Weather Normalization For Evaluating National Airspace System (NAS) Performance

Jing Xiong, Mark Hansen
University of California, Berkeley
July 3rd 2007
ATM 2007
Barcelona, Spain

Agenda

- Overview
- Model Description
- Model Estimation Results
- Conclusion
- Future Research

Performance Measurements

- Dimensions of National Airspace System (NAS) Performance
  - Safety
  - Capacity
  - Cost
  - Delay
- What are sources of delay?
  - High traffic volume
  - Airport capacity reduction
  - Convective weather
  - Airline internal malfunctions
  - Air traffic management initiatives

Statistical Delay Models

- Relate NAS performance (delay) to causal factors such as traffic, en route wx, terminal wx
- Based on daily or monthly data
- Motivations include
  – Understanding causes of delay
  – Tracking ANSP performance
- Active research area in US (Klein, Jehlen, Ball, Wieland, Sridhar, Post, Knorr, MITRE)

Contributions of this Paper

- Innovative performance metrics (DAFT)
  - Decomposable
  - Avoid schedule padding effects
- Consider wx forecast errors
- “En-route” WITI time series characterization
- Comparison of factors in terms of
  - Relative importance in explaining performance
  - Performance sensitivity

Model Specification

- *Perf(t)=f(Tra(t),WITI(t), Wind(t), IFR(t), Fcst(t))+v(t)*
  - *Where:*
  - *Perf(t)* is some NAS performance metric in day *t*;
  - *f(.)* is a deterministic function;
  - *Tra(t)* is air traffic demand in day t;
  - *WITI(t)* is a vector characterizing the en-route WITI in day *t*;
  - *Wind(t)* is average wind speed at major airports in day *t*;
  - *IFR(t)* is proportion of flights scheduled to land under IFR conditions in day *t*;
  - *Fcst(t)* is a vector capturing the weather forecast errors in day *t*;
  - *v(t)* is stochastic error term;

Variables Description

- Performance metrics
  - ASPM 75 daily average delay
  - Deviation of Average Flight Time Index (DAFT)
- Air traffic demand
- En route convective weather (WITI)
- Terminal weather (Wind and IFR)
- Weather forecast performance metrics

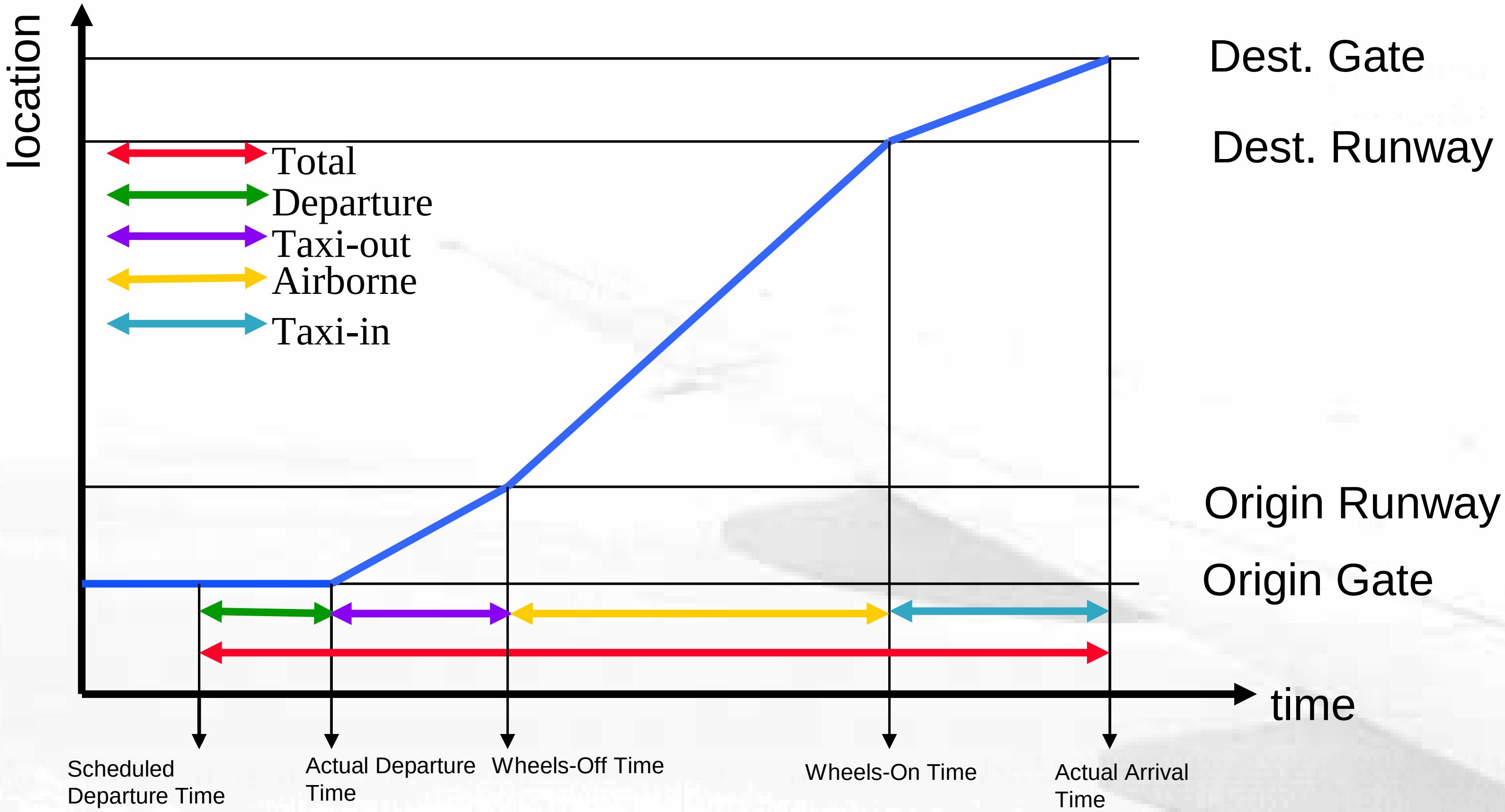
ASPM Daily Average Delay

- Total arrival delay against schedule divided by total completed arrivals
- Negative delay (arrive early) counted as zero
- 75 benchmark airports

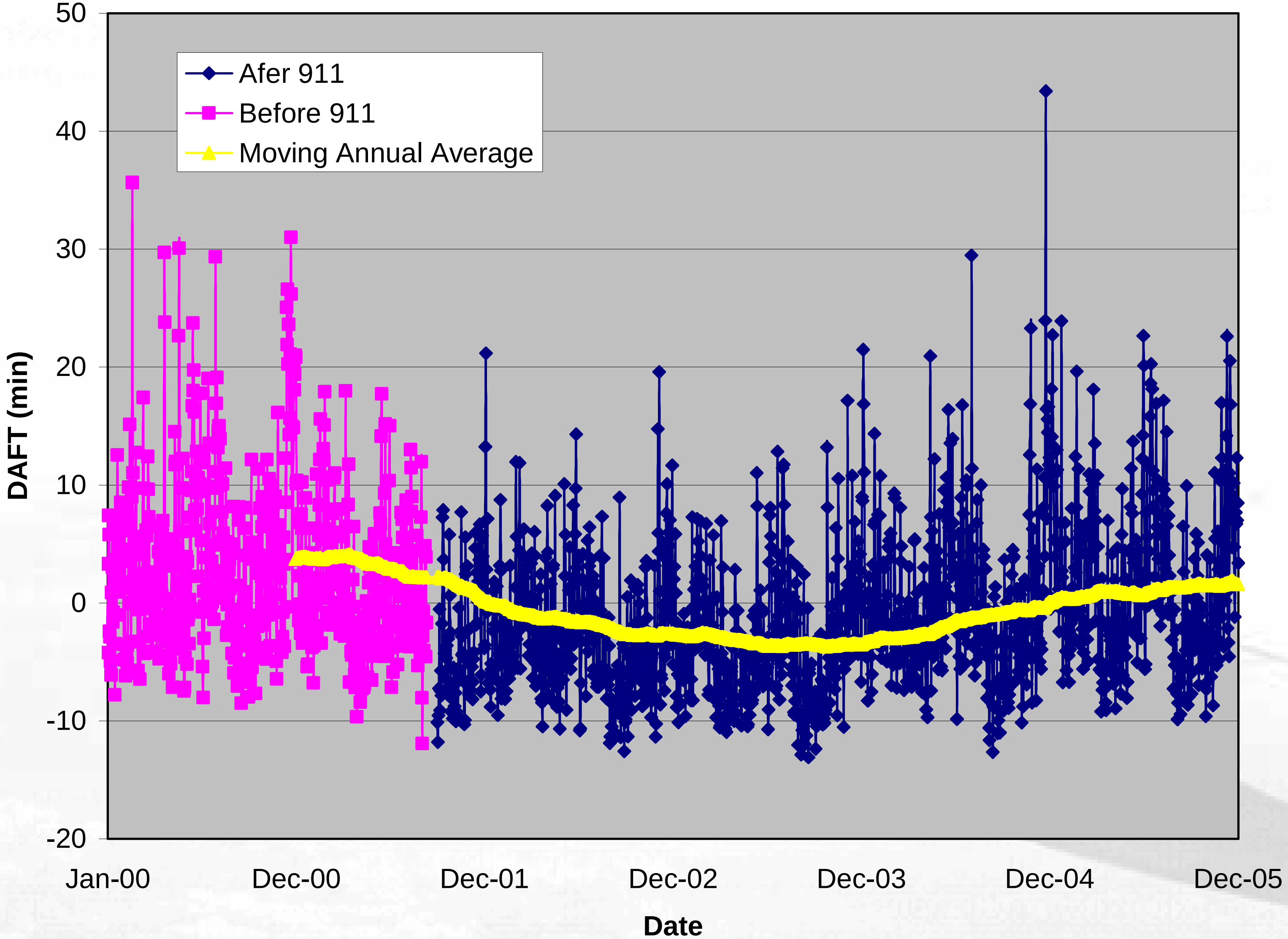
DAFT (Deviation of Average Flight Time)

- "Consumer price index" of flight times
- Market basket of OD pairs with fixed weights based on flight volume
- 0 values corresponds to average over 2000-2006 period
- Contains different phase of flight: gate delay, taxi-out time, airborne time, taxi-in time

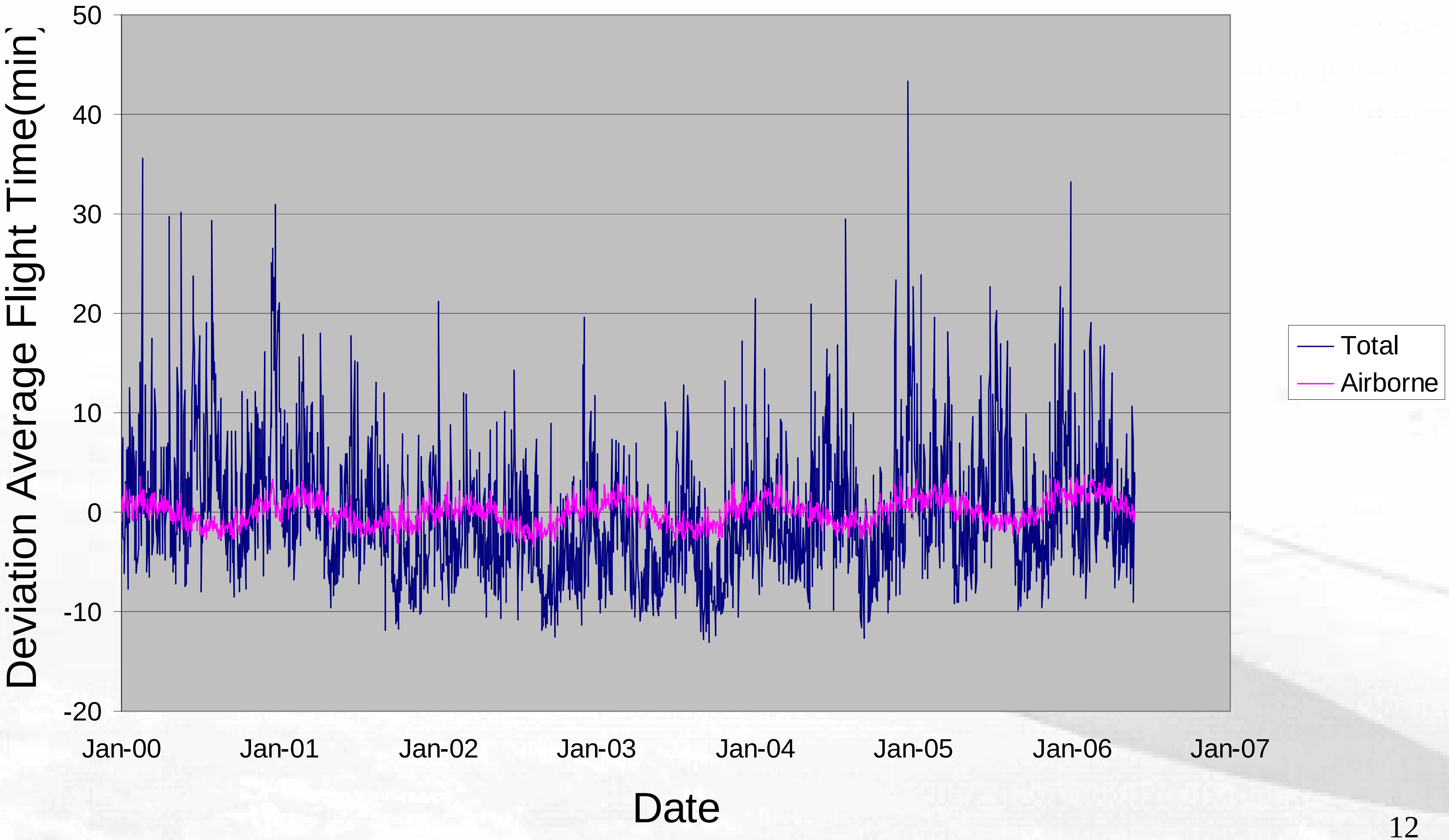
DAFT and its Components



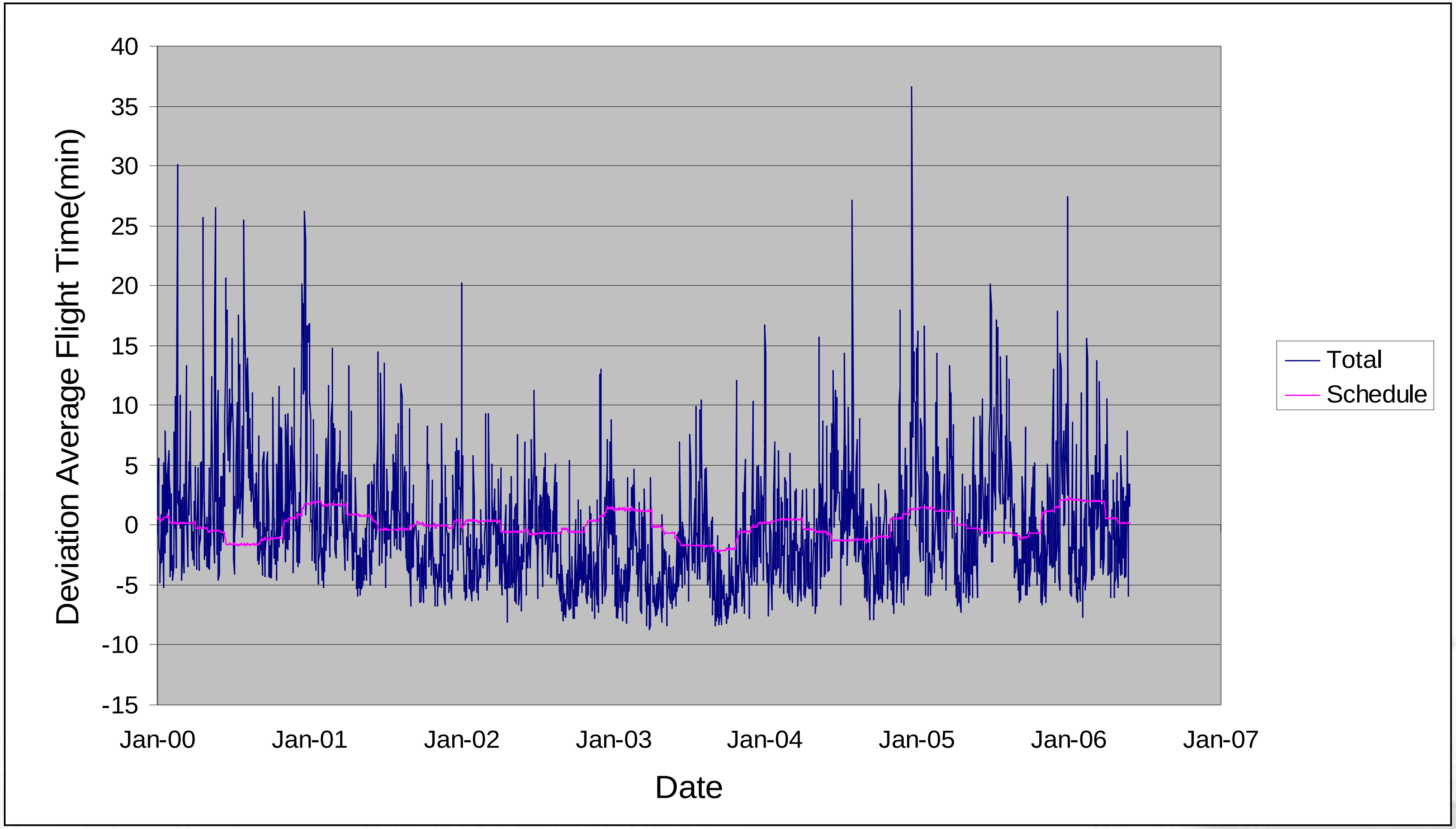
DAFT Trends 2000-2005



DAFT Total and Airborne



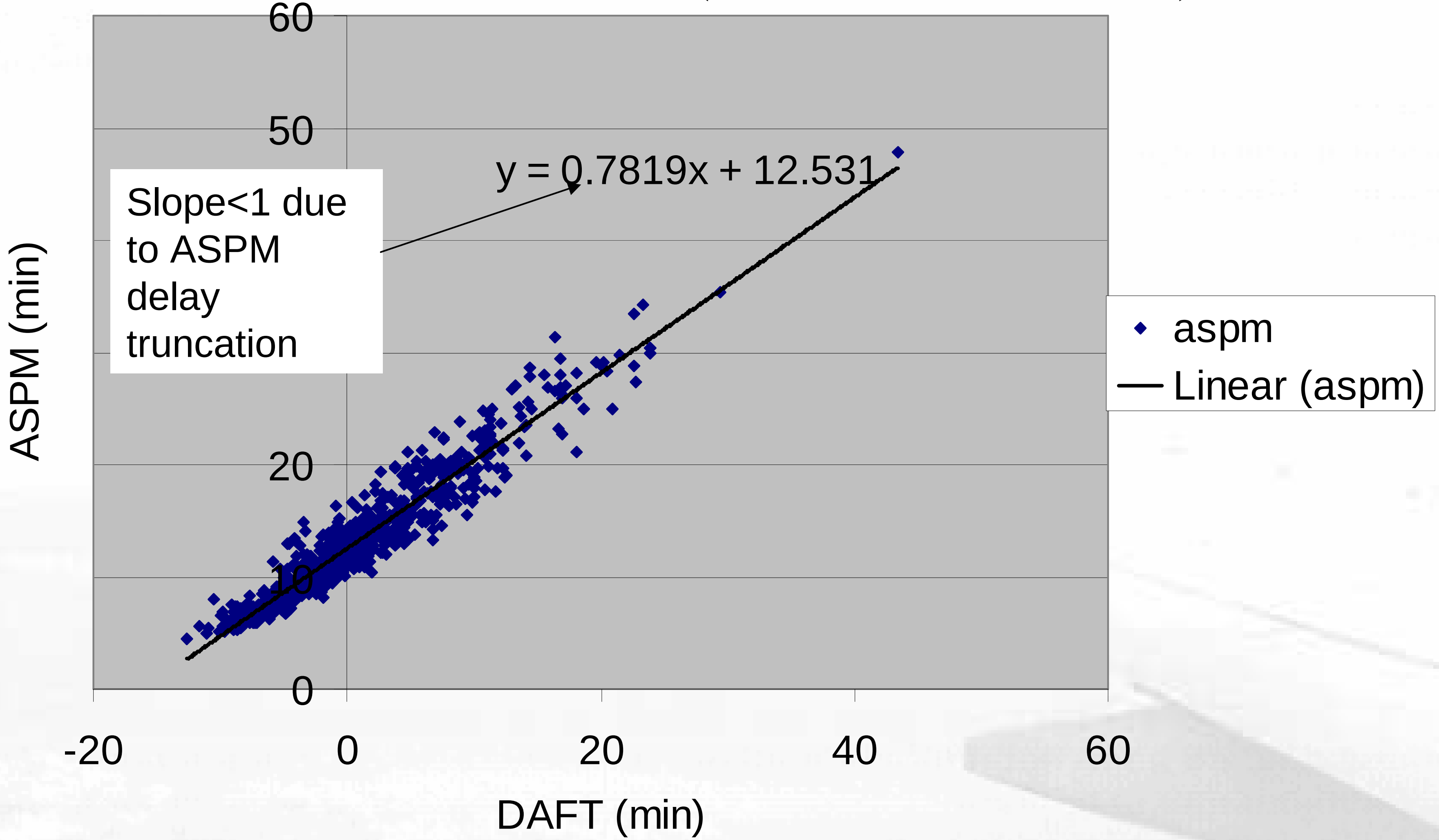
DAFT Total and OAG Schedule



ASPM Daily Average vs DAFT

- Comparing with schedule vs Comparing with “Average” over the analysis period
- Padding effect vs no padding
- Truncation of negative delay vs no truncation
- Total delay vs decomposed to four phases of the flight

DAFT vs ASPM (2004 and 2005)



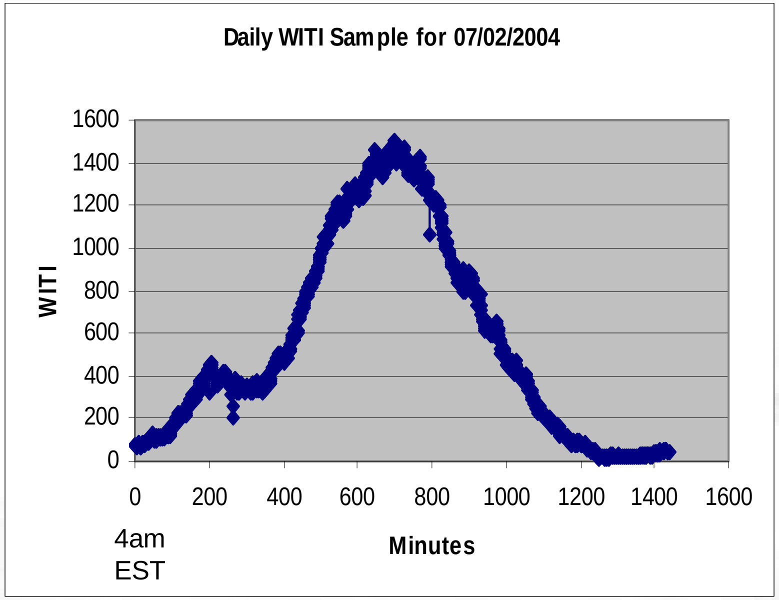
WITI Development (Sridar et al)

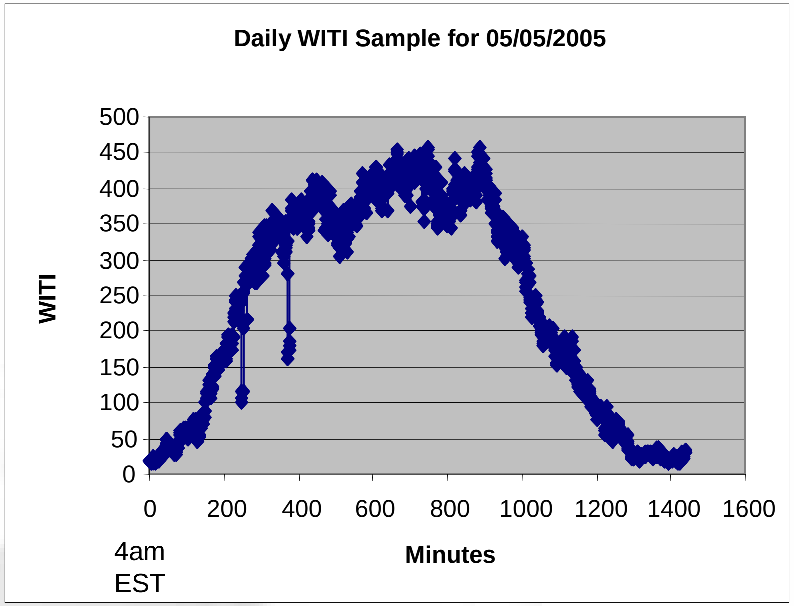
- WITI (t) = $\sum_{i,j} W_{i,j}(t) \cdot T_{i,j}(t)$
- $W_{i,j}(t)$ is
  - Severe convective weather incidence in cell *(i,j)* at time *t*
  - Binary data developed from NOWRAD
  - Five-minute interval
  - Extended to 20 miles
- $T_{i,j}(t)$ is
  - traffic counts in cell *(i,j)* at time *t*
  - Reference day ETMS actual trajectories
  - One-minute interval
- Reference day
  - a day with low OPSNET delay but high traffic

Klein's WITI vs Sridar's WITI (En-route)

- Great circle trajectories vs ETMS trajectories
- Actual flight schedule on the flows on a particular day vs Reference day for all sample days

WITI Daily Time Series





- 60x24=1440 observations
- Jul, Sep and Oct in 2004; Apr, May Jun, Jul and Aug in 2005; 189 sample days in total

WITI Characterization

- Used small number of variables to characterize daily time series
  - Mean value scaled by actual scheduled traffic
  - Distribution around mean
  - Time of day when wx/trajectory interactions occur
  - Based on distribution moments
- Significant but little improvement in model performance compared to using daily mean WITI only

Weather Forecast Performance Metrics

- CCFP verification statistics
  - Compare realized weather with forecasted
  - 40 x 40 km cells

	Observation	
Forecast	No	Yes
No	**NN**	**NY**
Yes	**YN**	**YY**

  - 2, 4, 6 hours forecast results are all included
- Two sets of variables to measure weather forecast accuracy
  - False-positive (YN) & False negative (NY)
  - Over-forecasting Ratio & Under-forecasting Ratio

More Variables…

- Air Traffic Demand
  - Total daily scheduled arrivals at ASPM 75 airports
  - Obtained from ASPM
- Wind
  - For each flight, find wind speed at destination airport when it is scheduled to land
  - Average over all flights
- IFR
  - MC is binary data, 1 when airport is operated under IFR condition, 0 otherwise
  - Fraction of flights scheduled to land in IFR conditions

Estimation Procedure and Results

- Simple linear model (first order approximation)
- Ordinary least squares
- Developed parsimonious model
  - Estimated full model using ASPM delay as performance variable
  - Eliminated insignificant variables to obtain parsimonious model
  - Estimated parsimonious model for
    - ASPM Delay
    - DAFT (Total and Components)

Full Model

Category	Variable	ASPM Models: ASPM delay 1 Estimate	ASPM delay 1 Std. Err.	ASPM delay 2 Estimate	ASPM delay 2 Std. Err.
	Intercept	***-76.079***	***32.061***	***-73.164***	***31.733***
	Scheduled Arr.	***40.918***	***20.108***	***48.924***	***21.775***
WITI	WITI mean	***0.104***	***0.011***	***0.084***	***0.019***
	WITI Std. Err.	***29.283***	***12.571***	***27.921***	***12.377***
	WITI Skewness	***-5.310***	***2.588***	***-5.367***	***2.559***
	Time mean	**0.038**	**0.021**	0.032	0.021
	Time Std. Err.	**52.712**	**28.112**	**50.360**	**27.640**
	Time Skewness	-0.995	2.763	-0.825	2.715
Terminal	Wind	**0.528**	**0.277**	***0.639***	***0.272***
	IFR	***15.109***	***4.234***	***16.431***	***4.236***
	False-positive			***10.068***	***2.870***
	False-negative			-6.841	4.542
	Over-forecasting Ratio	***12.349***	***3.609***		
	Under-forecasting Ratio	-1.121	1.557		
R-squared		0.624		0.636	

Values in ***bold italics*** are significant at 5% level
Values in **bold** are significant at 10% level

- Except for mean, WITI variables only marginally significant
- Forecasting error counts perform better than forecasting ratios
- False-positive highly significant but false negative is not

Parsimonious Models: ASPM and Total DAFT

Category	Variable	ASPM delay3 Estimate	ASPM delay3 Std. Err.	Total Estimate	Total Std. Err.
	Intercept	*-16.009*	*4.519*	*-45.847*	*5.268*
	Scheduled Arr.	*56.879*	*17.308*	*127.067*	*20.176*
WITI	WITI mean	*0.065*	*0.013*	*0.065*	*0.015*
	WITI Std. Err.				
	WITI Skewness				
	Time mean				
	Time Std. Err.				
	Time Skewness				
Terminal	Wind	*0.834*	*0.252*	*0.798*	*0.294*
	IFR	*18.279*	*3.878*	*21.062*	*4.520*
	False-positive	*6.387*	*1.544*	*8.870*	*1.800*
	False-negative				
	Over-forecasting Ratio				
	Under-forecasting Ratio				
R-squared		0.617		0.655	

- Most coefficients have similar magnitudes
- Coefficient on traffic much higher in DAFT model (masking due to truncation?)
- DAFT model has slightly higher R^2

Unit-free Measures: Beta Weight and Elasticity

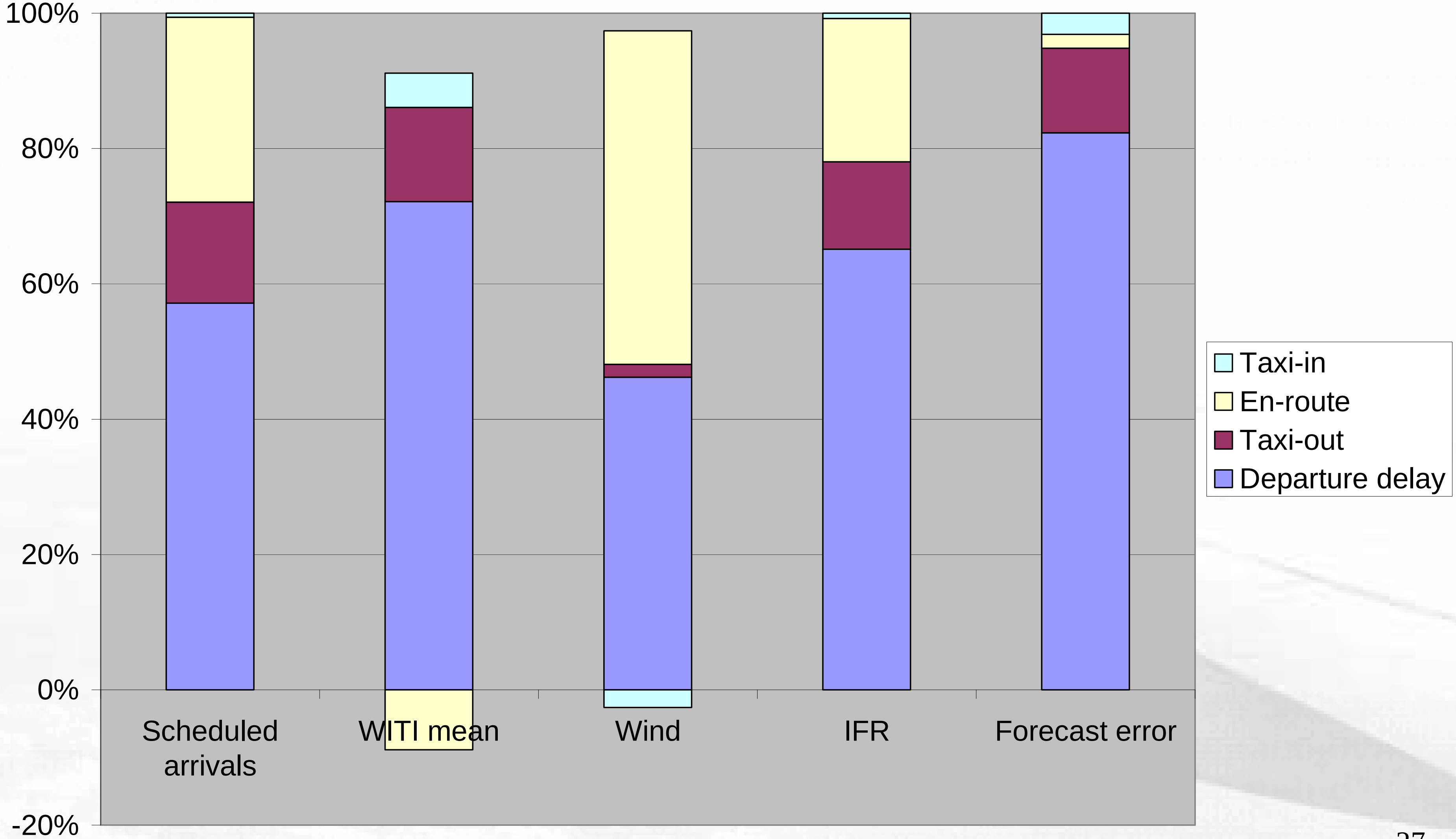
- Beta weight: measures contribution of variation in explanatory variable to variation in dependent variable
- Elasticity: measures sensitivity of dependent variable to explanatory variable
- WITI and False-positive have highest betas
- Traffic has highest elasticity

		Scheduled flights	WITI Mean	Wind	IFR	False-positive
ASPM Model	Beta Weight	0.16	0.41	0.16	0.24	0.32
	Elasticity	1.01	0.35	0.50	0.16	0.19
DAFT Total Model	Beta Weight	0.30	0.34	0.12	0.23	0.36
	Elasticity	2.13	0.33	0.45	0.18	0.25

DAFT Component Results

- DAFT is sum of four components
- DAFT regression coefficients are sums of regression coefficients for the four components
- Contributions for different components
  - Provide insight into causal mechanism
  - Have implications for cost of delay

DAFT Component Results



Conclusions

- **WITI Characterization** -- Details of daily WITI time series much less important than mean value (but statistically significant)
- **DAFT Decomposition**
  - Components respond differently to various causal factors
  - ***Departure delay and taxi-out***: traffic, IFR, WITI, forecast errors
  - ***En route time***: traffic, wind, IFR

Conclusions (continued)

- **Convective Forecast Errors**
  - False positive storm forecasts increase delay. Without this contribution, the total delay would be ~20% less
  - Under-forecasting does not appear to affect the delay in our model. However this might suggest that failure to foresee convective weather affects cost impacts (and perhaps safety) of delay rather than its quantity

Future Research

- Incorporate higher order terms into delay model
- WITI refinements
- Improve forecast error characterization
- Assess realized performance against theoretical optimum
- Value function development