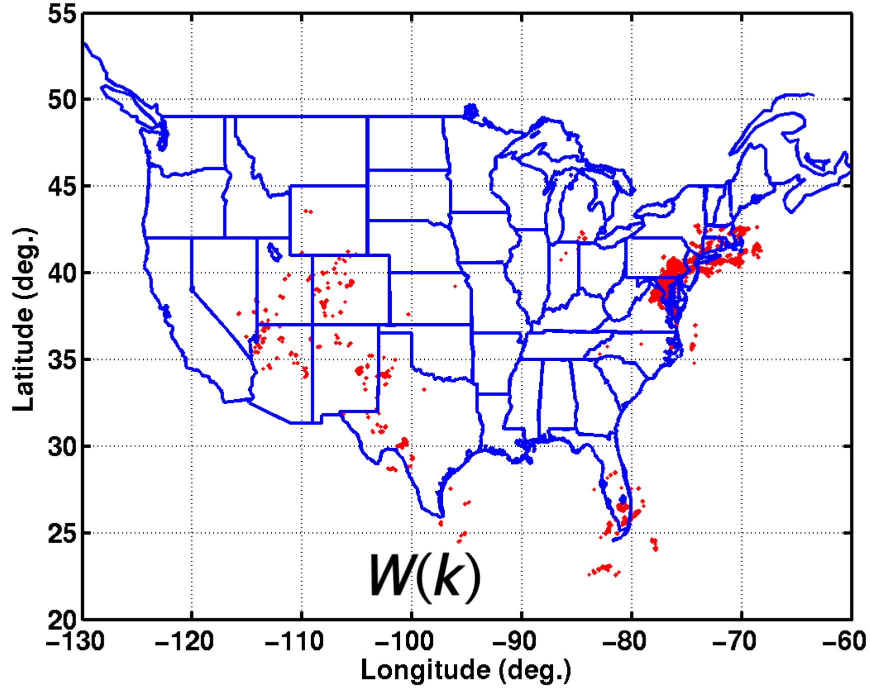


Weather Impacted Traffic Index (WITI)

Aircraft positions



Severe weather



$$WITI(k) = \sum_{1 \le j \le m} \sum_{1 \le i \le n} T_{i,j}(k) W_{i,j}(k)$$

Weather Impacted Traffic Index (WITI)

- Grid-based WITI
- National Weather Index (NWX)
  - En-route WITI (E-WITI), representing convective weather impact on major flows between city pairs
  - Terminal WITI (T-WITI), representing weather impact on major airports
  - Airport Queuing Delay (Q-Delay), representing surface and terminal-airspace weather impact on major airports in a non-linear fashion

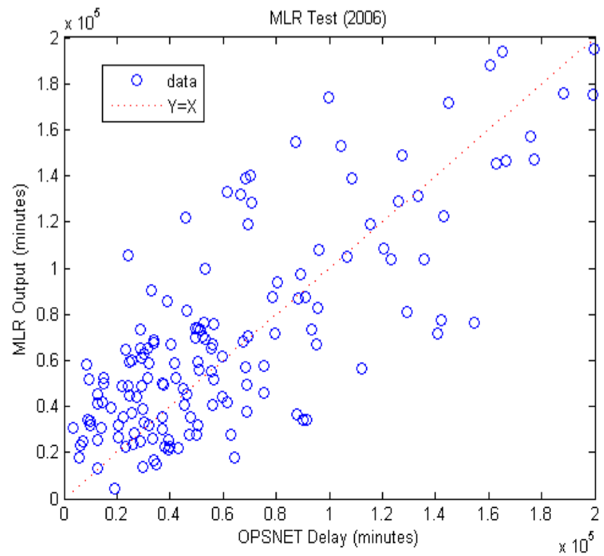
Modeling/Estimation of Metrics

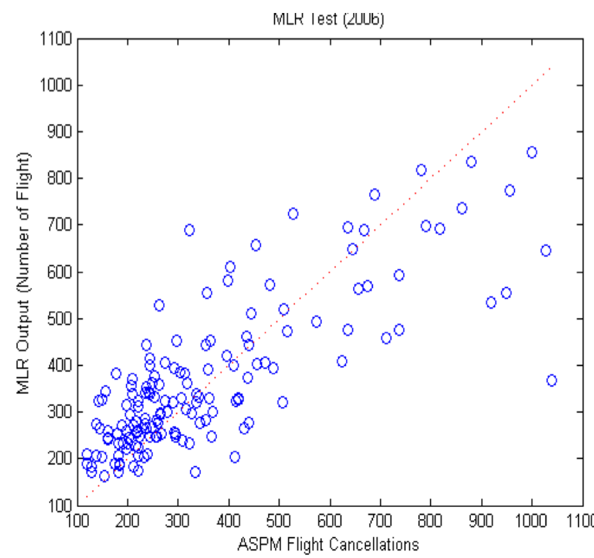
- Number of aircraft affected by weather (X)
- Number of aircraft affected by weather in each Center (X_p)
- Performance metric (δ)
- Models
  - Linear Regression (LR) $\delta = \alpha X + \beta$
  - Multiple Linear Regression (MLR) $$\delta = \sum_{p=1}^{20} \alpha_p X_p + \beta_p$$
  - Neural Networks $$\delta = f(X_p)$$
  - Dynamic Models $$\delta(t) = f(X_p(t-k), \ldots X_p(t-1), X_p(t), X_p(t+1) \ldots X_p(t+r))$$

Performance of Regression Models

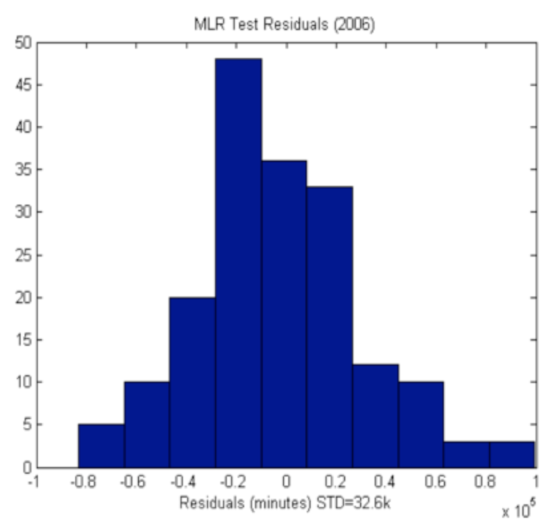
OPSNET Total Delay | **Flight Cancellations**

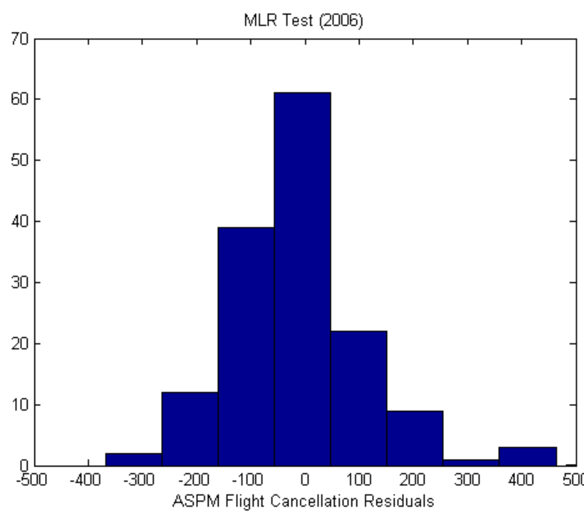
Estimate





Error

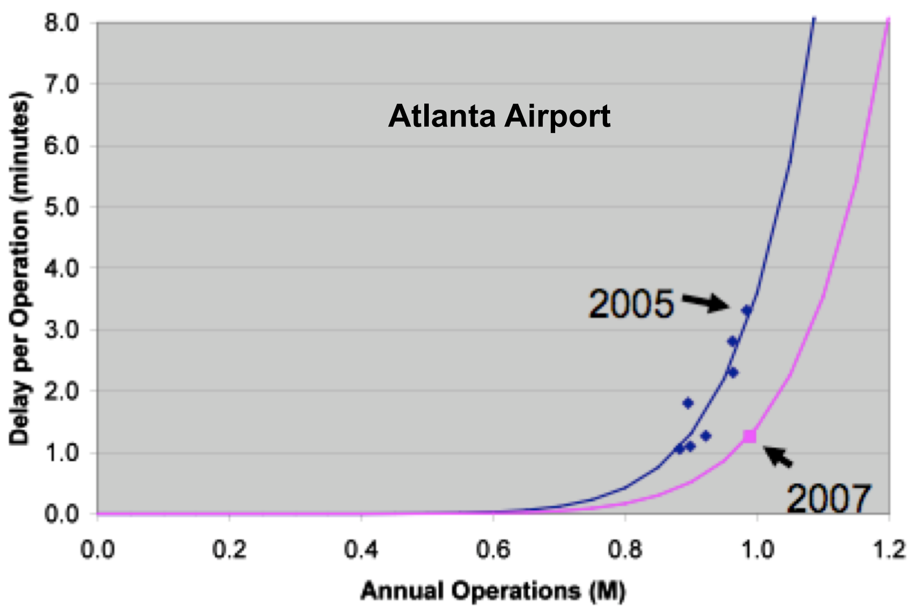




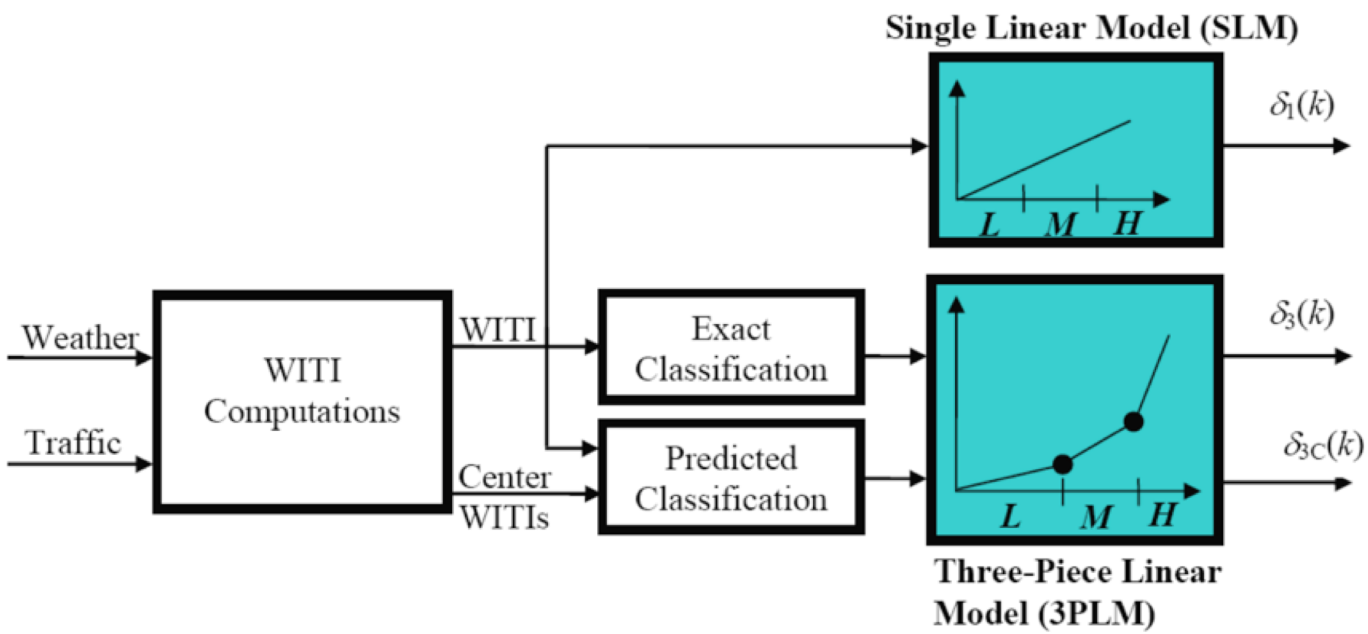
Performance of Regression Models

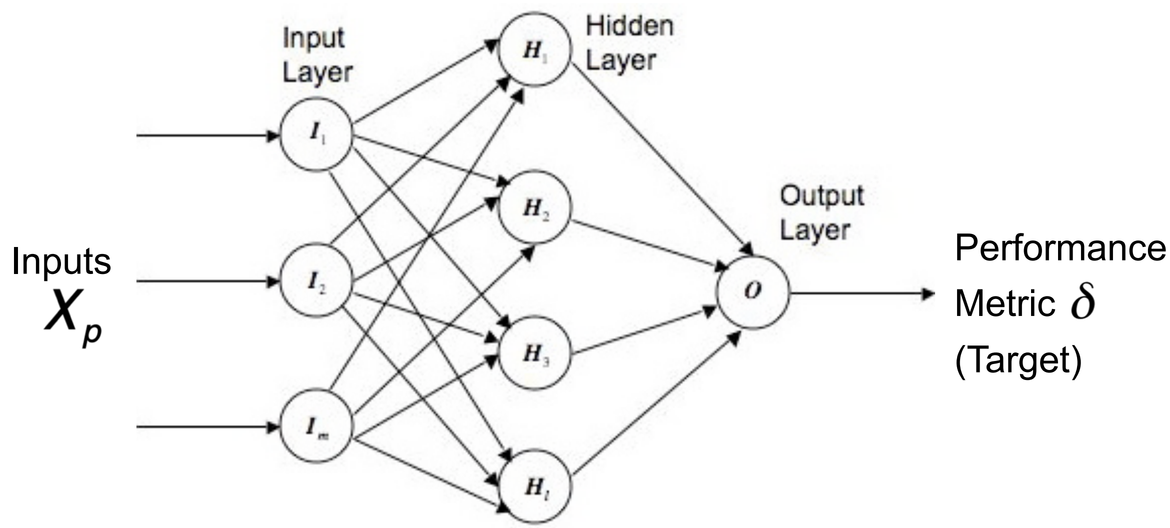
Type of Metric	Correlation Coefficient	Root Mean Squared Error	Maximum Absolute Error
OPSNET delay (LR)	0.71	32,700 minutes	26,600 minutes
OPSNET delay (MLR)	0.77	31,200 minutes	24,500 minutes
Scheduled delay	0.75	99,200 minutes	74,300 minutes
Flight Cancellations	0.77	131 flights	94 flights

- Regression models perform a good job of accounting for the impact of weather on delays and flight cancellations
- For systems with demand-capacity imbalance, growth in delay is non-linear



Nonlinear Models





Performance of National Model

Table 4.1 Performance of OPSNET national delay models

Type of Model	*r*	RMSE	MAE
LR	.71	32,700 minutes	26,600 minutes
MLR	.77	31,200 minutes	24,500 minutes
Neural Network	.80	30,000 minutes	23,300 minutes
Neural Network (5C)	.80	29,100 minutes	22,000 minutes

Table 4.3 Performance of ASPM flight cancellation models

Type of Model	*r*	RMSE	MAE
LR	.73	146 flights	106 flights
MLR	.77	131 flights	94 flights
Neural Network	.79	131 flights	93 flights
Neural Network (5C)	.79	139 flights	97 flights

- Neural Network models perform slightly better

Performance of different WITI definitions

Table 7. Correlation between performance metrics and WITI definition

	Airport Delay	OPSNET Delay	Flight-plan Delay	Schedule Delay	Flight Cancellation
E-WITI (LR)	.77	.70	.73	.76	.68
Grid WITI (LR)	.76	.72	.76	.76	.71
NWX (LR)	.84	.80	.84	.84	.74
E-WITI (MLR)	.80	.75	.78	.80	.78
Grid WITI (MLR)	.83	.80	.82	.83	.80
NWX (MLR)	.88	.84	.88	.88	.82

- Models using NWX perform slightly better
- Difference not significant while using MLR or NN

Seasonal performance of national delay models

Table 6. Seasonal Performance of national delay models

Type of Model	Summer Correlation		Winter correlation	
	Training	Test	Training	Test
MLR	.85	.75	.74	.71
Neural Network	.85	.76	.75	.72

- Higher correlation during summer
- Lower correlation in winter may be due to higher number of cancellations on days with heavy snow, very low ceilings/visibility

Airport delay models using Regression analysis

• 34 major airports on the OEP-list

Airport	γ_{LR}	γ_{MLR}	$\gamma_{MLR} / \gamma_{LR}$ - 1
ORD	0.743	0.803	0.08
ATL	0.752	0.777	0.03
EWR	0.640	0.725	0.13
PHL	0.764	0.805	0.06
DFW	0.577	0.646	0.12
JFK	0.618	0.670	0.08
LGA	0.685	0.723	0.06
LAX	0.195	0.496	1.54
IAH	0.684	0.725	0.06
DEN	0.550	0.664	0.21

• Good delay estimates for ORD, ATL,..
• Delay at ten airports in Eastern U.S not influenced by NWX in the neighboring Centers

Behavior of airport models

Table 10. Behavior of airport delay models (Training 2005-2007; Testing 2008)

Airport	Training γ_{MLR}	Testing γ_{MLR}	Training γ_{NN}	Testing γ_{NN}
ORD	.80	.79	.83	.80
ATL	.79	.72	.80	.72
EWR	.74	.64	.76	.68
PHL	.81	.78	.83	.80
DFW	.65	.60	.69	.63
JFK	.67	.67	.72	.68
LGA	.74	.64	.77	.67
LAX	.54	.34	.55	.35
IAH	.73	.72	.74	.72
DEN	.66	.68	.67	.68

Default Customer View

APF Dashboard

For Wednesday, August 13, 2008

Welcome Joe User

Today is Thursday, August 14, 2008

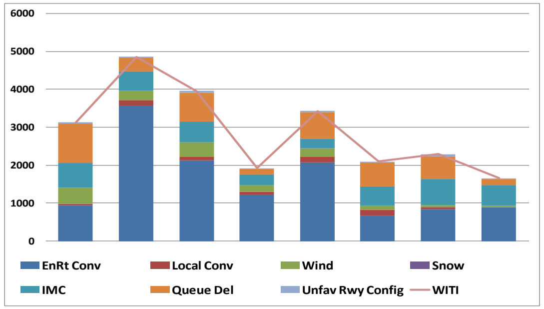
Print | Help | Log Out | Customize

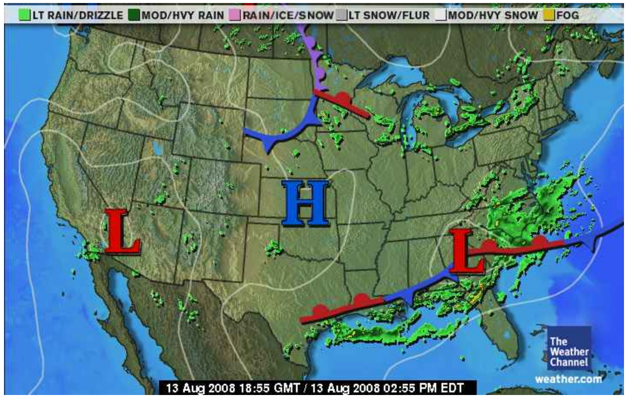
Customer View | ATO View

NAS | Service Area | ARTCC | Airport

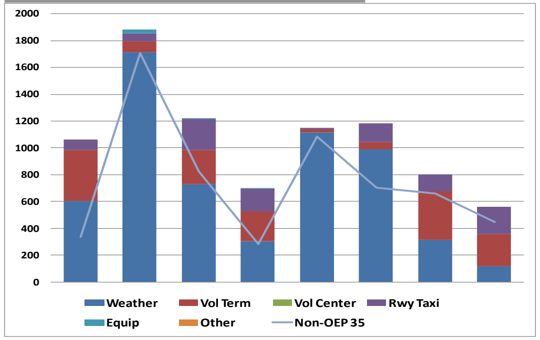
Days 1 3 8 30

WITI Last 8 Days

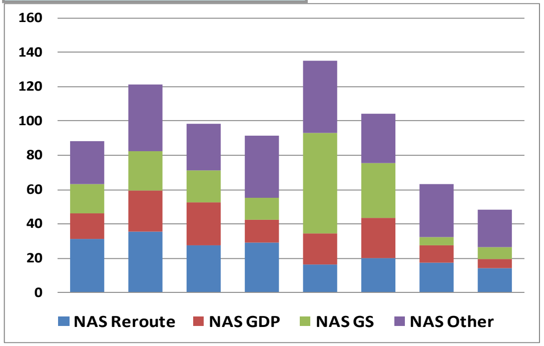




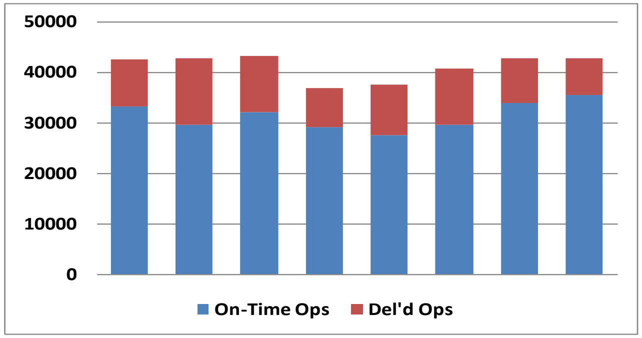
OPSNET Delays Last 8 Days



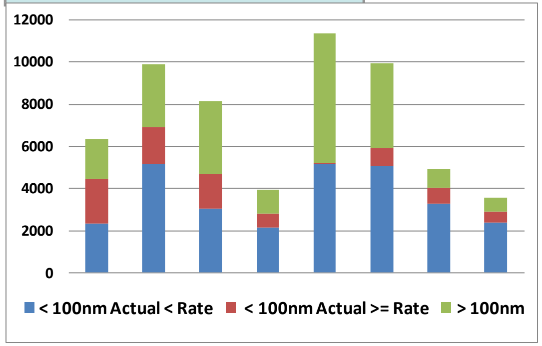
Advisories Last 8 Days



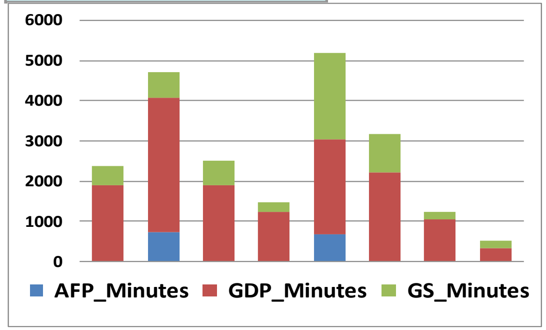
OPS – Delayed & On-Time Gate Arrivals Last 8 Days



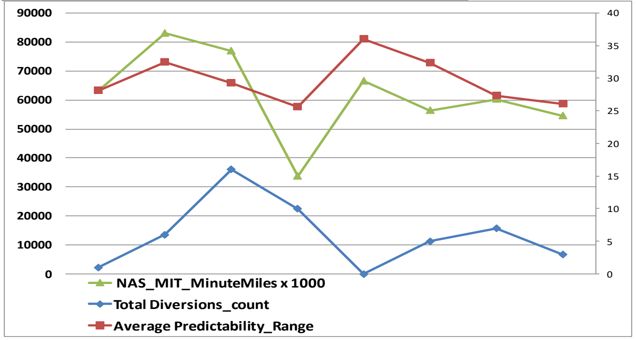
Holding Minutes Last 8 Days



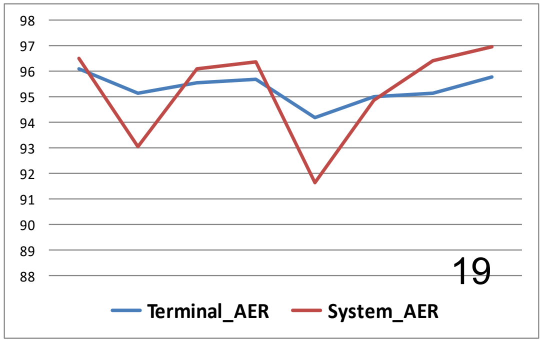
TMI Minutes Last 8 Days



Diverts, MIT, Predictability Last 8 Days



Efficiency Last 8 Days



Concluding Remarks

- Estimation/Modeling of performance metrics resulting from the use the two databases are comparable
- For all metrics, neural networks produce higher correlation and reduced errors than regression methods
- Different methods of reducing neural network complexity produce similar results