Accuracy in Predicting GDPs and Airport Delays from Weather Forecast Data

David Smith
Center for Air Transportation Systems Research
George Mason University
Fairfax, VA

September 6, 2007

CENTER FOR AIR TRANSPORTATION SYSTEMS RESEARCH

GEORGE MASON UNIVERSITY

Acknowledgements

- Funding
  - US Army
  - GMU/CATSR Research Foundation
- Technical Assistance
  - Ved Sud (FAA)
  - Mike Brennan (MAI)
  - Mark Klopferstein (MAI)
  - Dave Rodenhuis (University of Victoria)

Story

- Determining GDPs is difficult due to:
  - Uncertainty
  - Evolving nature
- How accurate can automation be when designed to predict uncertain systems
  - GDPs
  - Delays
- How good is good enough?

Agenda

- Motivation
- Background
  - TAF
  - Tool Requirements
- Method
- Tool Results: Accuracy of Predictions
- Conclusions

Forecasting Impact of Weather in Air Traffic Flow Management

- Forecasting is “too hard”
- Assumption – conditions are too stochastic
- Too many possibilities
- Plans fall apart
- Forecasts are too inaccurate

Forecasting & Military Planning

- The combat environment is one of the most chaotic
  - Fear
  - Morale
- NOT lock-step
- Leads to decision points
- Decision points lead to a finite set of options

Decision Points

- A decision point is a step in the planning process that, once determined, gives direction for a set of specific planning details
- Each decision point is followed by a list of planning questions and issues to be answered in the planning process
- Options are reduced to a manageable finite number
- Intelligence tools aid us to determine these points

Decision Point Example

- Delay forecast tool indicates convective weather over Philadelphia at 2300 Zulu
- Determine effects
  - Arrival delays
  - Reduce airport capacity
- Generate set of branch plans
  - MIT for Washington ARTCC
  - FCA for Cleveland ARTCC boundary
  - GDP at PHL
  - Do nothing

Intelligence

- What is the enemy to traffic flow?
  - Bad weather
  - Scheduled congestion
- Simulation techniques can be used to predict congestion
- Weather
  - Stochastic
  - Forecasts
    - Time
    - Chance
- Forecast gives us a decision point – time and place
- How do we create a branch plan?
  - Predict effect on operations – GDP, AAR ect.
  - Formulate a plan to counter the effects
- Solution – use Terminal Aerodrome Forecast (TAF) to predict effect on operations

What is the TAF?

- TAF - a concise statement of the expected meteorological conditions at an airport during a specified period (usually 24 hours)
- A TAF report contains the following sequence of elements in the following order:
  - Type of Report
  - ICAO Station Identifier
  - Date and Time of Origin
  - Valid Period Date and Time
  - Forecast Meteorological Conditions
  - Written in TAF “code”

TAF Example

TAF KEWR 161732Z 161818 24017G27KT P6SM SCT040 BKN250
FM1930 29018G32KT 4SM TSRA BR BKN040CB
FM2200 22009KT 6SM SHRA BR OVC040CB
FM0400 33006KT 6SM -SHRA BR OVC040
FM0800 34006KT 6SM BR OVC040
FM1400 26005KT P6SM BKN040

TAF Prediction Tool

- Use TAF to predict
  - Delays
  - GDP’s
  - AAR’s
- Tool provides a chance of occurrence of GDP
- Traffic flow managers develop a plan based on the prediction and the chance of it happening

Prediction Tool

Newark Airport TAF Delay Predictor

Enter each element of the TAF into a separate cell. Remove all plus (+), minus(-), and equal (=) signs

Station						
KEWR	010537Z	10606	30012KT	P6SM	SCT060	
	FM1100	27010KT	P6SM	SCT150	SCT250	
	FM1400	23010KT	P6SM	SCT060	BKN100	BKN250
	TEMPO	1618	BKN035			
	FM1800	18012KT	5SM	RA	OVC025	
	TEMPO	1819	2SM	SNRA	BR	BKN015
	FM2200	20012KT	5SM	RA	SCT020	OVC035
	FM0200	24005KT	6SM	RA	BKN035	OVC060
	TEMPO	205	5SM	RA	BR	BKN025

Input TAF

Time period delay

Input day of the week

Average Daily Delay Forecast

There is a 76% chance that delays WILL exceed 45 MINUTES

GDP Forecast

There is a 73% chance that there WILL NOT be a GDP at EWR

Estimated Delay	
1100 Zulu	3 minutes
1500 Zulu	10 minutes
1900 Zulu	43 minutes
2300 Zulu	88 minutes

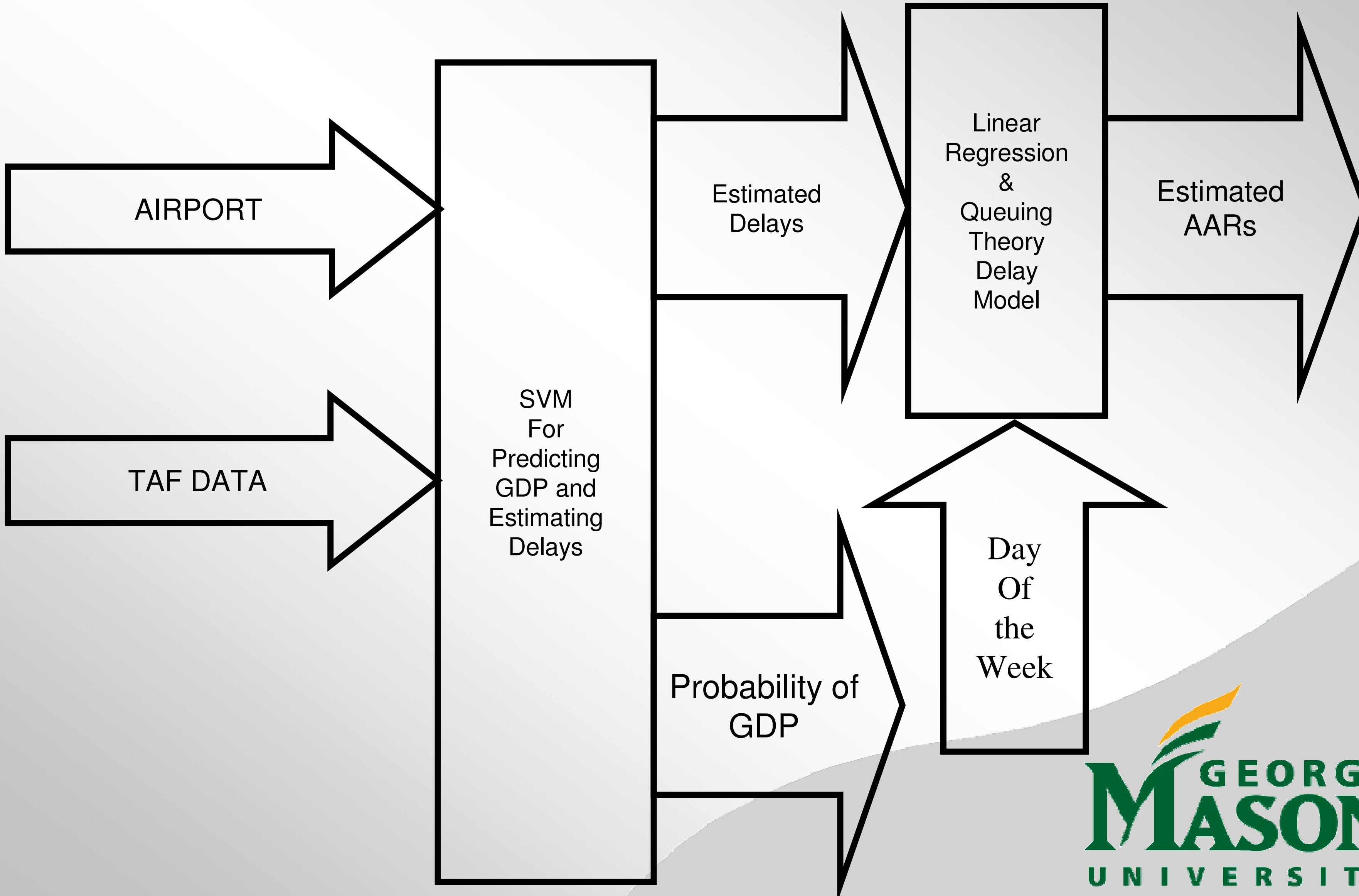
Estimated Arrival Rate		
Choose Day of the Week		Tuesday
1100 Zulu	35	
1500 Zulu	30	
1900 Zulu	27	
2300 Zulu	27	

AAR Forecast

GDP Forecast

Daily delay forecast

Method Summary



Method

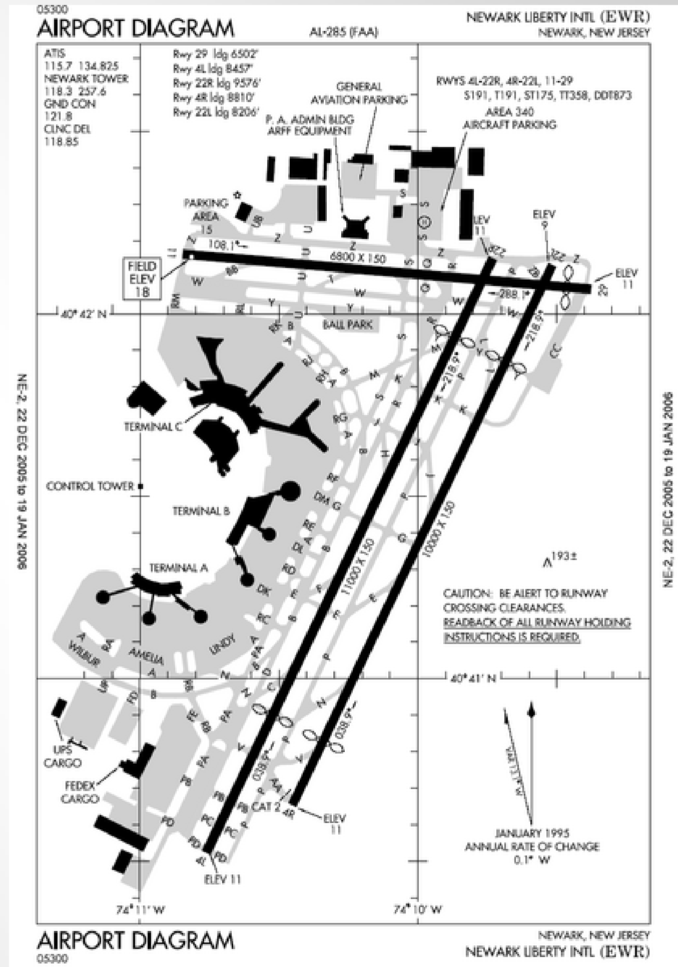
- Convert TAF format to a vector data set
- Use a pattern recognition tool called the support vector machine (SVM)
- SVM is trained with past data
- Develop functions to present day data to predict outcomes
- SVM does not require a linear relationship
- Use individual data points with few possible choices to product predictive functions

Extracted Data

- Time periods
  - 1100 Zulu
  - 1500 Zulu
  - 1900 Zulu
  - 2300 Zulu
- Wind speed
- Visibility
- Ceiling
  - Overcast
  - Scattered
  - Broken
  - Few
- **Cross Wind**
- Binary variables
  - Rain
  - Snow
  - Showers
  - Thunderstorms
  - Fog
  - Mist
  - Freezing

Crosswind

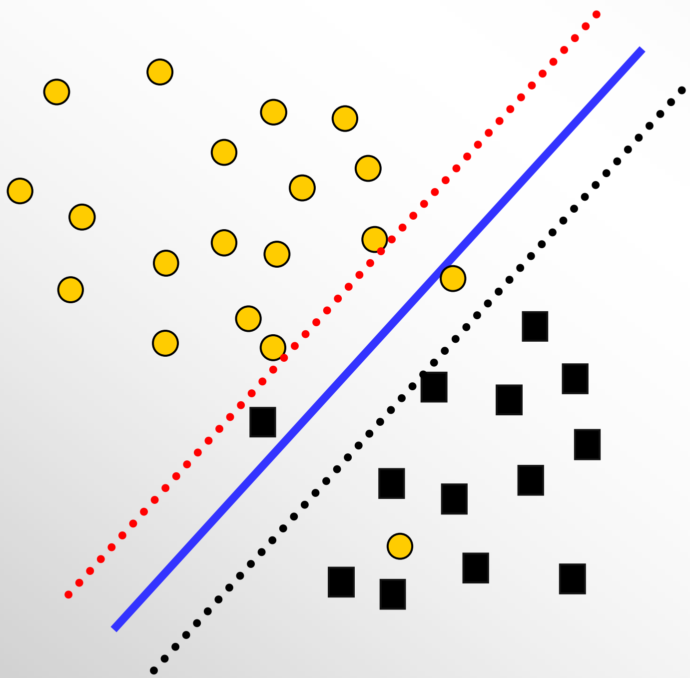
- Newark Airport
- Primary 4 – 22 R/L
- Crosswind runway 29/11
- Direction and wind is the deciding factor
- Assumed crosswind anywhere between 270 and 350 and between 170 and 90



Vector Data Set

Date	Wind Speed	Visibility (mi)	Rain (Y/N)	Snow (Y/N)	Showers (Y/N)	Thunder storms (Y/N)	Fog (Y/N)	Mist (Y/N)	Freezing (Y/N)	Overcast Ceiling	Scattered Ceiling	Broken Ceiling	Few Ceiling	Cross Winds
1	7	6	0	0	0	0	0	0	0				150	0
2	12	6	0	0	0	0	0	0	0		250			0
3	6	6	0	0	0	0	0	0	0			150		0
4	5	5	0	0	0	0	0	0	0			120	60	0
5	6	4	1	0	1	0	0	1	0			60		0
6	8	6	0	0	1	0	0	0	0	60	25	35		0
7	9	6	0	0	0	0	0	0	0		90			1
8	5	6	0	0	0	0	0	0	0			250		0
9	5	6	0	0	0	0	0	0	0		120	250		0
10	6	6	0	0	0	0	0	0	0		100	250		0
11	10	6	0	0	0	0	0	1	0					0
12	6	4	0	0	0	0	0	0	0		15			0
13	11	6	0	0	0	0	0	0	0		50	80		1
14	4	6	0	0	0	0	0	0	0					0
15	3	6	0	0	0	0	0	1	0			250		0
16	4	6	0	0	0	0	0	0	0					0
17	4	6	0	0	0	0	0	0	0					1
18	7	6	0	0	0	0	0	0	0		250			0
19	6	6	0	0	0	0	0	0	0		40	100		0
20	10	6	0	0	0	0	0	1	0			20		0
21	5	5	0	0	0	0	0	1	0					0
22	8	2	0	0	0	0	0	1	0			15		0
23	12	6	0	0	0	0	0	1	0		25			0
24	5	6	0	0	0	0	0	0	0					1
25	6	6	0	0	0	0	0	0	0		250			0
26	5	6	0	0	0	0	0	0	0			250		0
27	7	6	0	0	0	0	0	0	0		250			0
28	6	6	0	0	0	0	0	0	0			90	5	0
29	8	6	0	0	0	0	0	0	0					1
30	4	6	0	0	0	0	0	0	0		50	100		1
31	7	6	0	0	0	0	0	0	0		250			0

Linearly Inseparable Case: Supporting Plane Method



Just add non-negative error vector ***z***.

$$\min_{w,b,z} \quad \frac{1}{2}\|w\|^2 + C\sum_{i=1}^{N} z_i$$

$$s.t. \quad y_i\left(x_i^T w + b\right) + z_i \geq 1$$

$$z_i \geq 0 \quad i = 1,..,l$$

Data Collection

- TAF Data from National Climatic Data Center
  - Use Excel macro to convert to linear data
  - Used data from 2002 and 2006
- Bureau of Transportation Statistics
  - Excel macro retrieved average daily delay for each airport
  - Initial look is at Newark, O'Hare, Atlanta, and Philadelphia
- Ground Delay Program Data from Collaborative Decision Making (CDM) website

Predicting GDP

- Collect TAF data as the independent variable
- Quadratic program loaded into AMPL
- TAF data is transformed into the integer vector x
- GDP indicator variable y
  - -1 indicates no GDP
  - 1 indicates a GDP
- Program output
  - Solution w vector
  - Linear y-intercept vector b
- Prediction equation $w_i^T x_i + b$

Predicting Delay

- Delay data is not binary
- Several SVM runs
  - Whether or not there will be a x minute delay
  - x minute intervals from y to z minutes
- Creates 4 predictor functions

Predicting Arrival Rates

- M/M/1 queuing system
- Delay (D) is the average time the customer spends in the system
- Arrival rate per unit time (λ) is estimated by the total number of flights divided by 16 operating hours
- Aircraft arrival rate is the customer service rate (μ)

$$\mu = \frac{1}{D} + \lambda$$

GDP Results (Training Data)

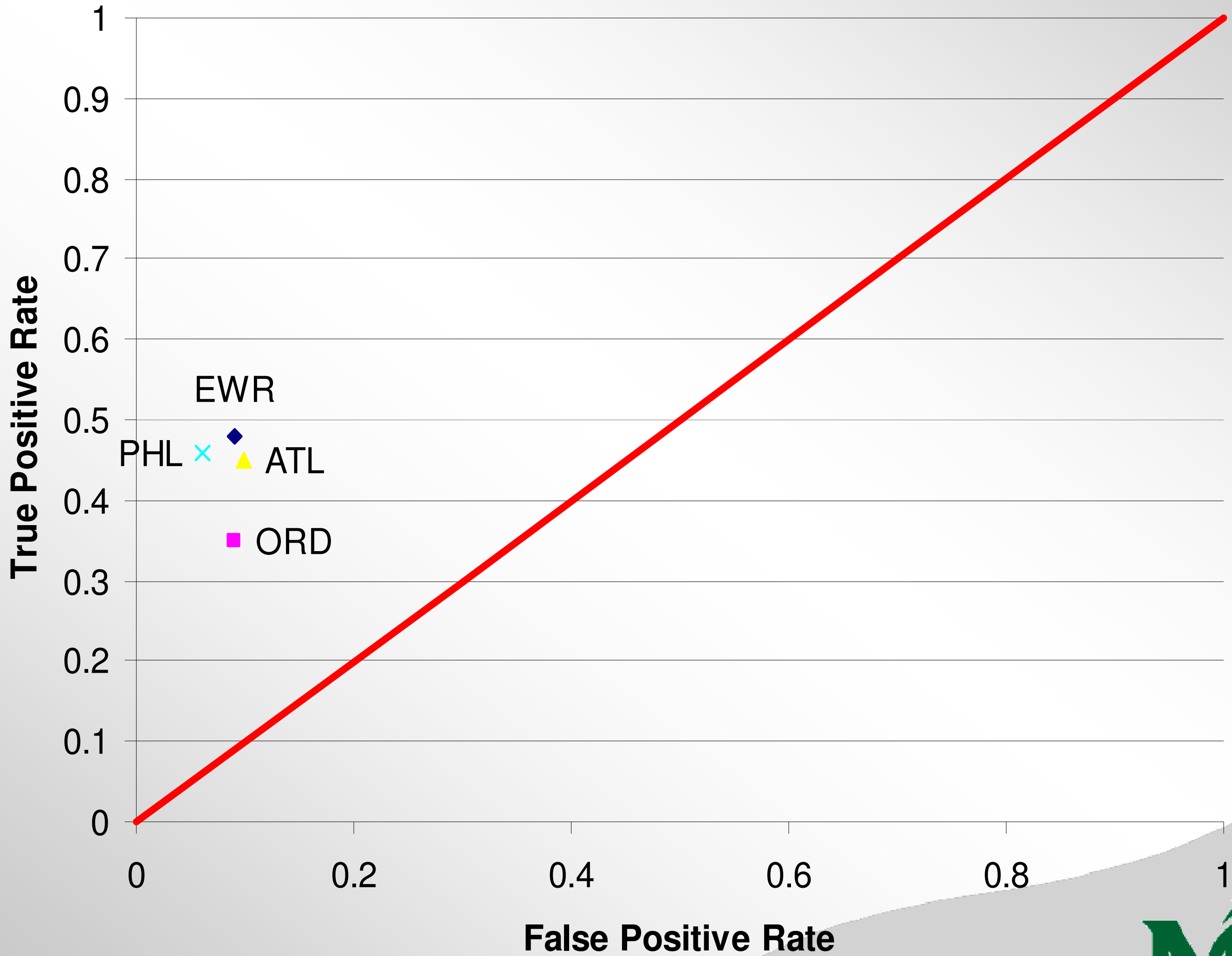
Airport	Sensitivity	Specificity	% Correct when predicted GDP	% Correct when predicted no-GDP	% Correct
EWR	0.48	0.91	0.74	0.77	0.76
ORD	0.35	0.91	0.67	0.72	0.72
ATL	0.45	0.90	0.63	0.81	0.78
PHL	0.46	0.94	0.69	0.86	0.84
				Average	0.78

- 1826 Days – JAN 2002 to DEC 2006
- Sensitivity – Proportion of all GDP airports that are identified by the model
- Specificity – Proportion of non-GDP airports that are identified by the model

Receiver Operating Characteristic (ROC) Curve

- True Positive Rate vs. False Positive Rate
- Evaluates how the model represents the data
- Red line is the line of no-discrimination
  - Indicates no better than random chance
  - Model is discarded if on line
- The farther the data point is in the upper left hand corner the better

GDP ROC Graph (Training)

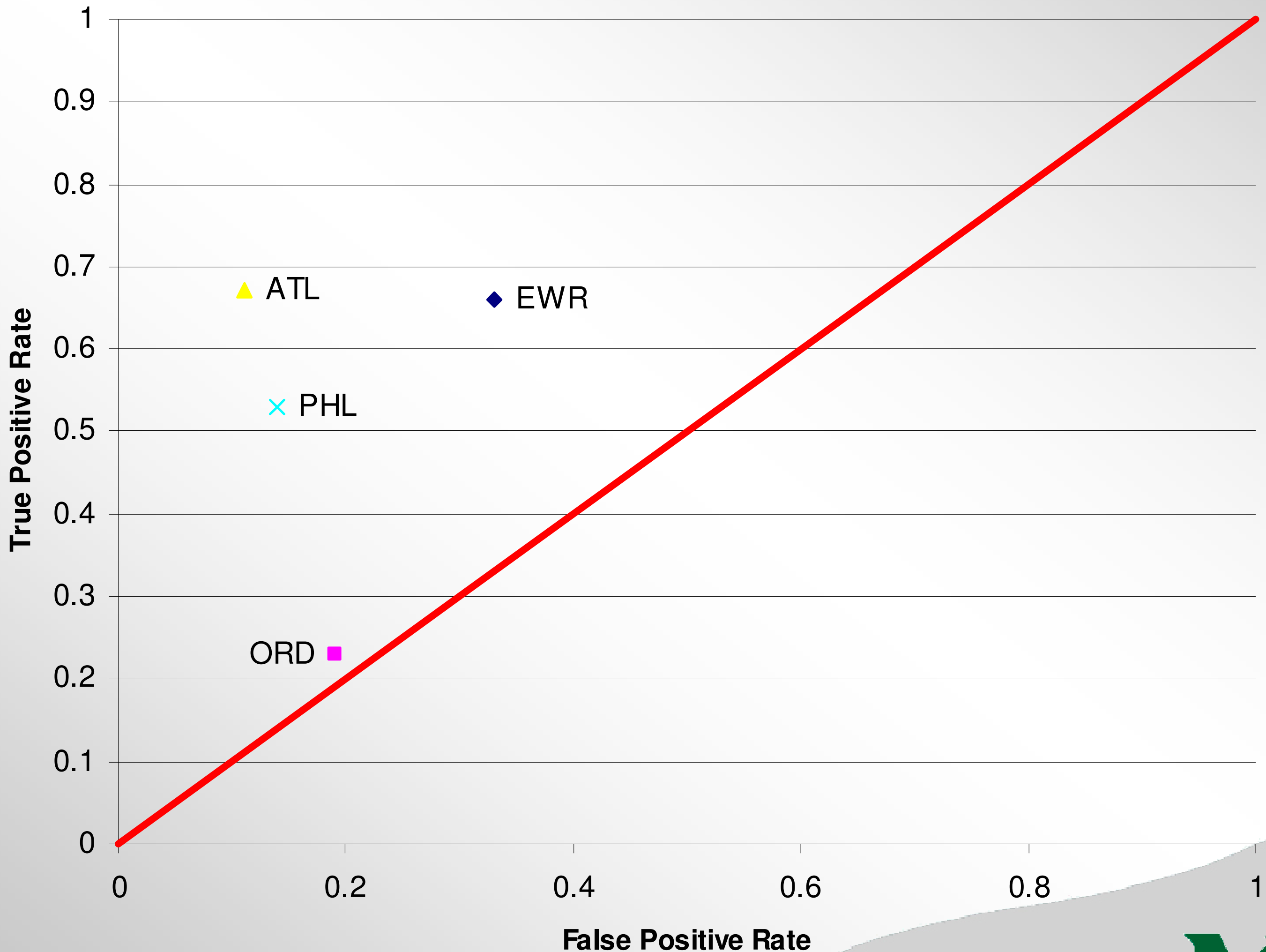


GDP Results (Test Data)

Airport	Sensitivity	Specificity	% Correct when predicted GDP	% Correct when predicted no-GDP	% Correct
EWR	0.66	0.67	0.78	0.52	0.66
ORD	0.23	0.81	0.42	0.63	0.58
ATL	0.67	0.89	0.25	0.98	0.88
PHL	0.53	0.86	0.57	0.84	0.78
				Average	0.73

JAN 2007 – APR 2007

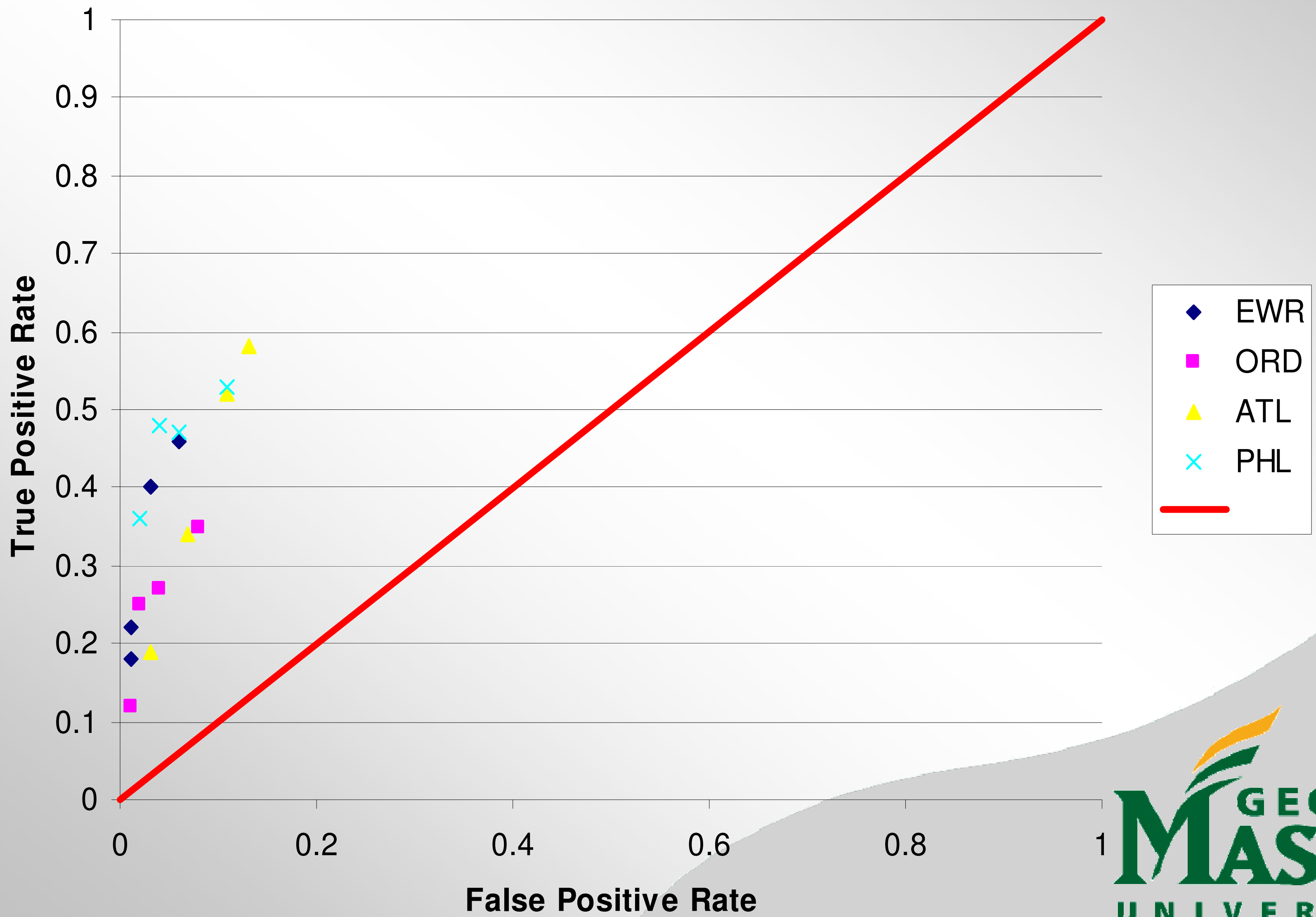
GDP ROC Graph (Test)



Delay Results (Training Data)

Airport	Average Delay	Sensitivity	Specificity	PPV	NPV	% Correct
EWR	15 min	0.46	0.94	0.78	0.79	0.79
EWR	30 min	0.40	0.97	0.75	0.89	0.88
EWR	45 min	0.22	0.99	0.78	0.94	0.93
EWR	60 min	0.18	0.99	0.97	0.98	0.97
ORD	7.5 min	0.35	0.92	0.69	0.72	0.72
ORD	15 min	0.27	0.96	0.68	0.80	0.79
ORD	22.5 min	0.25	0.98	0.74	0.85	0.84
ORD	30 min	0.12	0.99	0.76	0.88	0.88
ATL	5 min	0.58	0.87	0.78	0.72	0.74
ATL	10 min	0.52	0.89	0.69	0.81	0.78
ATL	15 min	0.34	0.93	0.62	0.83	0.81
ATL	20 min	0.19	0.97	0.58	0.97	0.88
PHL	5 min	0.53	0.89	0.77	0.74	0.74
PHL	10 min	0.47	0.94	0.75	0.83	0.81
PHL	15 min	0.48	0.96	0.74	0.88	0.86
PHL	20 min	0.36	0.98	0.73	0.90	0.88
					Average	0.83

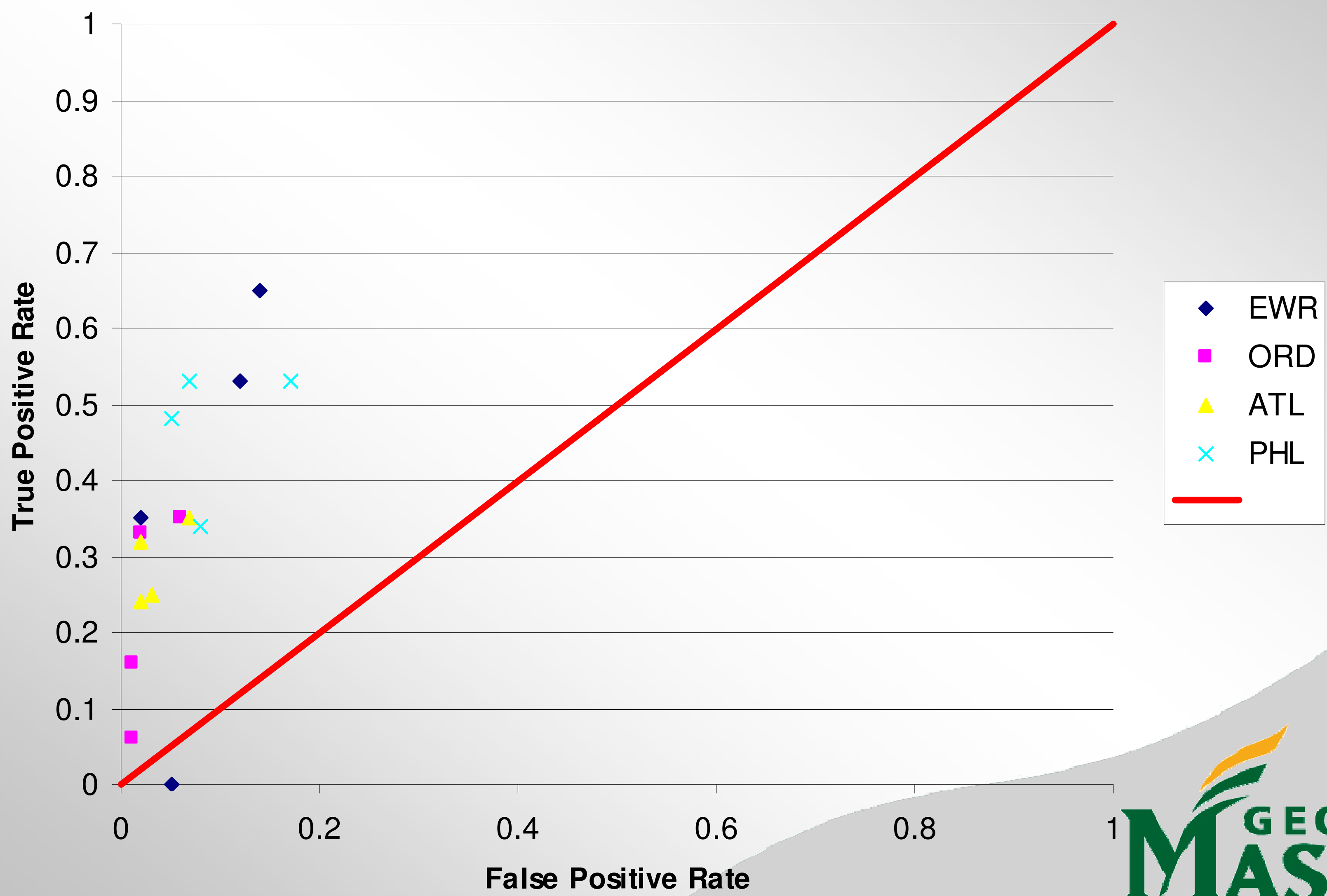
Delay ROC Graph (Training)



Delay Results (Testing Data)

Airport	Average Delay	Sensitivity	Specificity	PPV	NPV	% Correct
EWR	15 min	0.65	0.86	0.87	0.65	0.74
EWR	30 min	0.53	0.88	0.73	0.76	0.75
EWR	45 min	0.35	0.98	0.78	0.88	0.88
EWR	60 min	0.00	0.95	0.00	0.91	0.87
ORD	7.5 min	0.33	0.98	0.96	0.48	0.58
ORD	15 min	0.35	0.94	0.81	0.68	0.70
ORD	22.5 min	0.16	0.99	0.99	0.68	0.70
ORD	30 min	0.06	0.99	0.99	0.74	0.74
ATL	5 min	0.35	0.93	0.76	0.70	0.71
ATL	10 min	0.32	0.98	0.82	0.83	0.83
ATL	15 min	0.24	0.98	0.67	0.89	0.88
ATL	20 min	0.25	0.97	0.50	0.92	0.90
PHL	5 min	0.53	0.83	0.83	0.52	0.64
PHL	10 min	0.34	0.92	0.79	0.61	0.65
PHL	15 min	0.53	0.93	0.77	0.81	0.80
PHL	20 min	0.48	0.95	0.72	0.86	0.84
					Average	0.76

Delay ROC Graph (Test)



Conclusions

- The rarer the event the easier to predict the negative
- Cost prediction can be easily derived
- Larger data set **may or may not** produce a better model
- Like to expand to other major airports to give an overall NAS forecast
- Benefits
  - Airlines can plan schedule and route changes
  - FAA can simulate traffic management decisions
  - Traveler can be advised to potential delays
  - Data mine controller reaction
  - Forecasters can estimate the cost of a forecast

Future Work

- Expand data collection
- Predict delays throughout the entire NAS
- Predict Flow Constrained Areas (FCA) and Ground Stop
- Compare SVM to other methods
  - Trees
  - Regression

Questions?

Backups

Results - Definitions

- GDP Days - % of GDP days
- Non GDP Days - % of non-GDP days
- Sensitivity – Of all GDP airports how many are identified by the SVM
- Specificity – Proportion of non-GDP airports that are identified by the SVM
- Positive predictive value – Probability that if the SVM predicts a GDP that one actually occurs
- Negative predictive value – Probability that if the SVM predicts no GDP that one does not occur

References

1. Kay, M.P., J.L. Mahoney, J.E. Hart, "An Analysis of CCFP Forecast Performance for the 2005 Convective Season." University of Colorado. Boulder, Colorado, 2006.
2. Torbert, D. D. Rodenhuis, "Operational Feedback Reports to Providers of Aviation's Collaborative Convective Forecast Product." Federal Aviation Administration. Herdon, Virginia, 2005.
3. Seseske, S., J.E. Hart, "Collaborative Convective Forecast Product (CCFP). Cooperative Issuance Analysis." Cooperative Institute for Research in Environmental Sciences. University of Colorado. Bolder, Colorado, 2006.
4. Hill, T., Lewicki, P., *STATISTICS Methods and Applications*. StatSoft, Tulsa, OK, 2006.
5. Winston, W.L., *Operations Research: Applications and Algorithm*. Brooks/Cole-Thomson Learning, Belmont, CA, 2004.

CATSR

GEORGE MASON UNIVERSITY

Benefits

- Airlines can adjust operations
- TFM specialists can use forecast patterns to link to specific outcomes
- Standardize procedures
- Improve simulations
- Forecasters can assess the “cost” of a forecast

What is New in this Research

- TAF research has focused
  - Forecast verification
  - How to improve the process
- No one has investigated the effect of the forecast on TFM managers
- The question is not did it rain, but instead did you bring an umbrella?