Models for Estimating Monthly Delays and Cancellations in the NAS

Avijit Mukherjee
Michael Ball
Bargava Subramanian

University of Maryland

Introduction

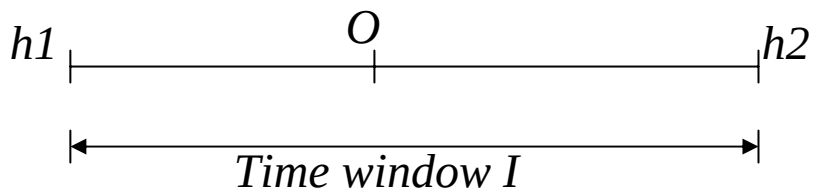
- Objectives
  - Develop a metric that indicates the level of congestion faced by a each operation (arrivals/departures) at an airport.
  - Estimate various percentiles of the distribution of the congestion metric across all operations in the NAS.
  - Model NAS average flight delay and cancellation probability as a function of the congestion level (i.e. the percentiles of the distribution of the congestion metric).

- Application in NAS Strategy Simulator
  - Estimating flight delays and cancellation under various demand growth or capacity increase scenarios in the NAS
  - Estimating passenger delay

Measure of Congestion

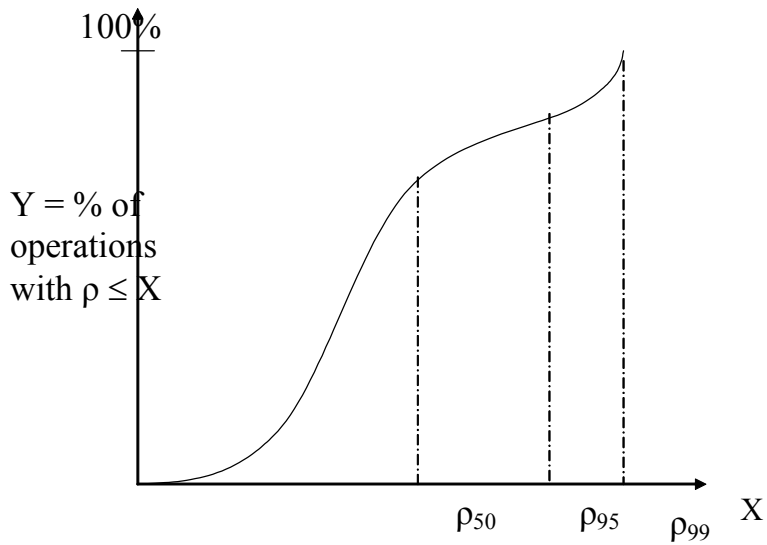
The demand/capacity ratio is a good measure of congestion. Each operation at an airport faces a congestion level depending on the capacity and demand at the airport in a time window at which the operation takes place.

Concept of Rho



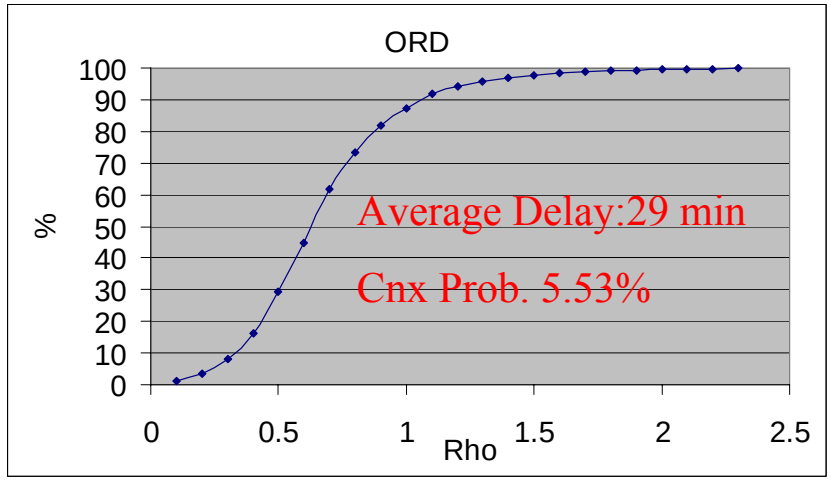
$$\rho = \frac{\#\text{ operations during } I \text{ at O's airport}}{\text{airport capacity during } I \text{ at O's airport}}$$

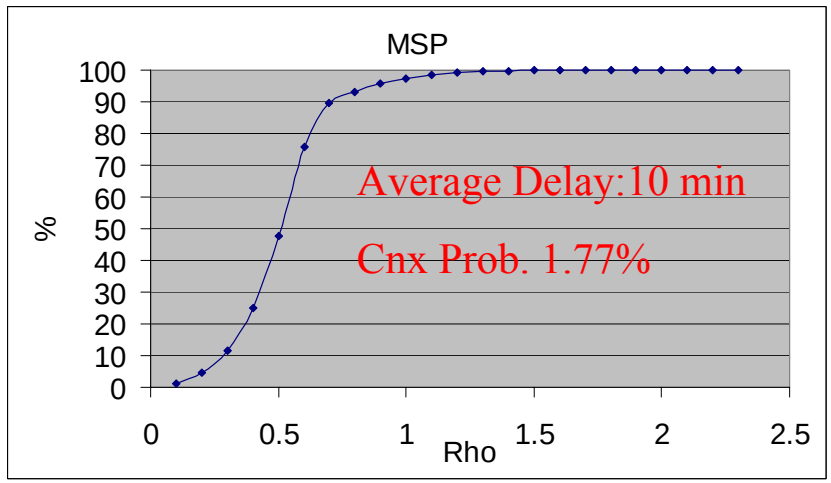
Percentiles of Rho



Estimating Rho at an Airport

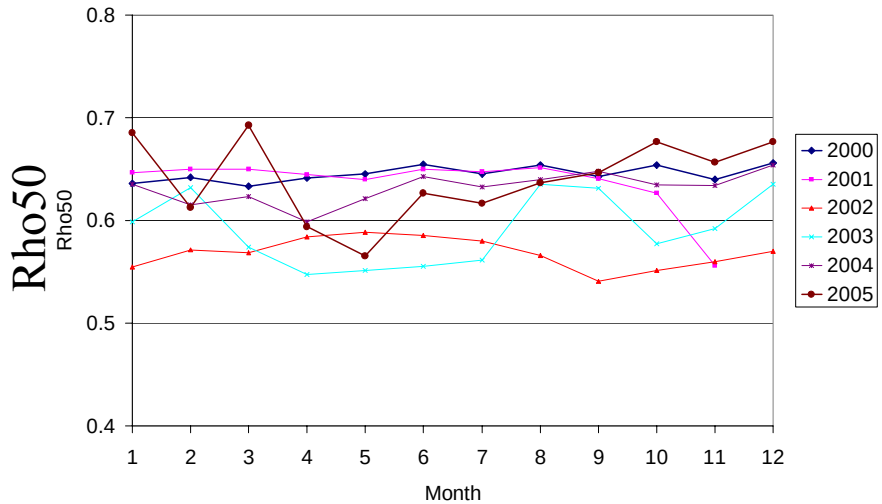
- Input
  - Hourly scheduled demand
  - GA demand
  - VMC / IMC capacities
- Associate demand/capacity in an hour as the Rho value for each operation during that hour
- Estimate the monthly distribution (or percentiles) of Rho at an airport

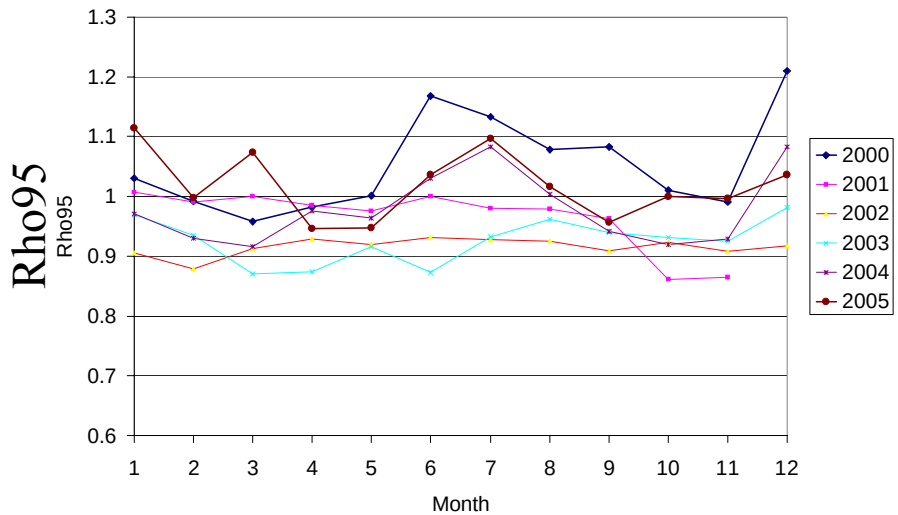




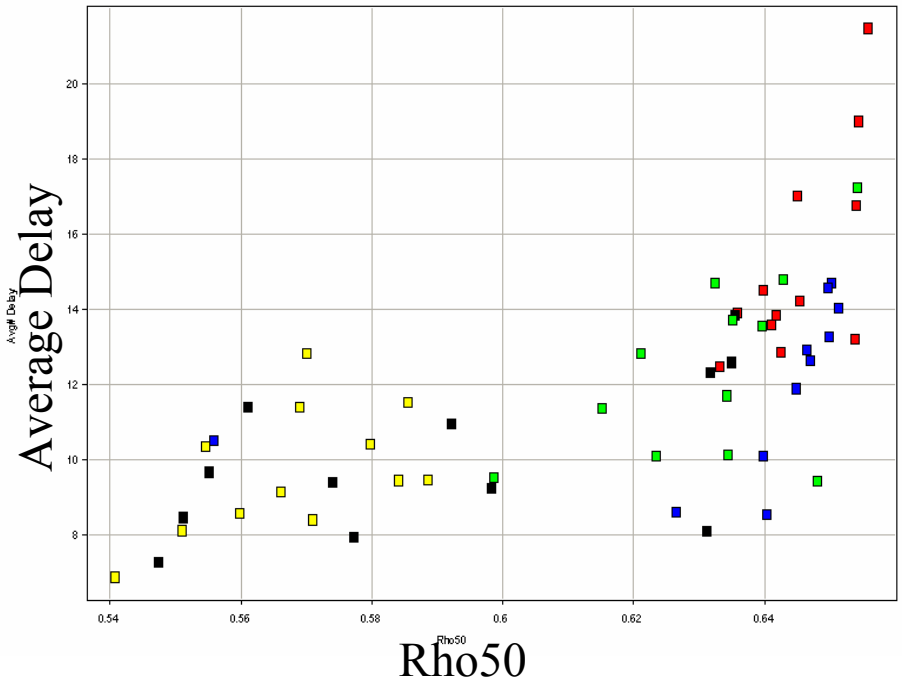
NAS Rho Percentiles

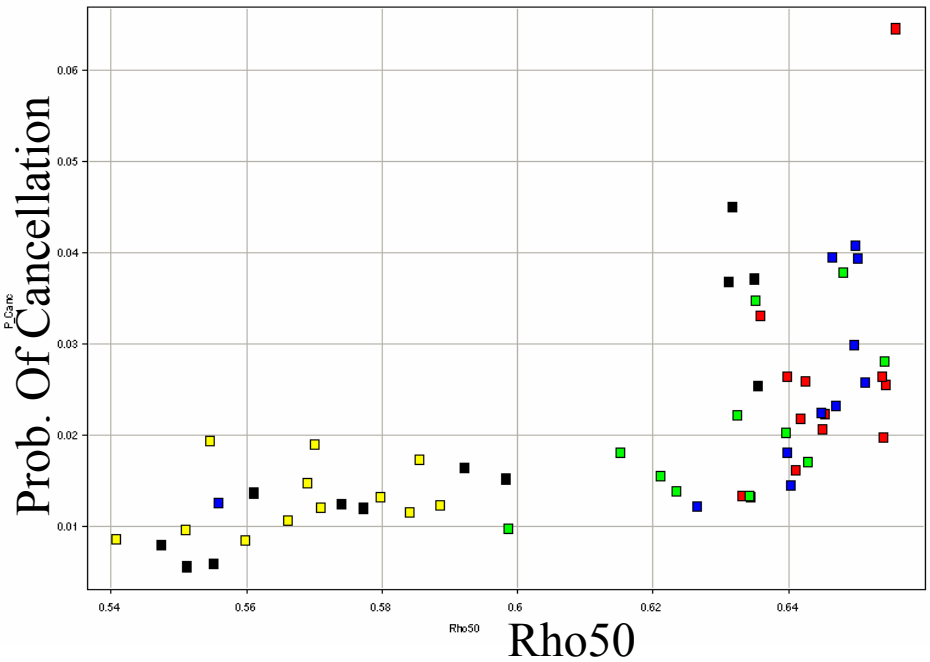
- Estimate Rho50 (50th percentile) and Rho95 (95th percentile) for all airports
- NAS Rho50 and Rho95 are weighted average of corresponding airport Rho's
- Weights are proportional to the fraction of NAS operations at each airport



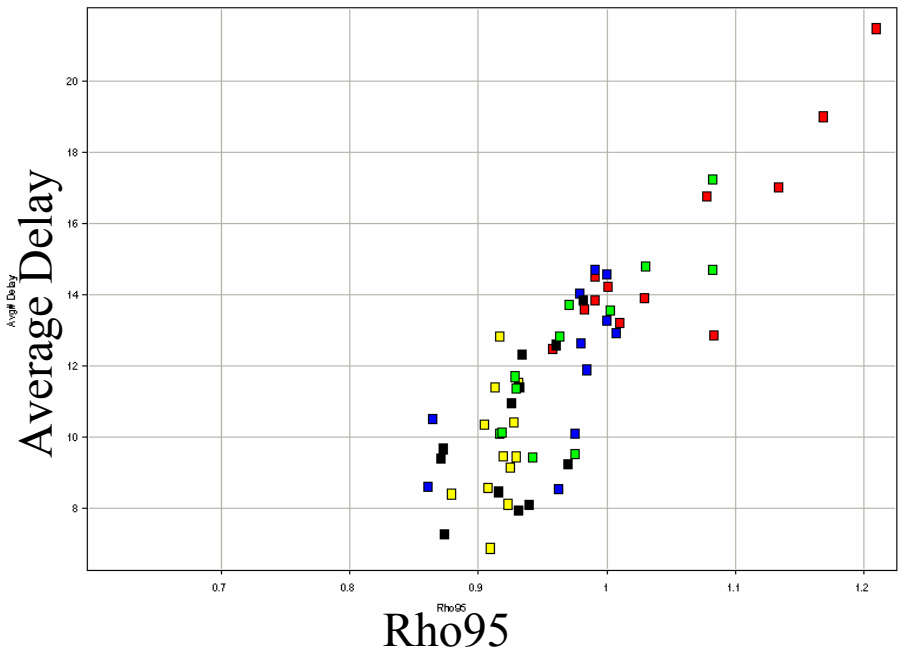


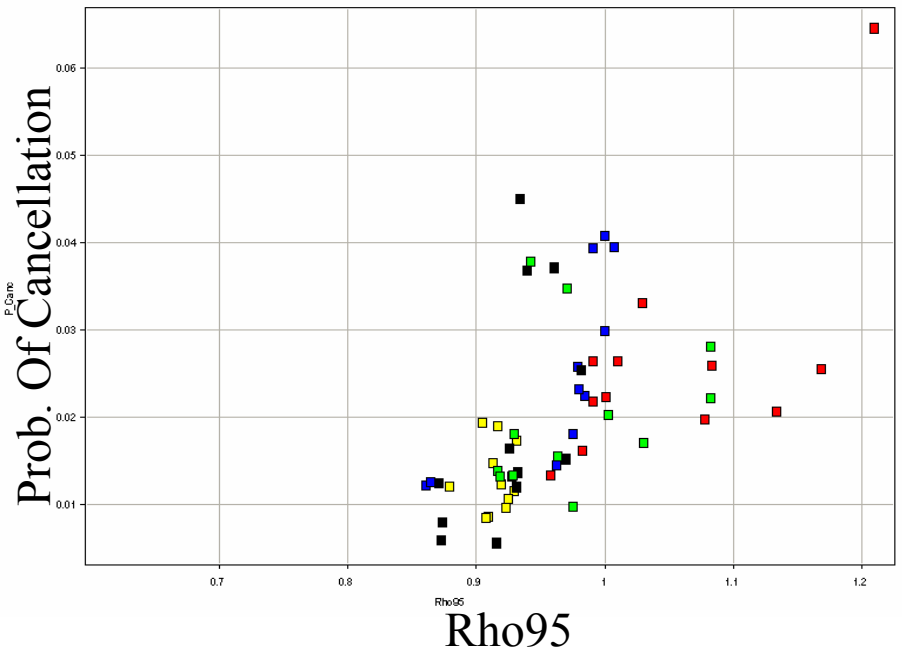
Impact of Rho50 on Delays and Cancellation





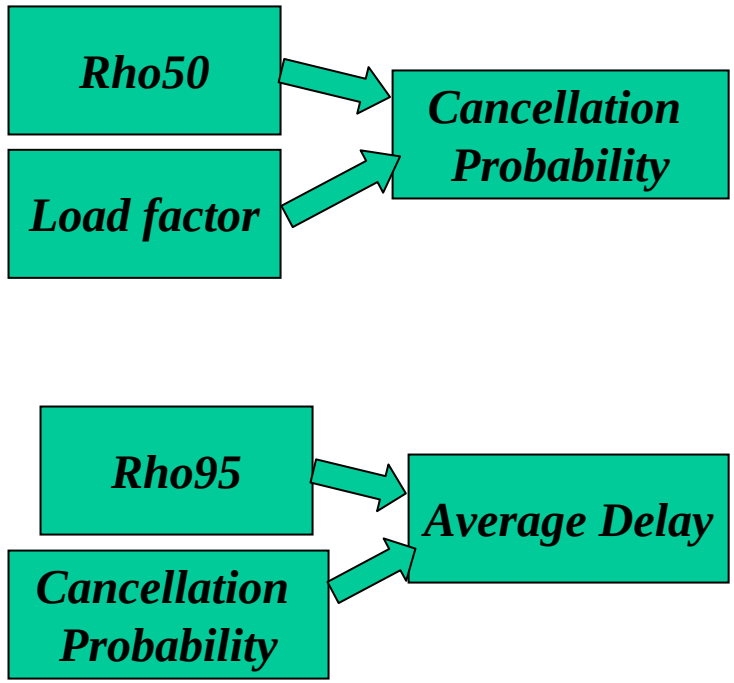
Impact of Rho95 on Delays and Cancellation



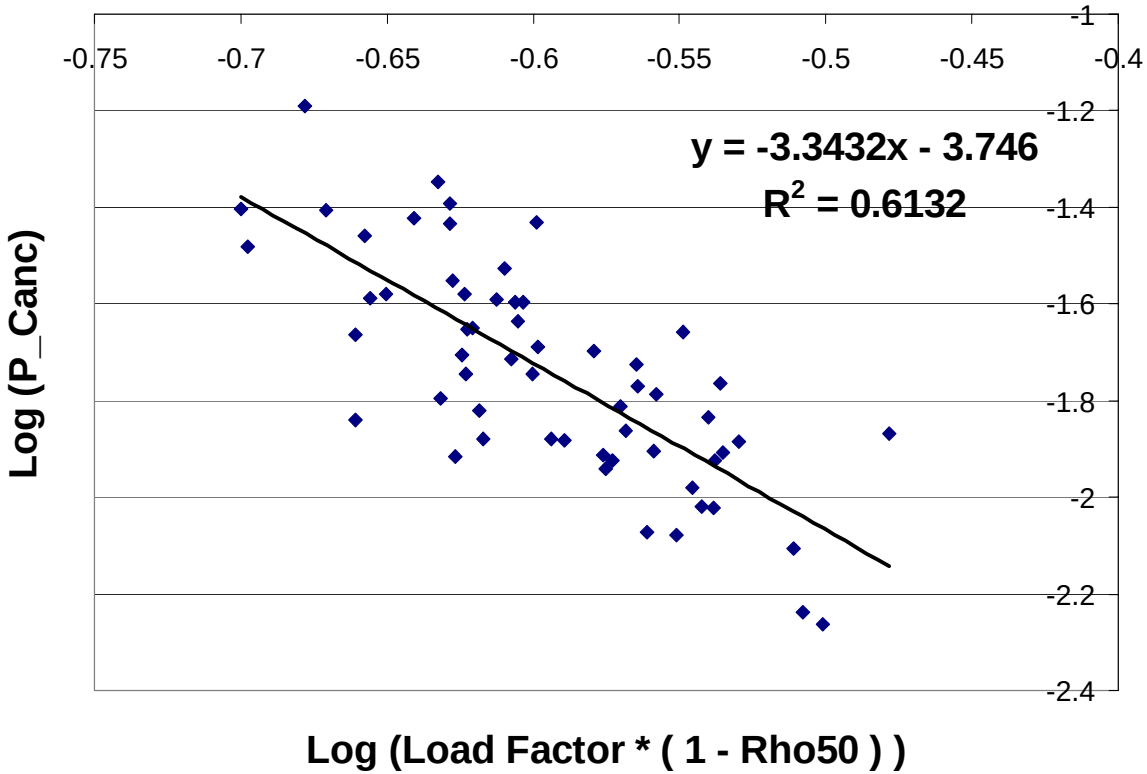


Estimating Flight Delays and Cancellation Probability

- Hourly scheduled demand and capacity at 35 major airports obtained from ASPM database
- Monthly load factor data obtained from Database Products Inc.
- Average flight delay and proportion of flights cancelled obtained from ASPM

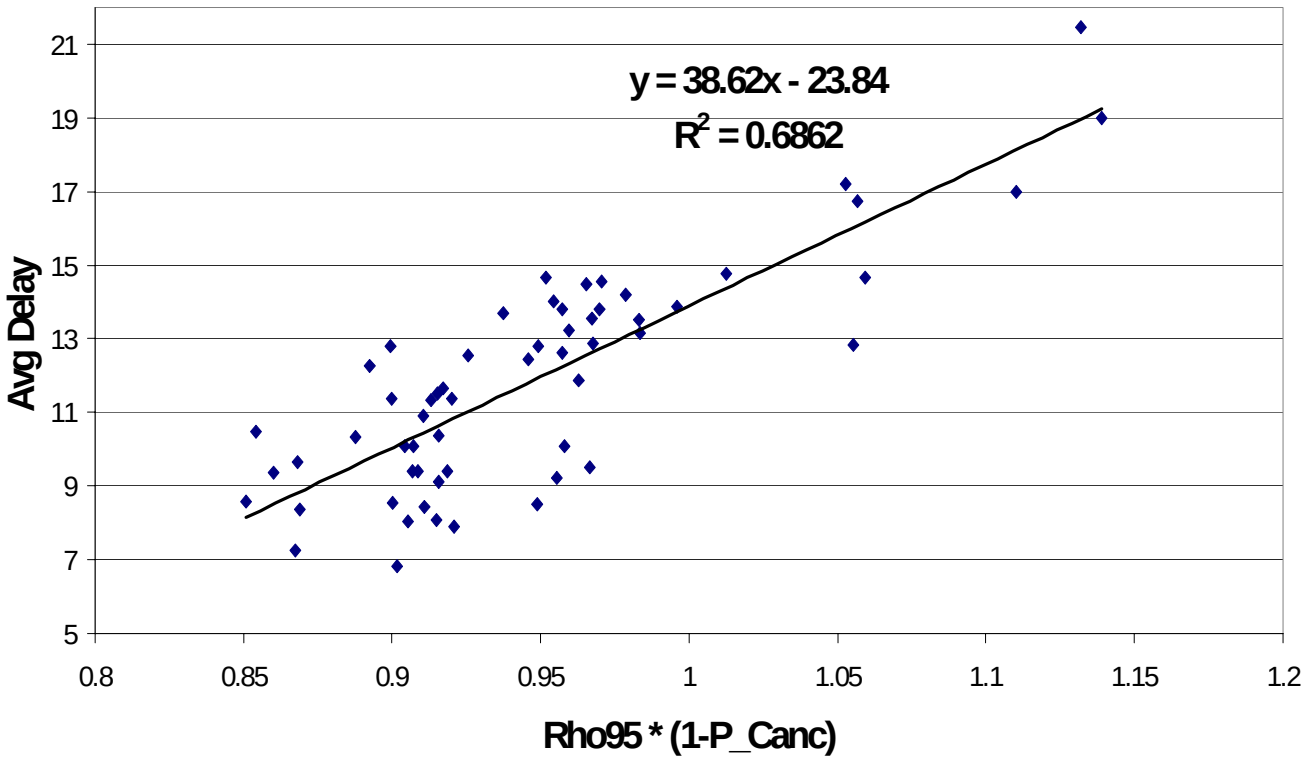


Monthly Cancellation Model



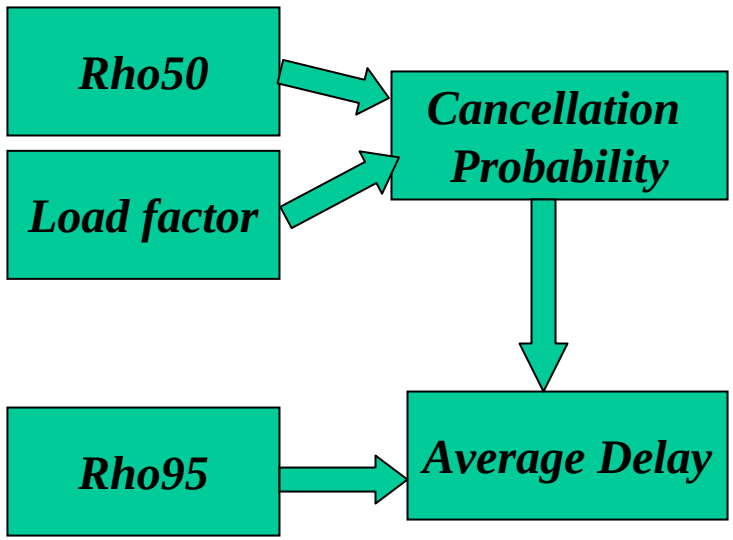
September 2001 considered outlier, hence excluded

Monthly Delay Model



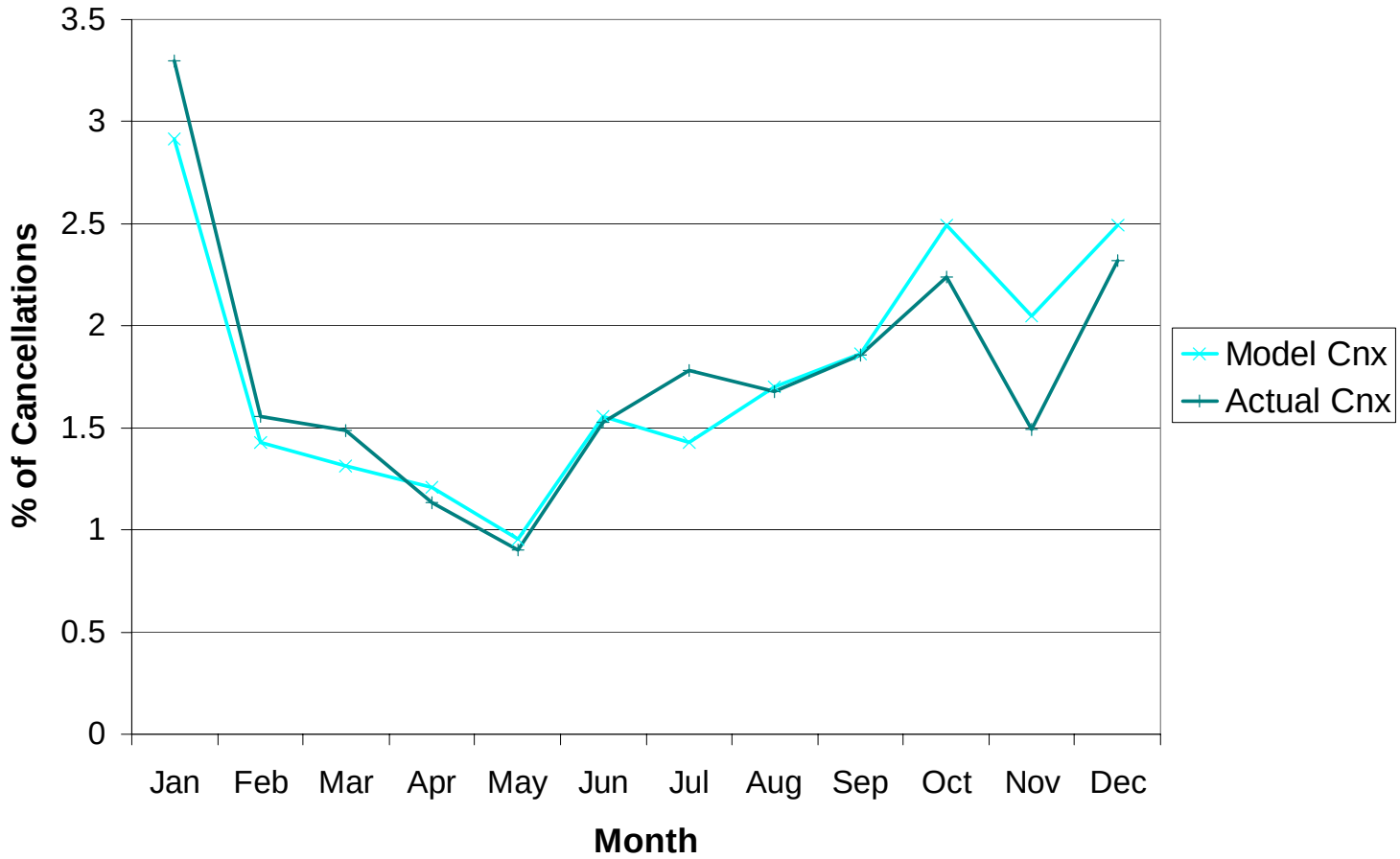
Functional Relationships

Canc. Probability = $e^{-3.75} * [\text{Loadfactor} * (1 - \text{Rho50})]^{-3.34}$

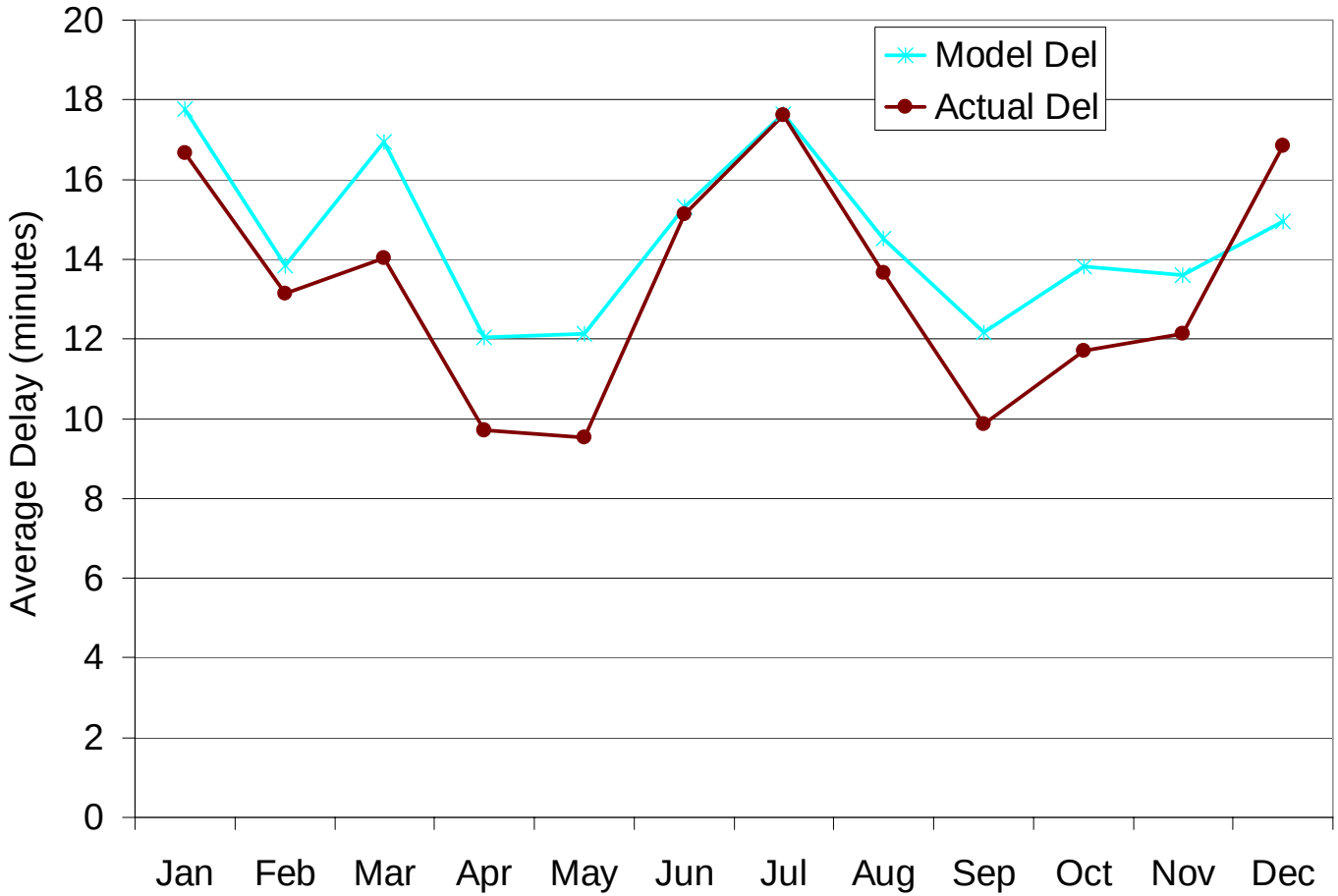


Average Delay = 38.62 * [Rho95 (1 – Canc. Prob.)] – 23.84

Predicted vs. Observed Prob. Of Cancellation in 2005

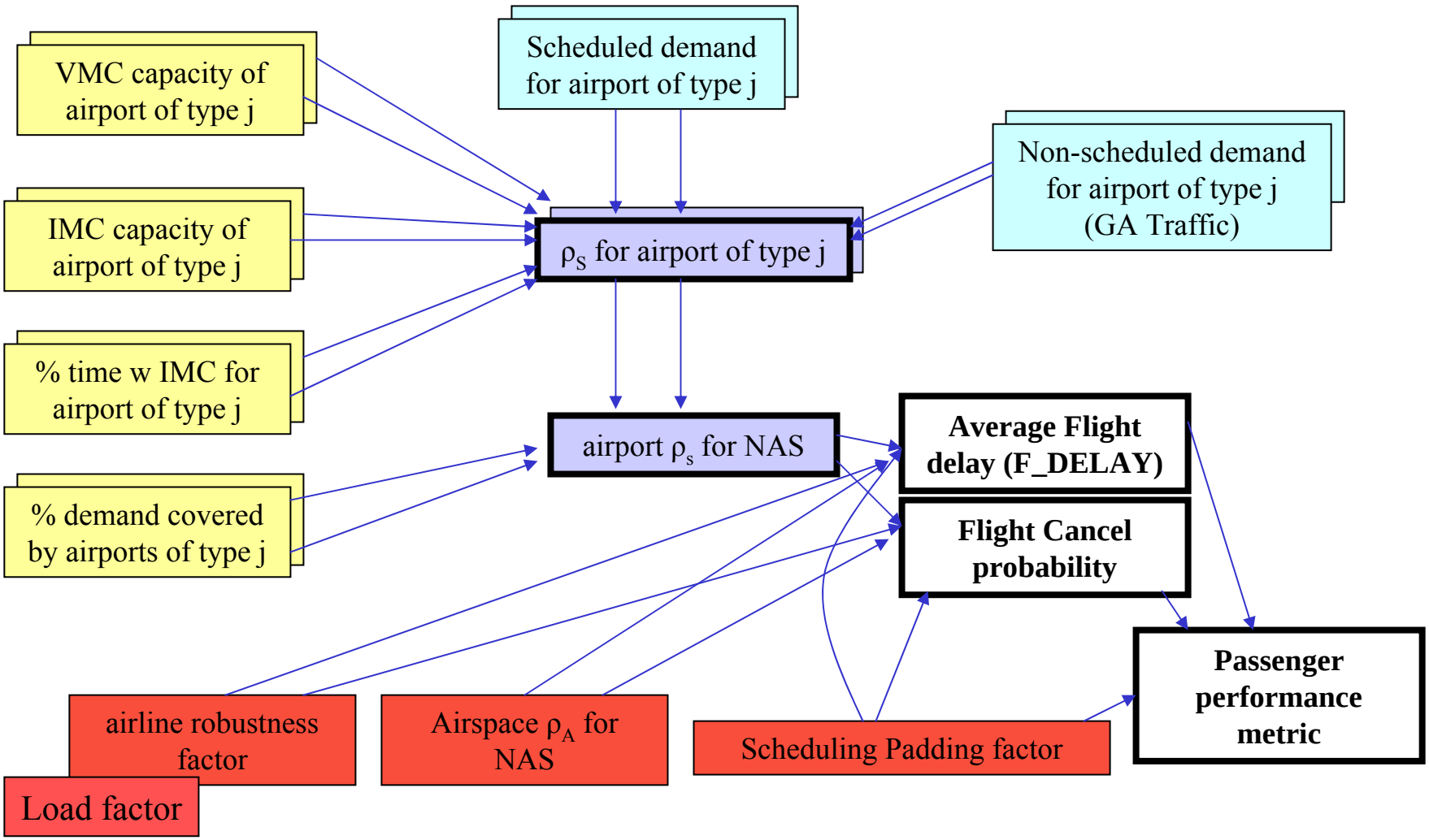


Predicted vs. Observed Delay in 2005



Applications

NAS Performance Models



Airport Categorization Based on Rho

Objective : To cluster airports using monthly Rho 50, Rho 95, Average Delays, Probability of Cancellation.

Method used : Clustering is carried out using K-Means clustering procedure. This is an iterative procedure wherein airports are assigned to various clusters so as to minimize the distance from the clusters' centroids.

Result : The airports can be divided into four categories :

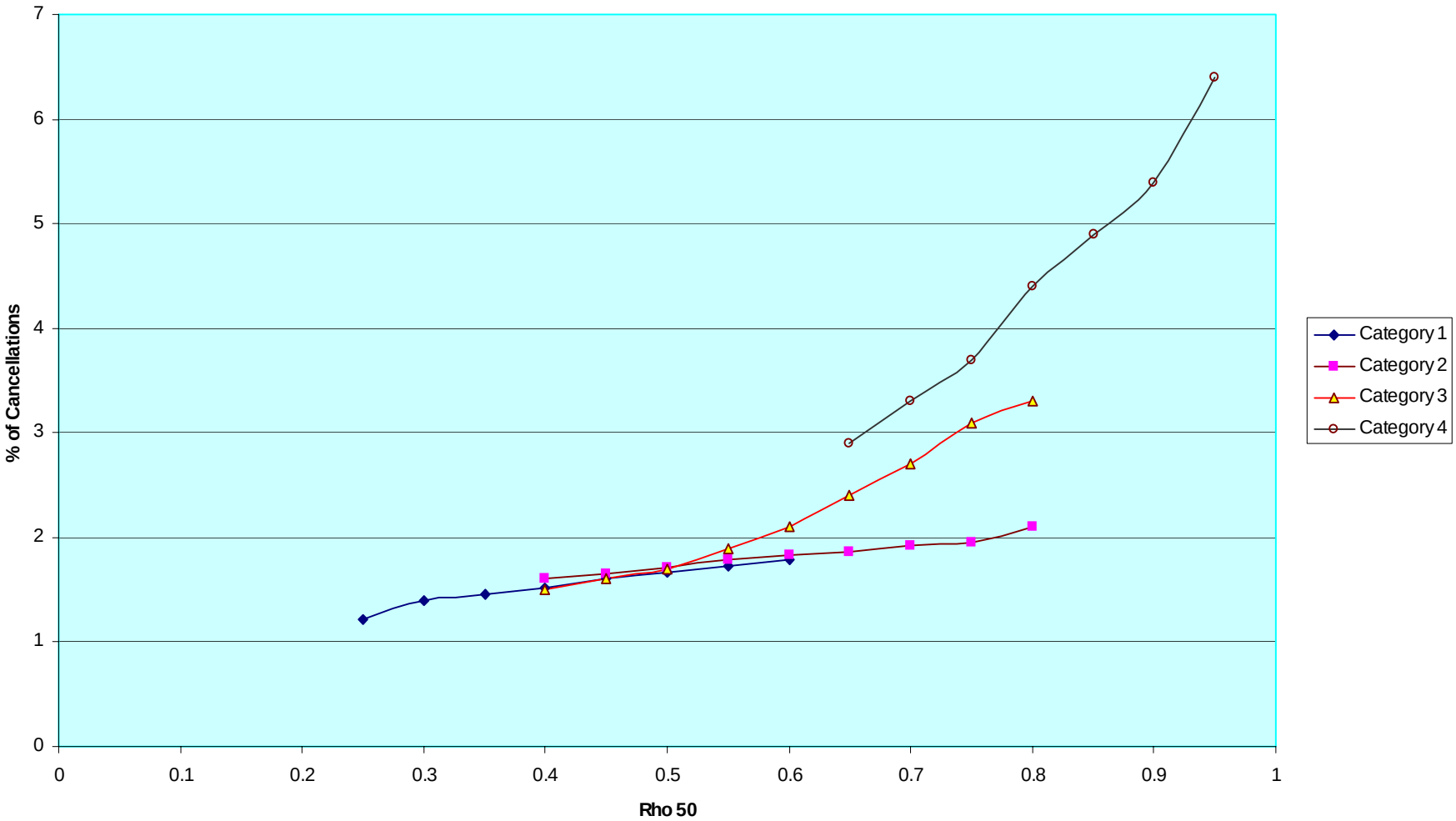
Category 1 : Low Rho50,Low Delay, Low Cancellation probability

Category 2 : Rho50, Delay, Cancellation probability similar to NAS average

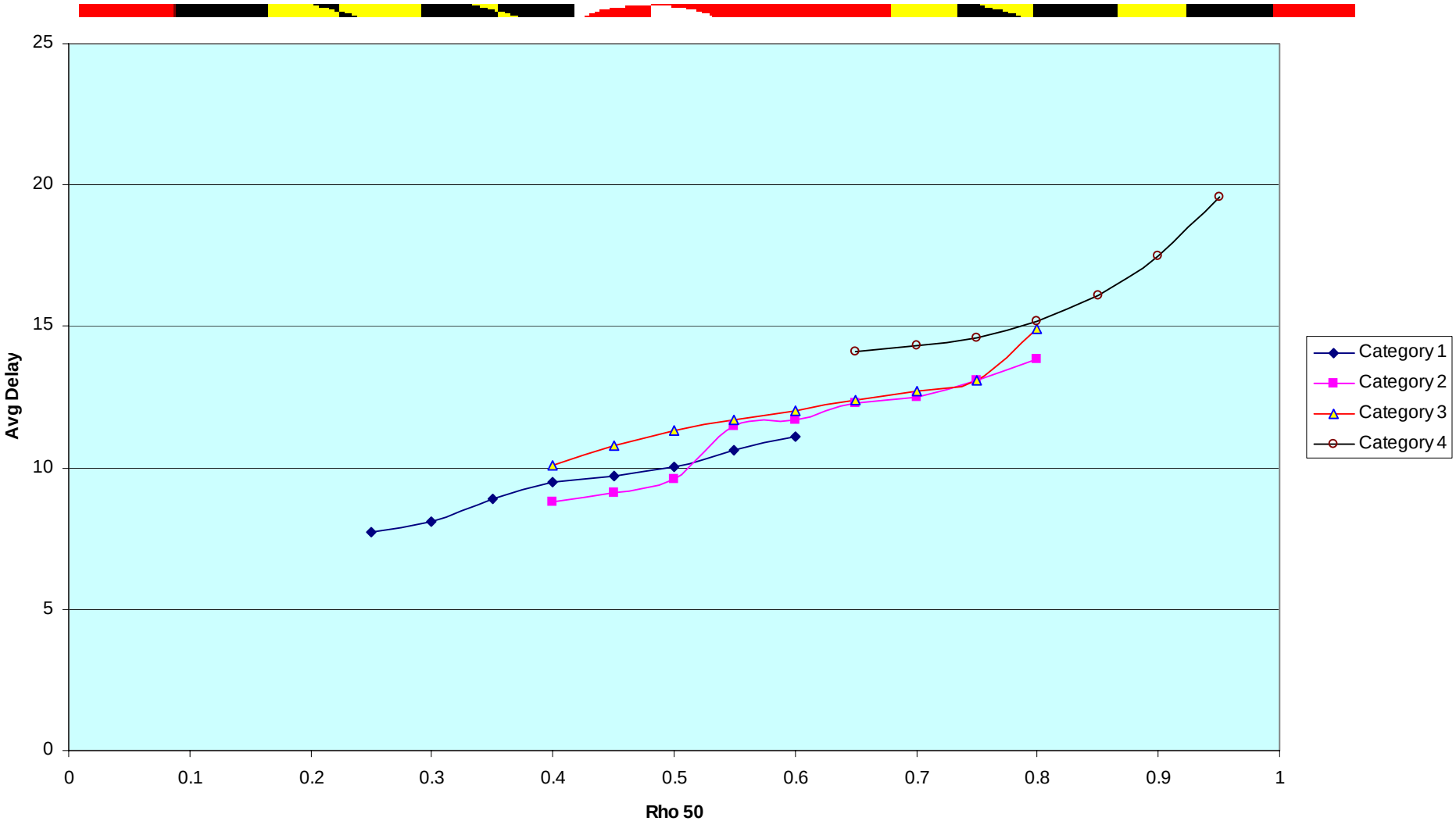
Category 3 : "Constrained" airport-set – delay similar to NAS average but cancellations increase beyond a particular Rho50.

Category 4 : High Rho, High Delay, High Cancellation Probability

Airport Categories



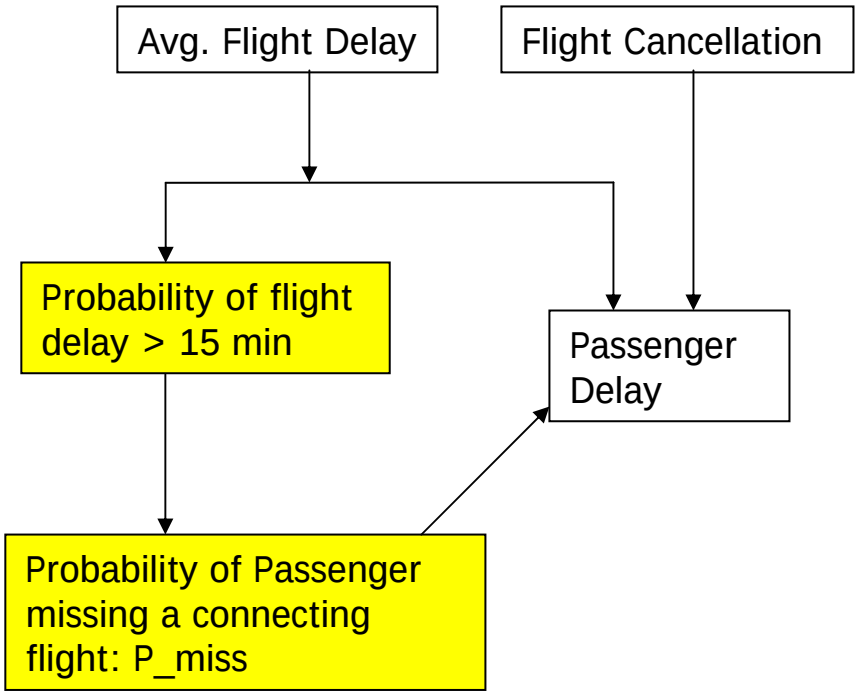
Airport Categories continued…



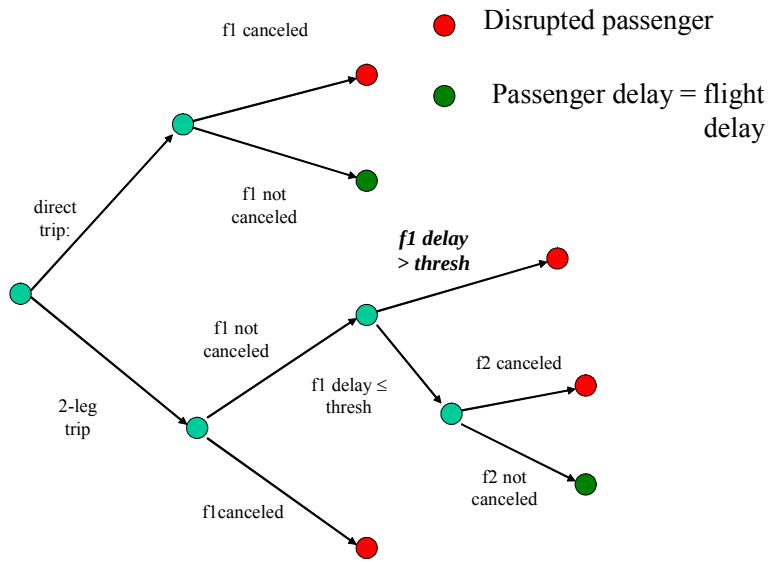
Airport Categories continued

Category 1	Category 2	Category 3	Category 4
TPA	PIT	DEN	PHL
MCO	MIA	IAD	BOS
SLC	CLT	DFW	EWR
BWI	FLL	DTW	ATL
PDX	LAS	DCA	ORD
	MDW	SFO	LGA
	CVG	JFK	
	SAN	LAX	
	PHX		
	SEA		
	IAH		
	MSP		
	STL		

Application: Estimating Passenger Delay

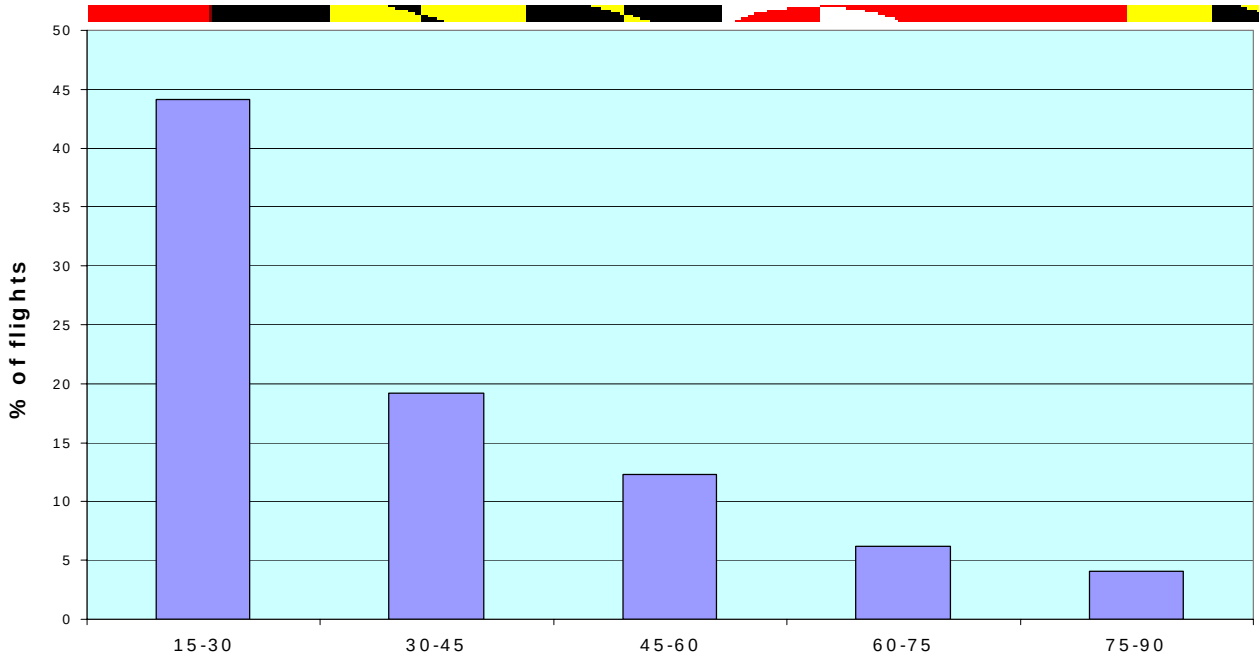


Decision Tree



$$P\_miss = Prob\{delay > layover\ time \mid delay > 15\ minutes\}$$

Flight Delay Distribution of Delayed Flights



From Individual flight data, construct histograms of delayed flights in each time interval for each month. Need to determine distribution of flight delays

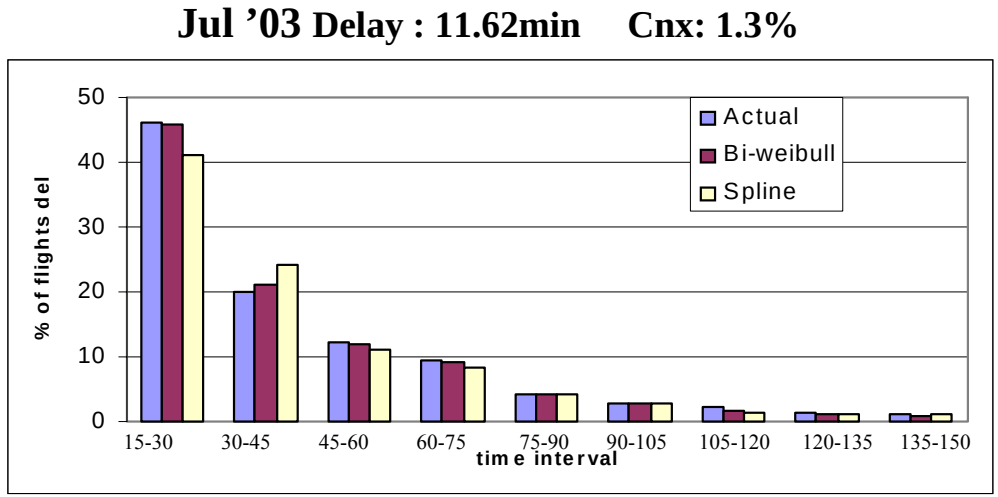
Two approaches to achieve the objective

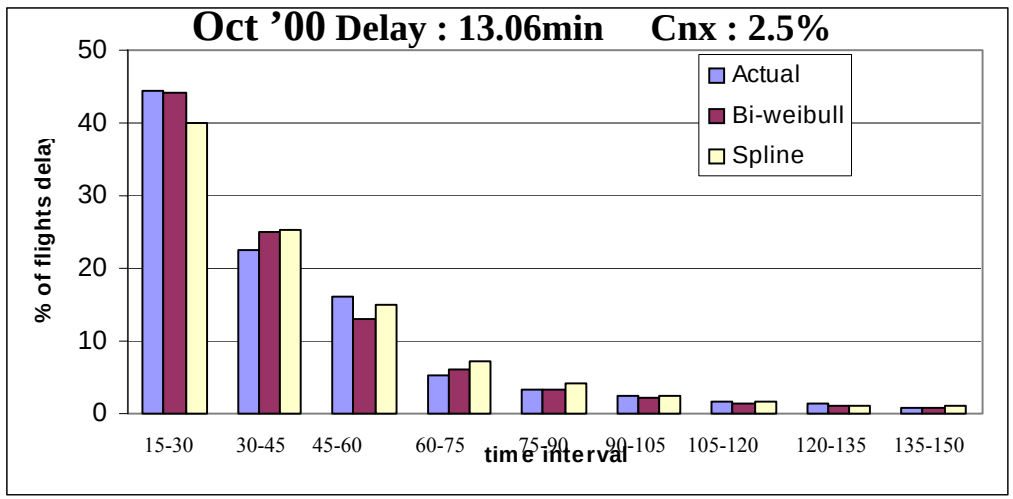
1) Fit one family of distribution for each month
2) Regress shape parameters to obtain distribution of flight delays as a functional form of average delay and cancellation probability

1) Regress each time interval to have a functional form in terms of average delay and cancellation probability.
2) Fit a smoothing spline to the time interval.
3) Normalize it to obtain distribution of flight delays

Comparison of two methods

- Data calibrated over 48 months ranging from Jan'00 to Dec '04
- 10 months used for validation – 5/00,10/00,7/01,4/02, 10/02, 12/03, 7/03, 3/04, 8/04,11/04
- Both the curves fit well.
- But Bi-weibull distribution is marginally better than bezier curve because the rate of descent initially(15-30 and 30-45 min) in bi-weibull matches more closely to the actual than bezier curve.





Work in Progress

- Introducing a metric (similar to Rho) for enroute congestion

- Evaluate the impact of demand growth at various airport categories on NAS Rho ➔ Delays and Cancellation