# Modeling delays and cancellations for collaborative strategic planning at single airports

Avijit Mukherjee[1], David Lovell[1,2], Michael Ball[1,3], Andrew Churchill[2], Amedeo Odoni[4,5]

[1] Institute for Systems Research
[2] Department of Civil and Environmental Engineering
[3] R.H. Smith School of Business
*University of Maryland, College Park*

[4] Department of Aeronautics and Astronautics
[5] Department of Civil and Environmental Engineering
*Massachusetts Institute of Technology*

National Airspace System Performance Workshop, Asilomar Conference Center, Pacific Grove, CA, March 2006

# Outline

**I. Strategic planning context**

**II. Delay model validation**

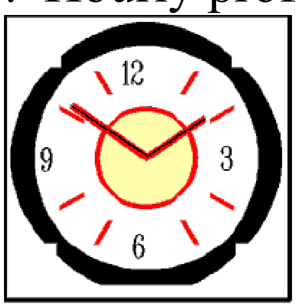
II.1. DELAYS model

II.2. Data filtering

II.3. Experimental design

II.4. Profile matching

II.5. Hourly profile plots



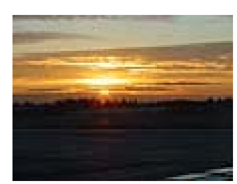

**III. Cancellation model**

III.1. Network flow model

III.2. Daily plots

III.3. Hourly profile plots


**IV. Conclusions**



# I. Strategic planning context

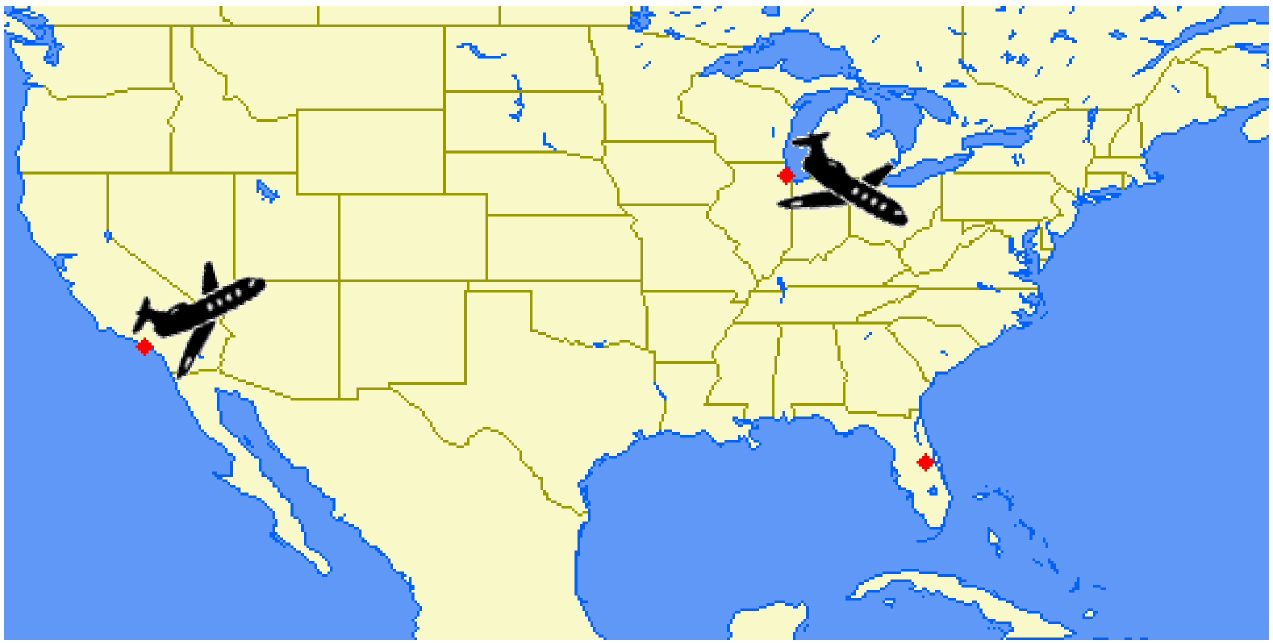
- Multiple carriers
  - Input data consist of scheduled flights only
  - Broker required to conserve confidentiality and prevent collusion
  - Broker produces estimates of delays and cancellations
- Single airport
  - Historical norms reliable in the rest of the NAS
  - Limited up and downstream interaction effects
- Applications
  - Collaborative scheduling
  - Strategic simulations
  - Evaluation of market mechanisms for congestion mitigation

# II.1. DELAYS model

- Model the aircraft arrival process as a non-homogeneous Poisson process with Erlang-*r* service times (DELAYS© code, developed at MIT by Koopman, Kivestu, Malone)
- How DELAYS works:
  - It is not a simulation
  - Governing differential equations of the stochastic process are generated
  - An efficient approximation scheme is then used to evaluate them
- Stochastic model produces pdf's for relevant outputs
- Can only capture congestion-related delays at an arrival airport
- We would use the model conditionally on each of several capacity scenarios relevant for the airport in question
- For validation purposes, we are trying to compare actual arrival delay information from ASPM to predicted delays

# II.2. Data filtering

## Aircraft XYZ



Departs LAX 60 minutes after scheduled

Arrives ORD 60 minutes after scheduled

Extra turn-around time of 30 minutes required

Departs ORD 90 minutes after scheduled

Arrives MCO 120 minutes after scheduled

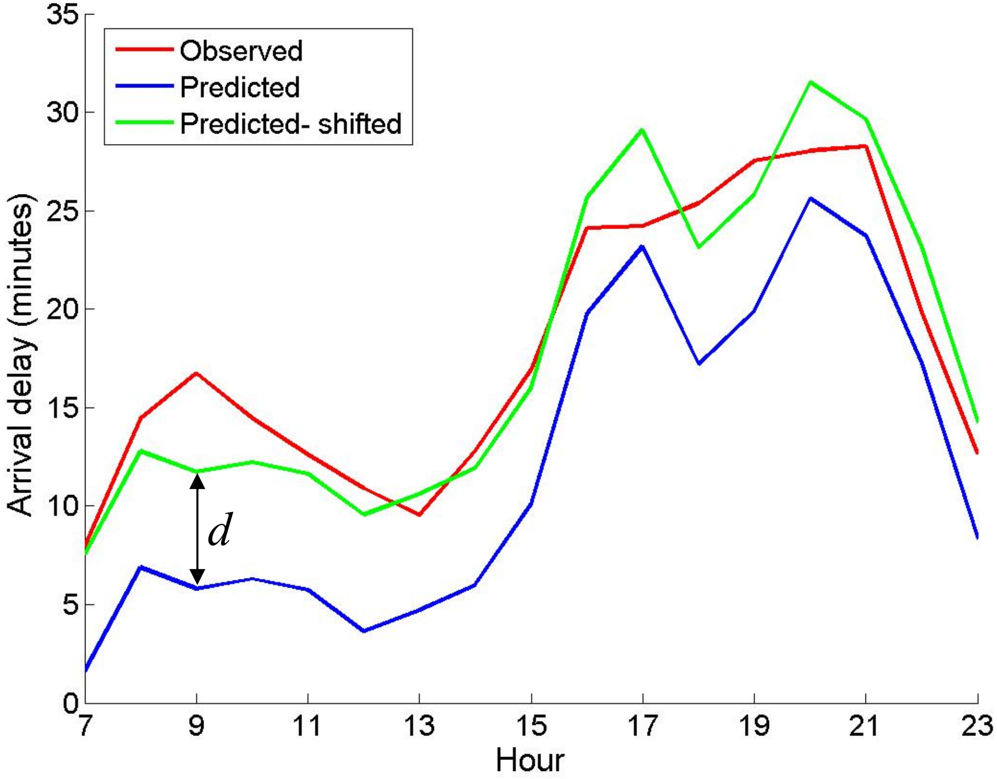
Reported delay for ORD-MCO segment: 120 minutes, includes 90 minutes of propagated delay

*Real* delay for ORD-MCO segment: 30 minutes

# II.3. Experimental design

- Airports:
  - Chicago O'Hare (ORD)
  - Atlanta Hartsfield (ATL)
- Time periods:
  - Monthly aggregation
  - January through December, 2004
- Inputs:
  - Demands = scheduled demands – cancellations
  - Capacities = AARs – unscheduled demand

# II.4. Profile matching




$$\min\left\{ f = \sum_{t=1}^{T} \left(O_t - P_t - d\right)^2 \right\}$$

*Example data:*
ATL, February 2004

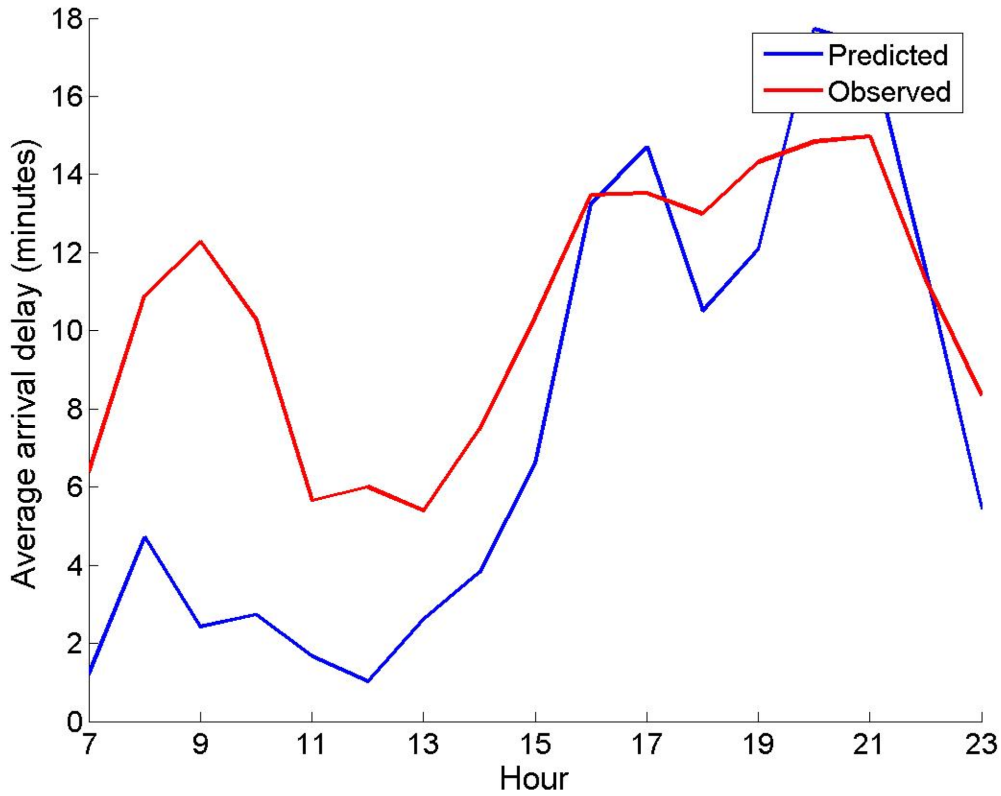
Predicted profile shifted up by 5.9 minutes, residuals of 100.4

Profile shape: primarily congestion impacts

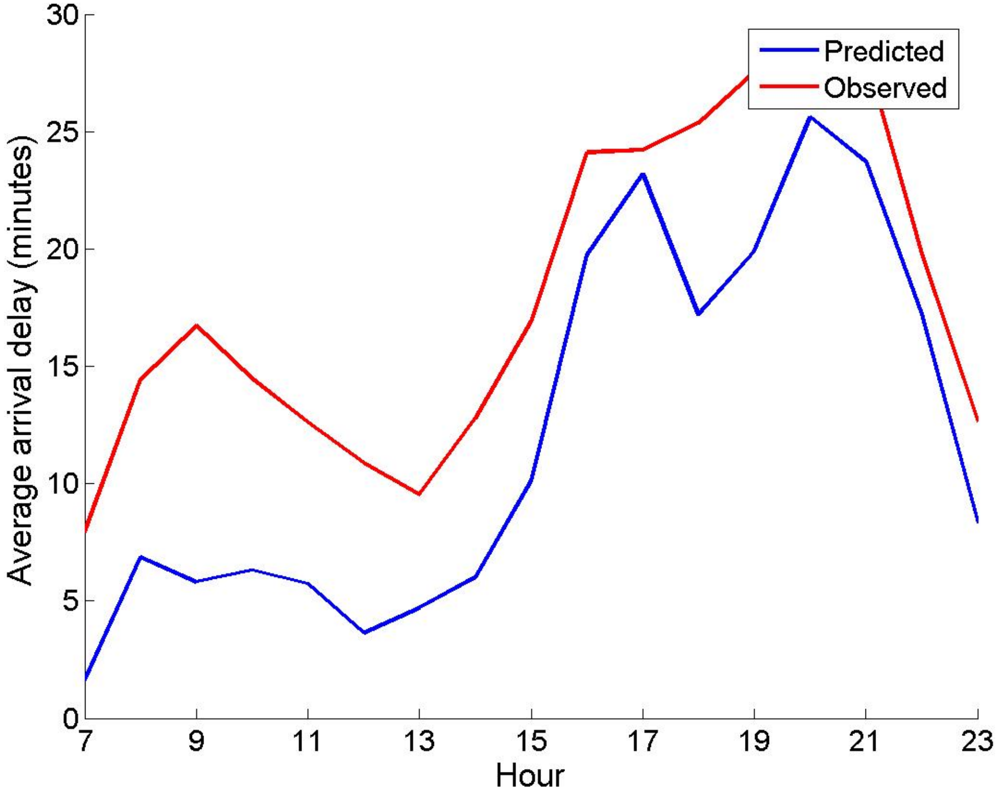
Profile magnitude: contains ambient causes

# II.5. Hourly profile plots

ATL, January 2004
Shift = 2.9
Residuals = 187.9

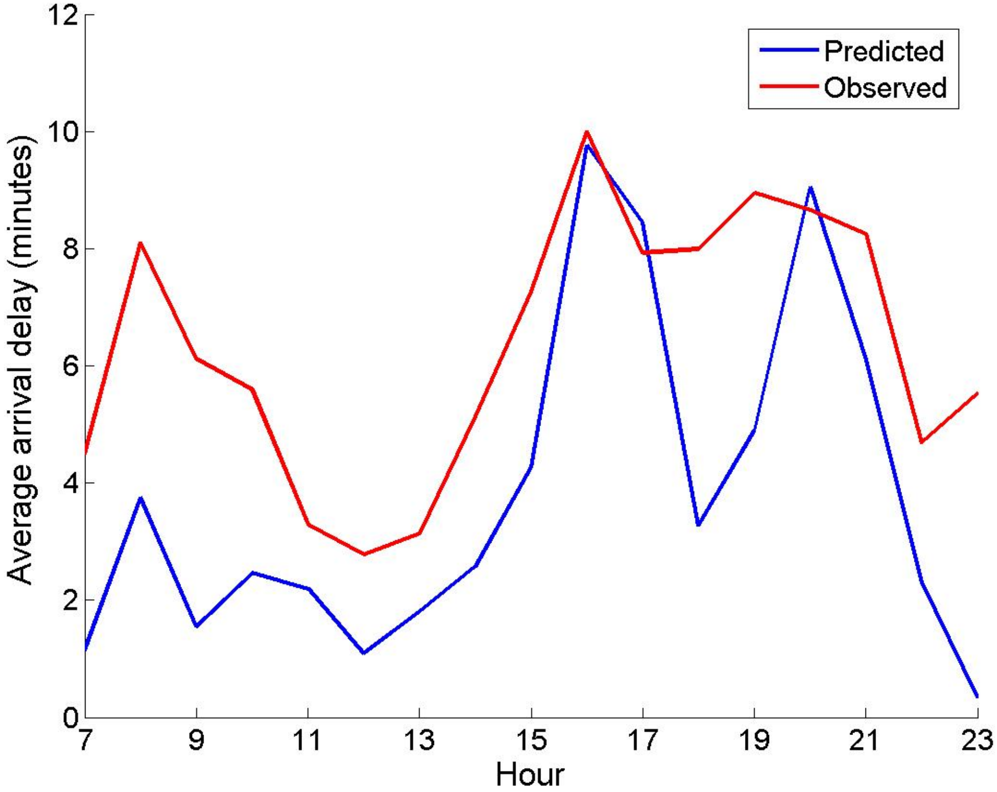


# II.5. Hourly profile plots



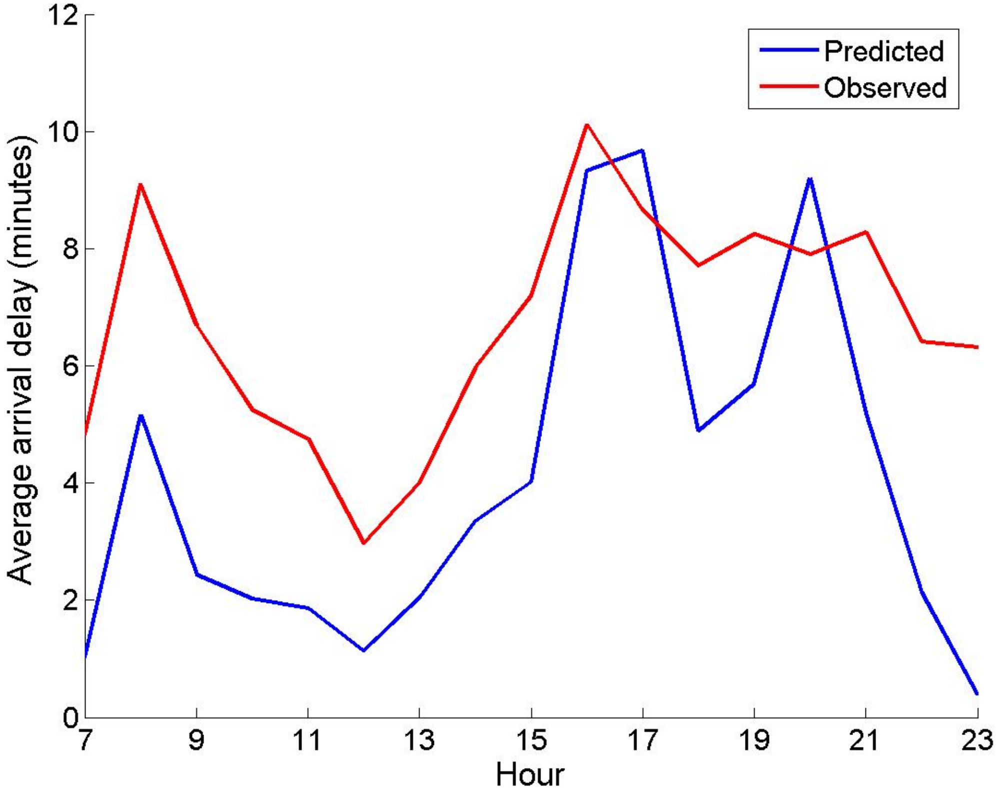

ATL, February 2004
Shift = 5.9
Residuals = 100.4

# II.5. Hourly profile plots



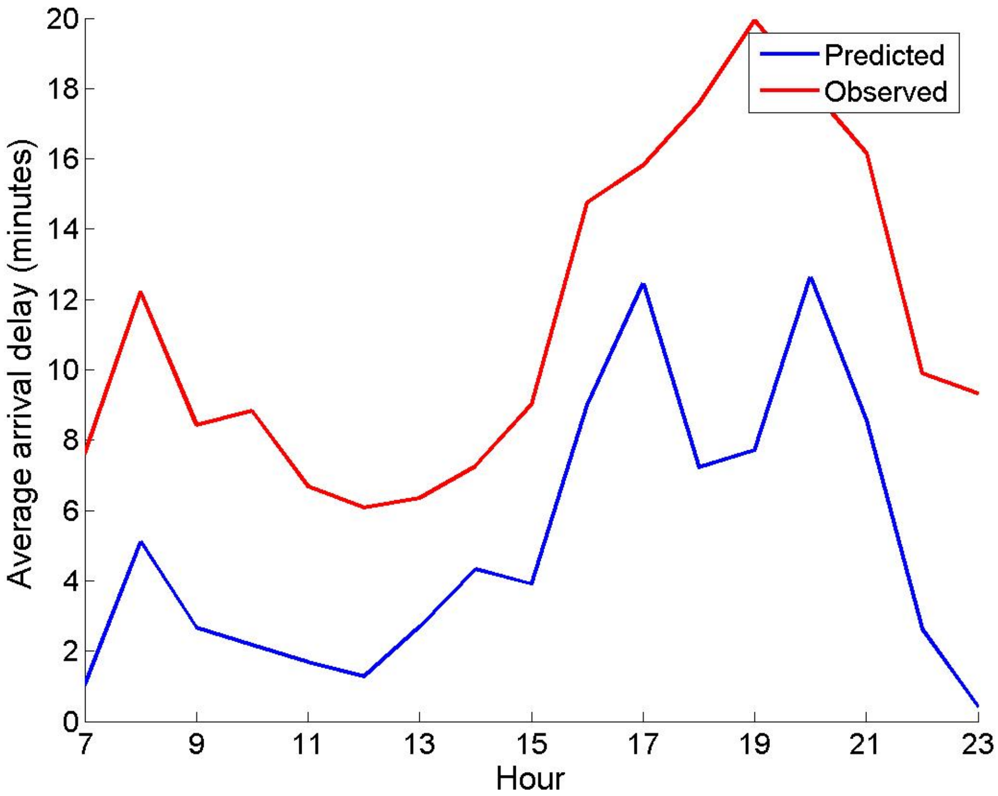

ATL, March 2004
Shift = 2.5
Residuals = 50.3

# II.5. Hourly profile plots



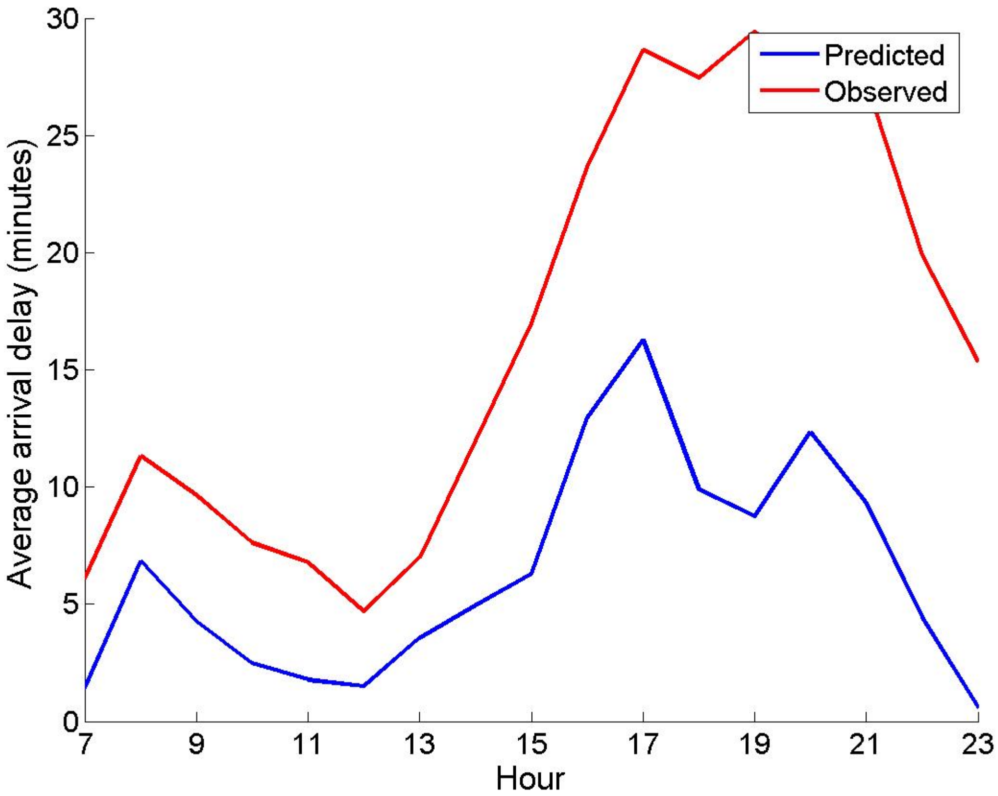

ATL, April 2004
Shift = 2.6
Residuals = 53.5

# II.5. Hourly profile plots




ATL, May 2004
Shift = 6.4
Residuals = 95.4

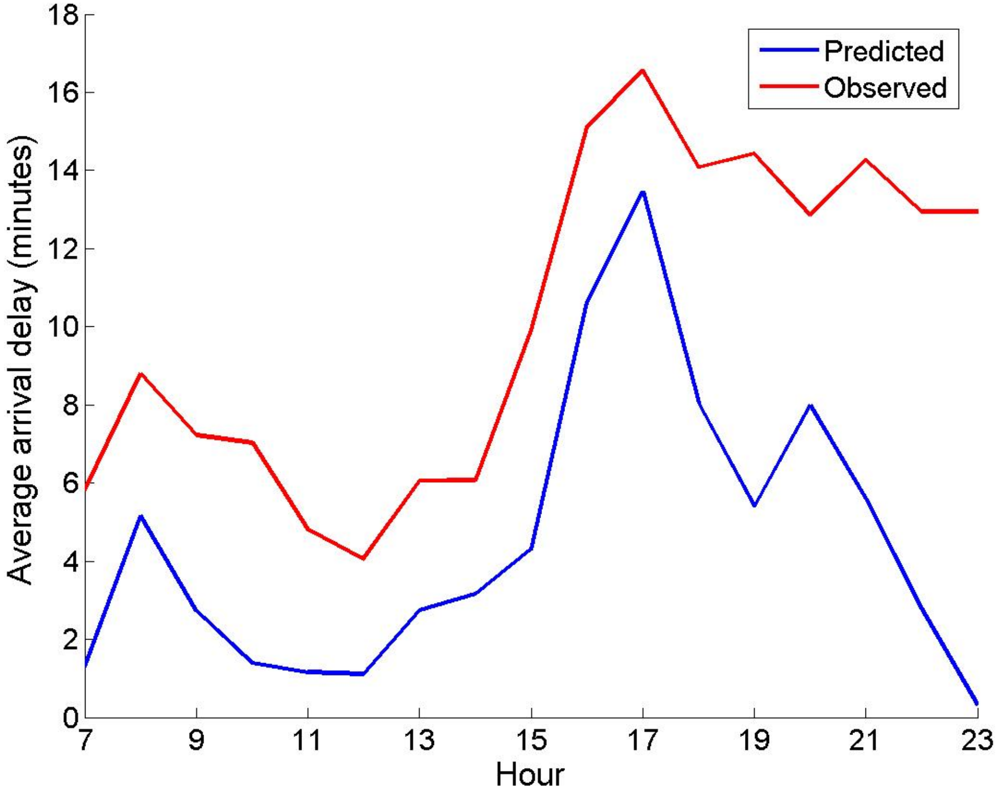
# II.5. Hourly profile plots



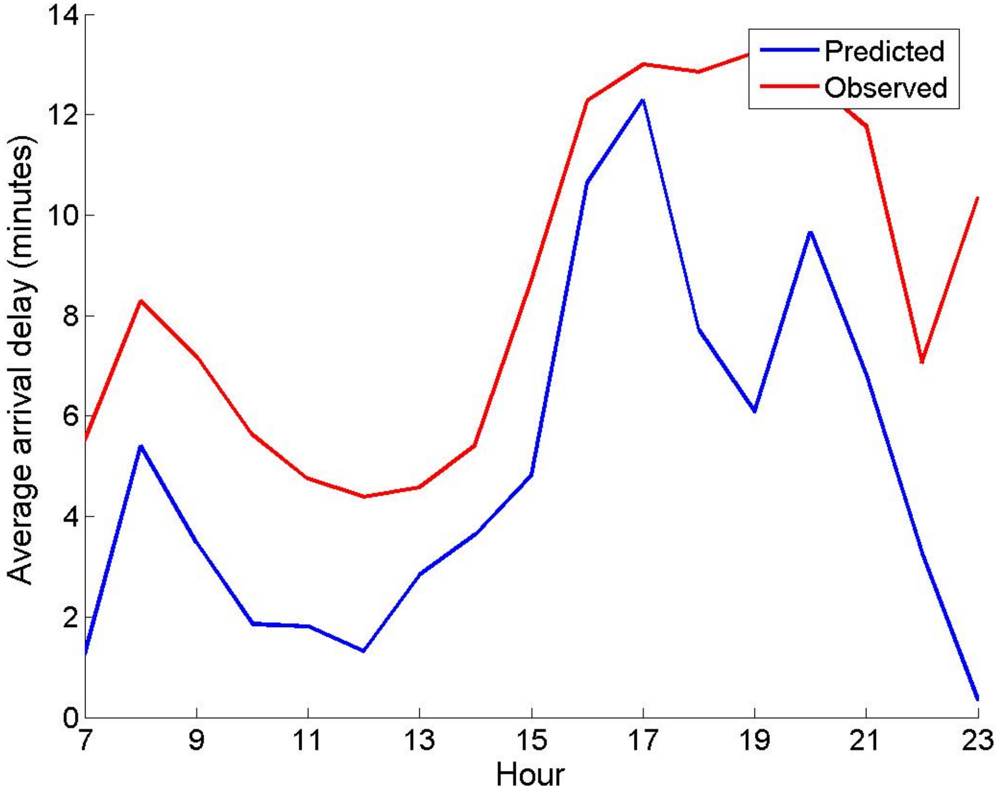

ATL, June 2004
Shift = 10.2
Residuals = 546.8

# II.5. Hourly profile plots

ATL, July 2004
Shift = 5.6
Residuals = 129.1



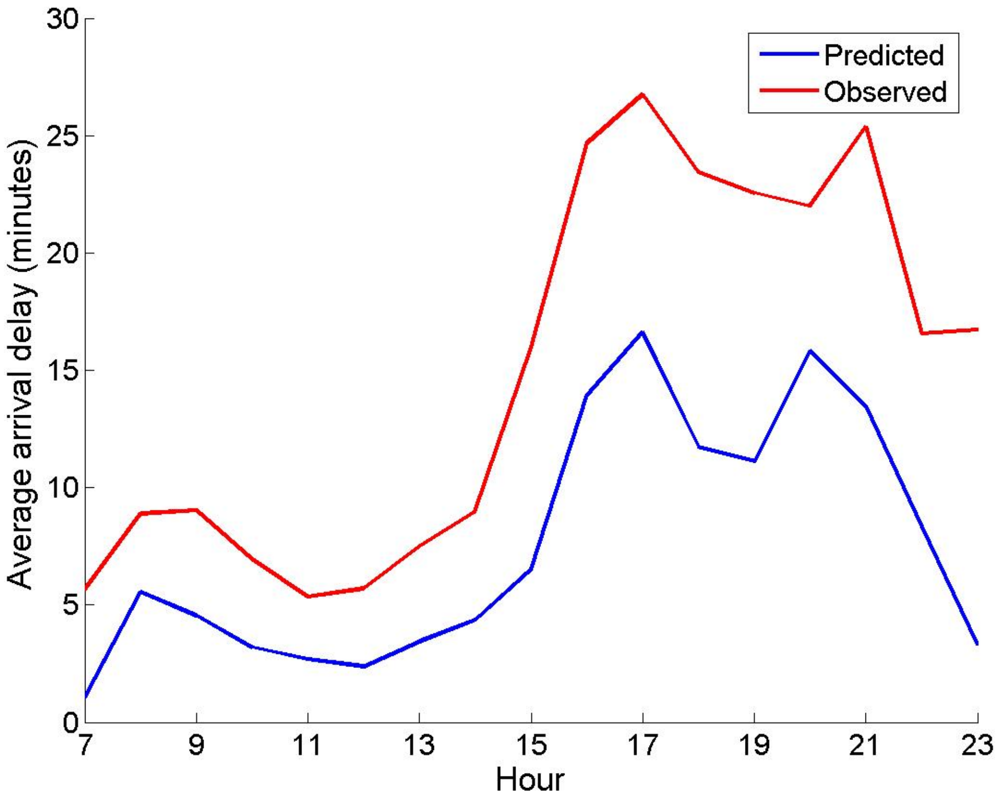
# II.5. Hourly profile plots



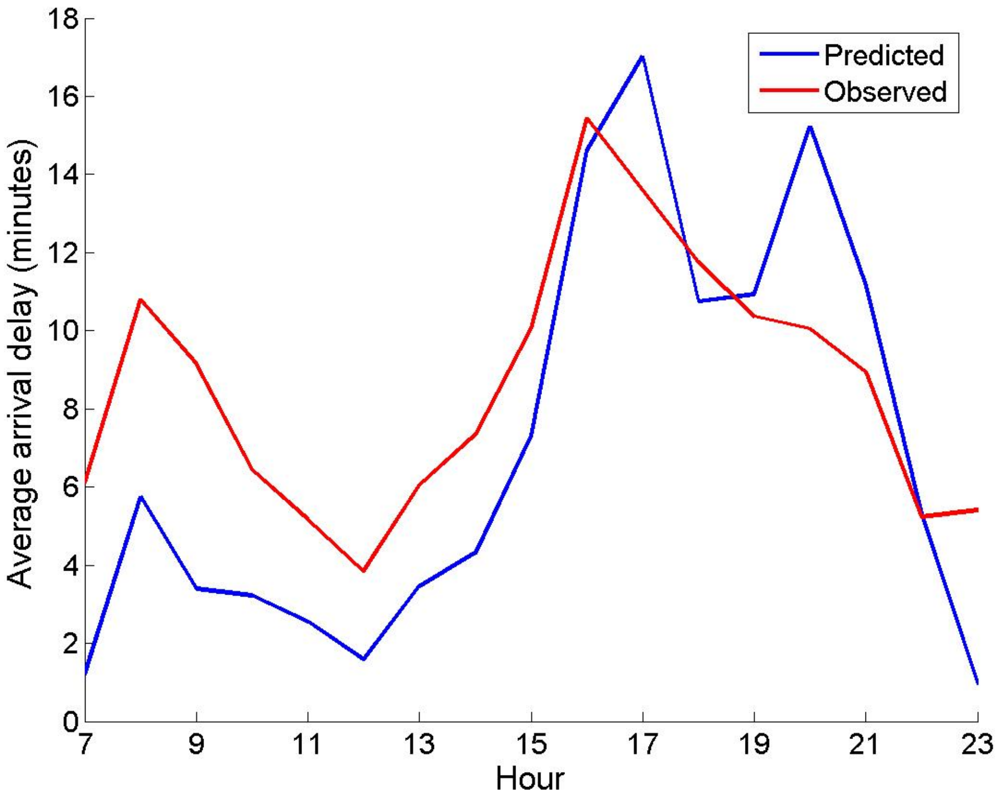

ATL, August 2004
Shift = 3.8
Residuals = 78.6

# II.5. Hourly profile plots

ATL, September 2004
Shift = 7.3
Residuals = 220.8



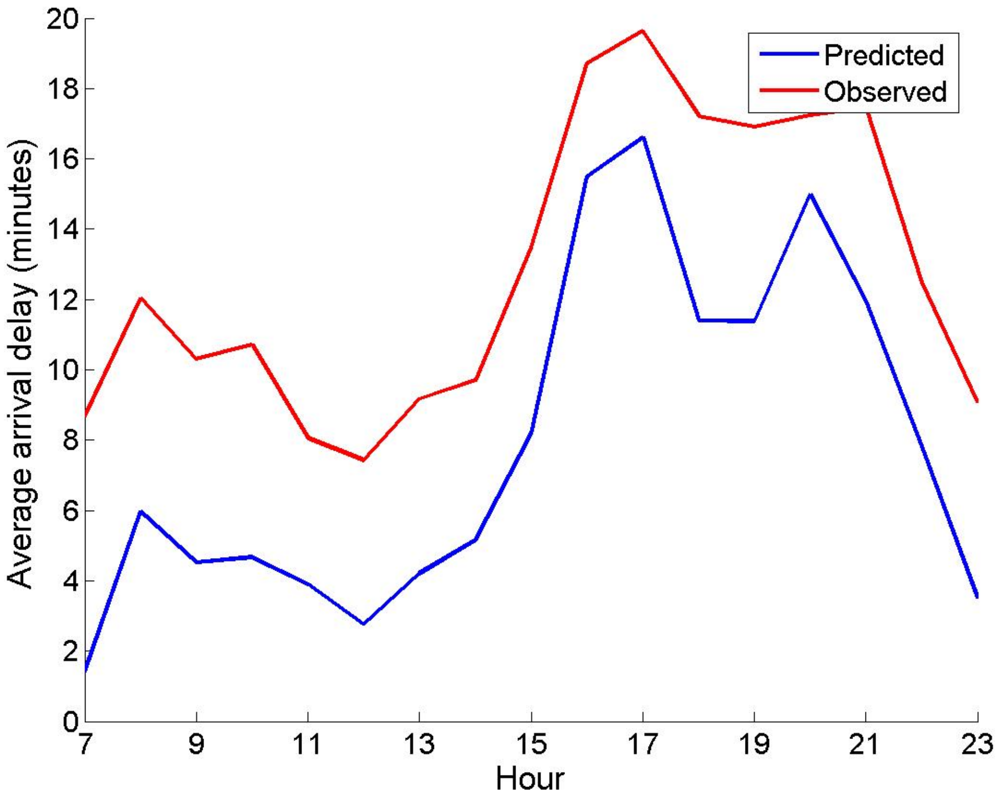
# II.5. Hourly profile plots



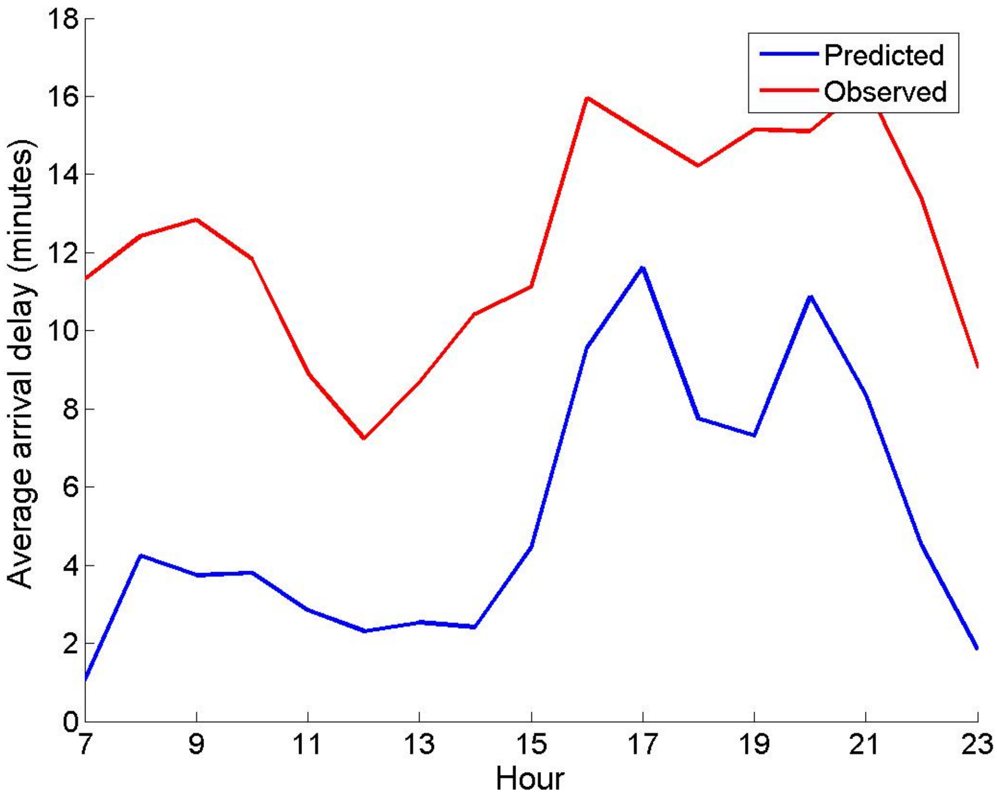

ATL, October 2004
Shift = 1.6
Residuals = 150.9

# II.5. Hourly profile plots

ATL, November 2004
Shift = 5.0
Residuals = 25.3

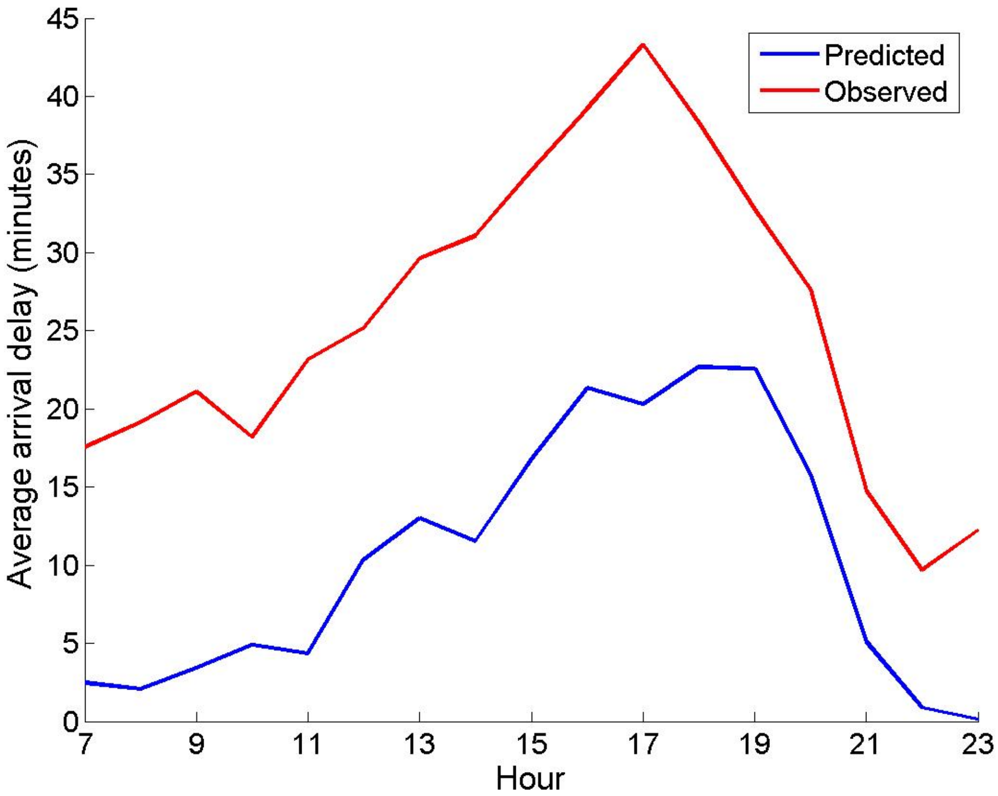


# II.5. Hourly profile plots




ATL, December 2004
Shift = 7.0
Residuals = 50.5

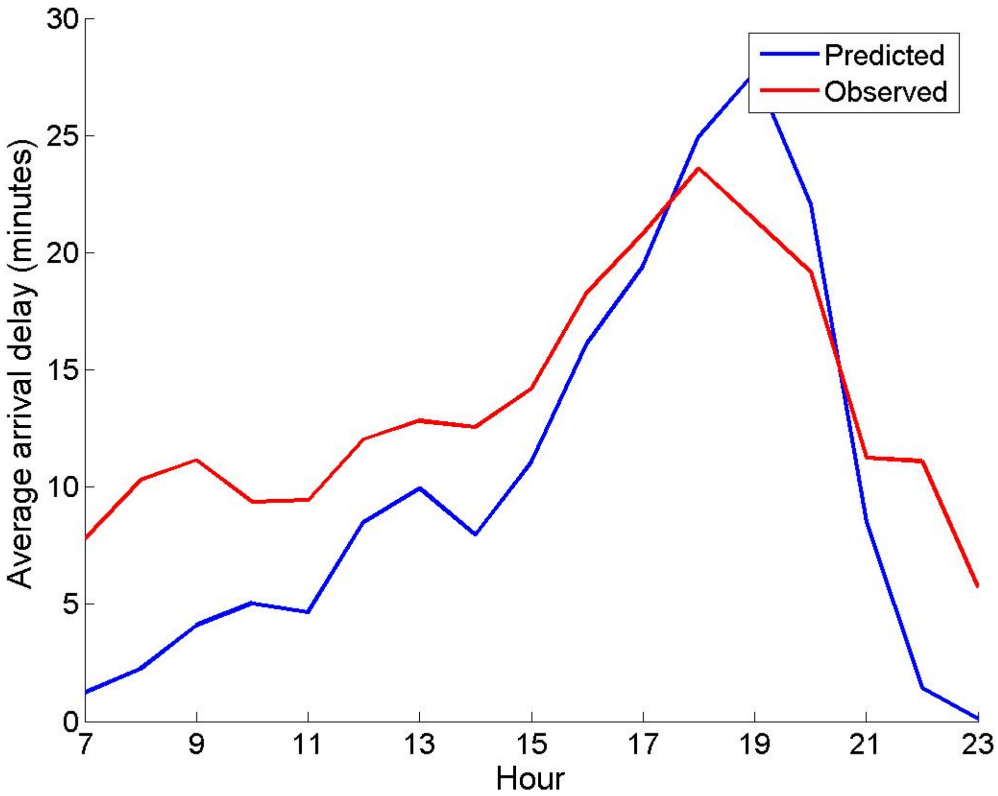
# II.5. Hourly profile plots



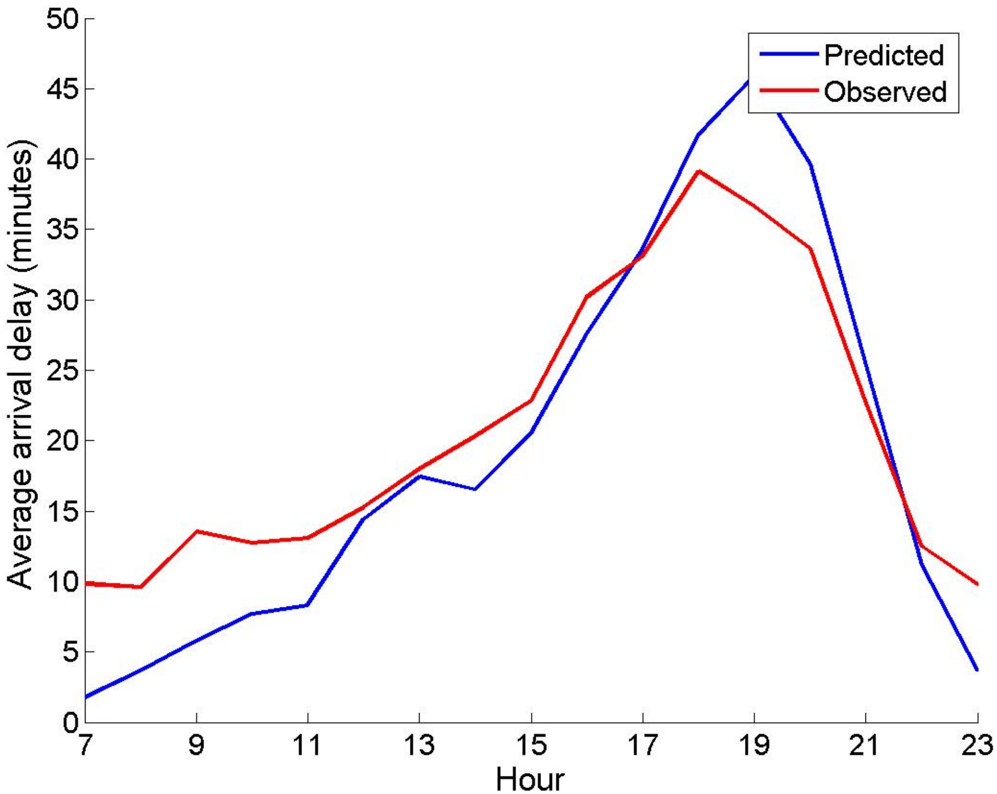

ORD, January 2004
Shift = 15.3
Residuals = 243.9

# II.5. Hourly profile plots

ORD, February 2004
Shift = 3.3
Residuals = 255.0



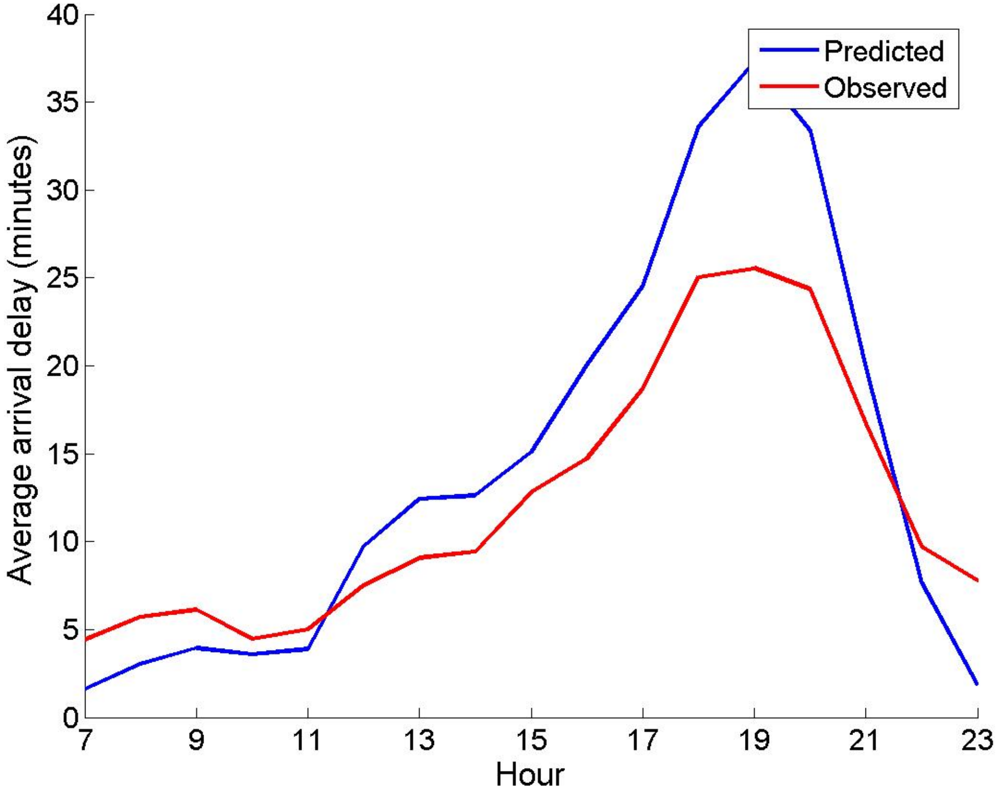
# II.5. Hourly profile plots



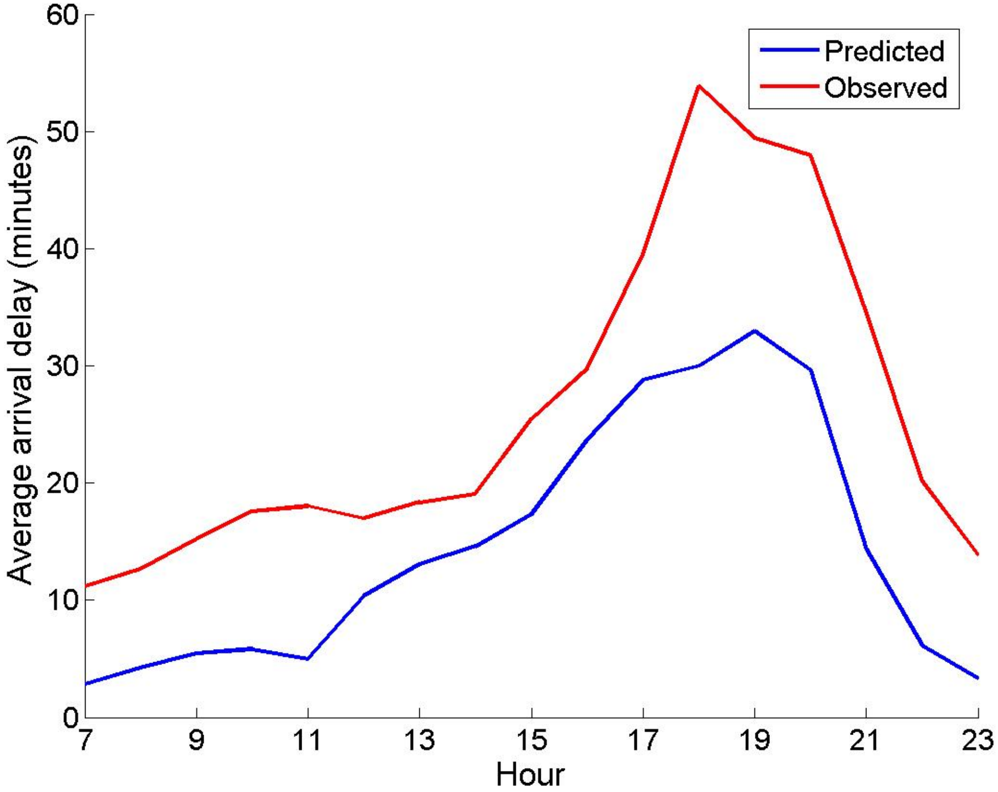

ORD, March 2004
Shift = 1.6
Residuals = 364.9

# II.5. Hourly profile plots

ORD, April 2004
Shift = -2.2
Residuals = 376.4

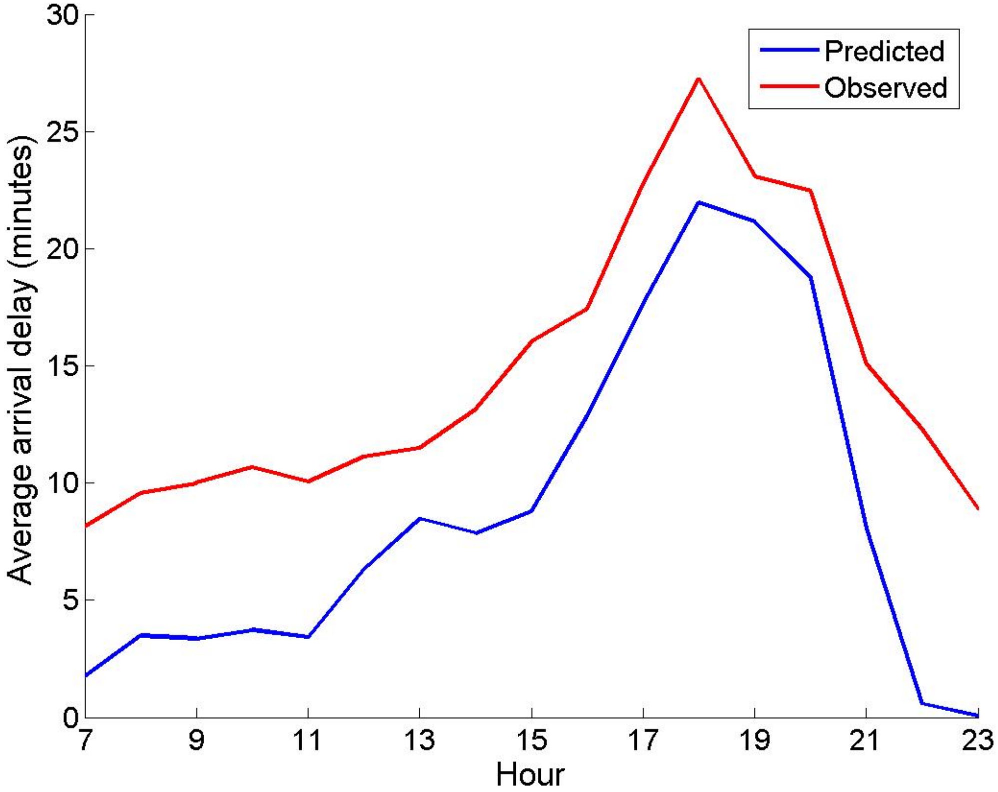


# II.5. Hourly profile plots



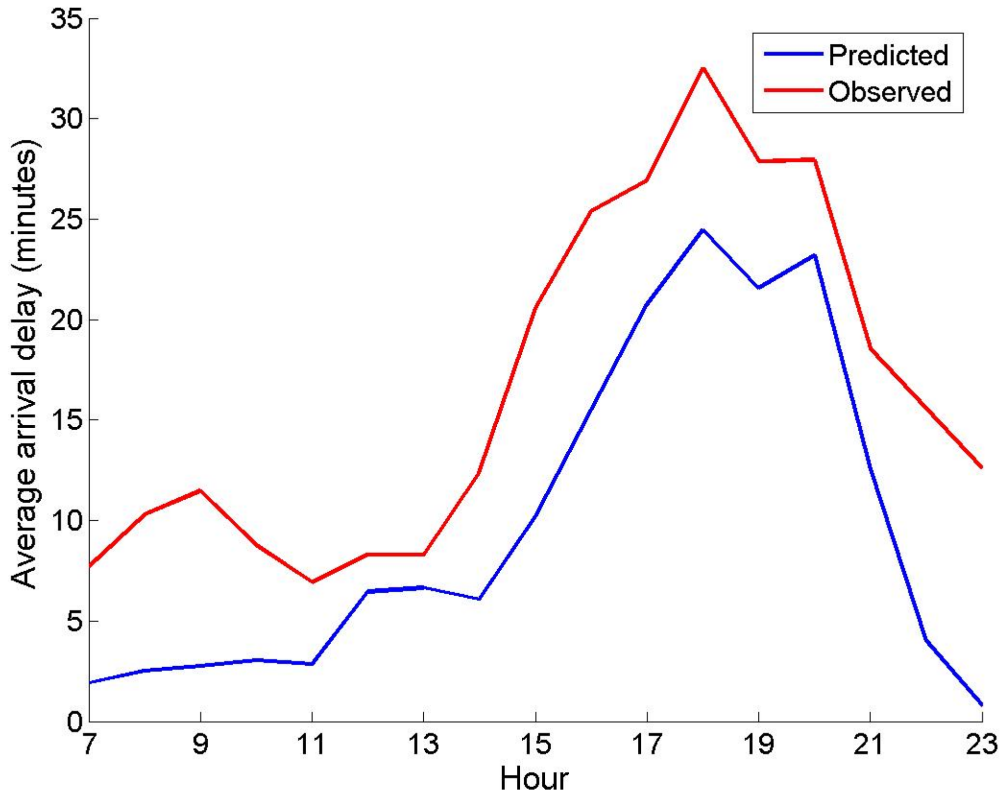

ORD, May 2004
Shift = 11.5
Residuals = 483.0

# II.5. Hourly profile plots



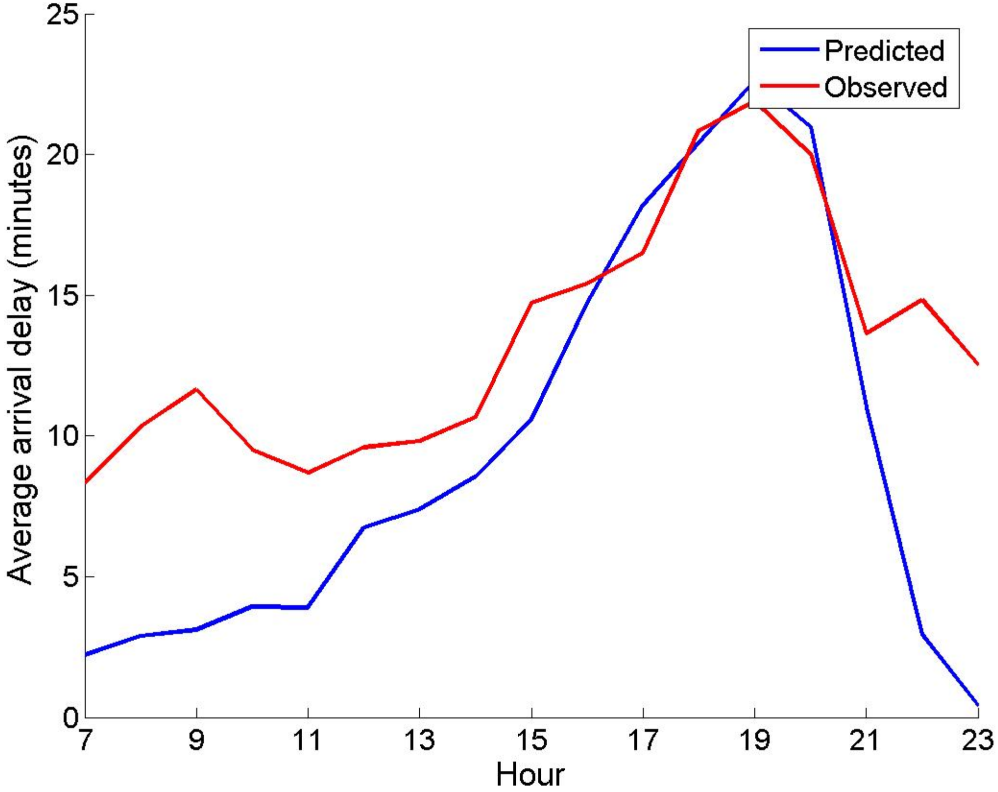

ORD, June 2004
Shift = 5.9
Residuals = 81.0

# II.5. Hourly profile plots



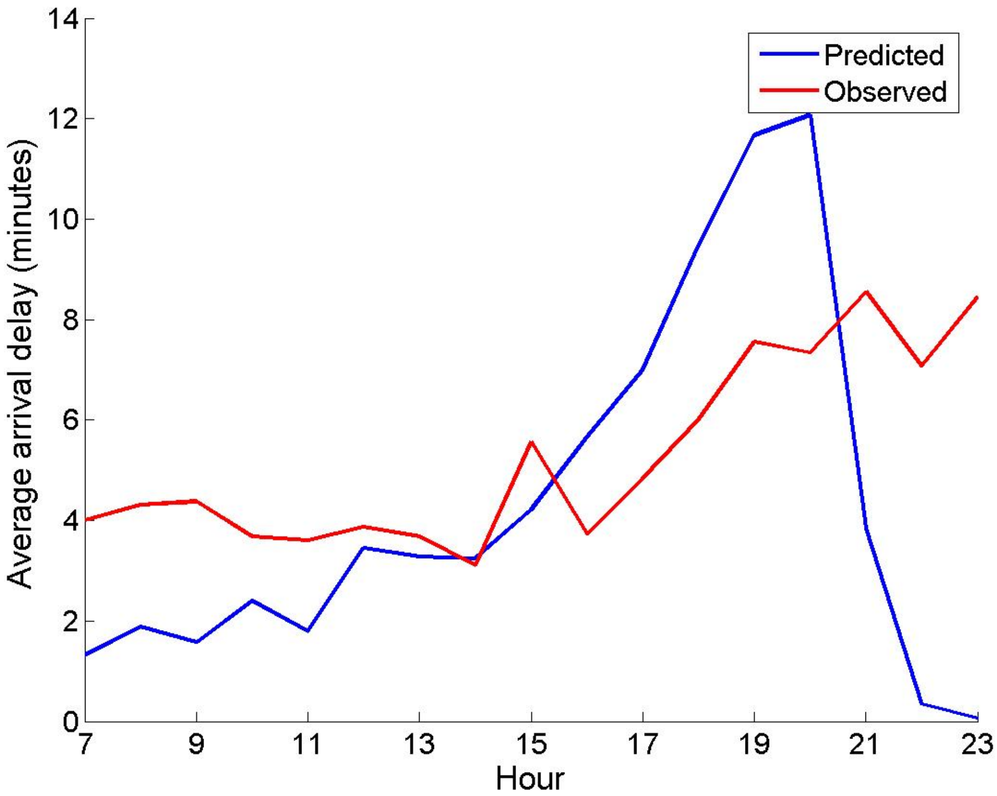

ORD, July 2004
Shift = 6.9
Residuals = 142.7

# II.5. Hourly profile plots



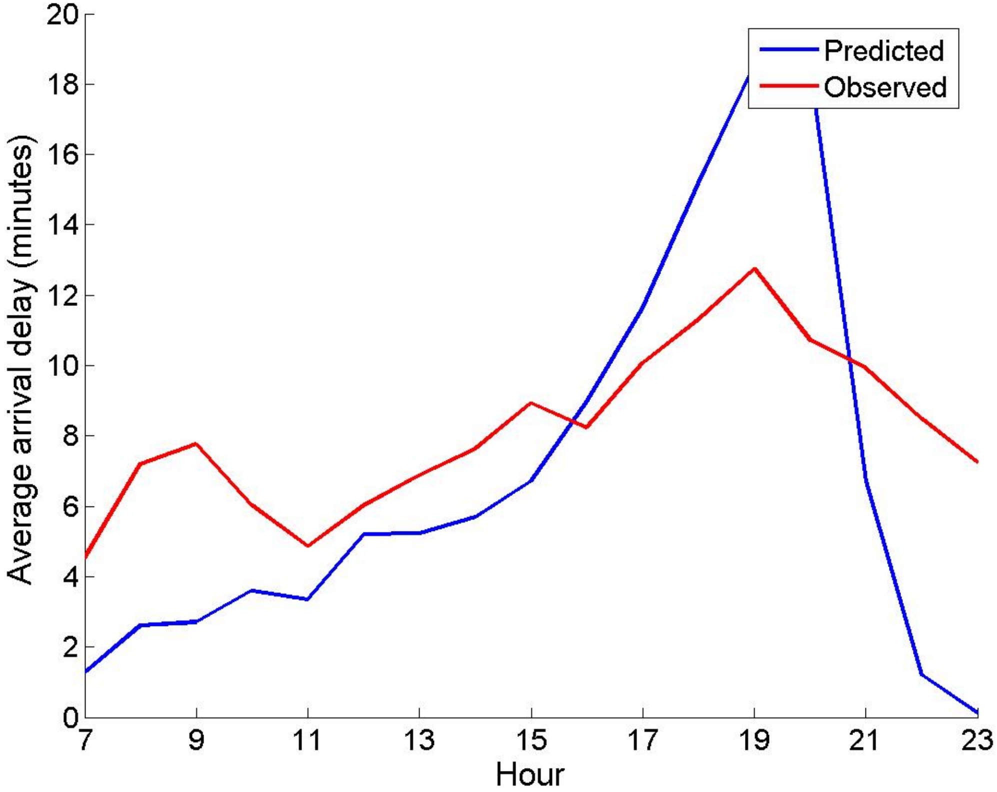

ORD, August 2004
Shift = 4.0
Residuals = 280.4

# II.5. Hourly profile plots



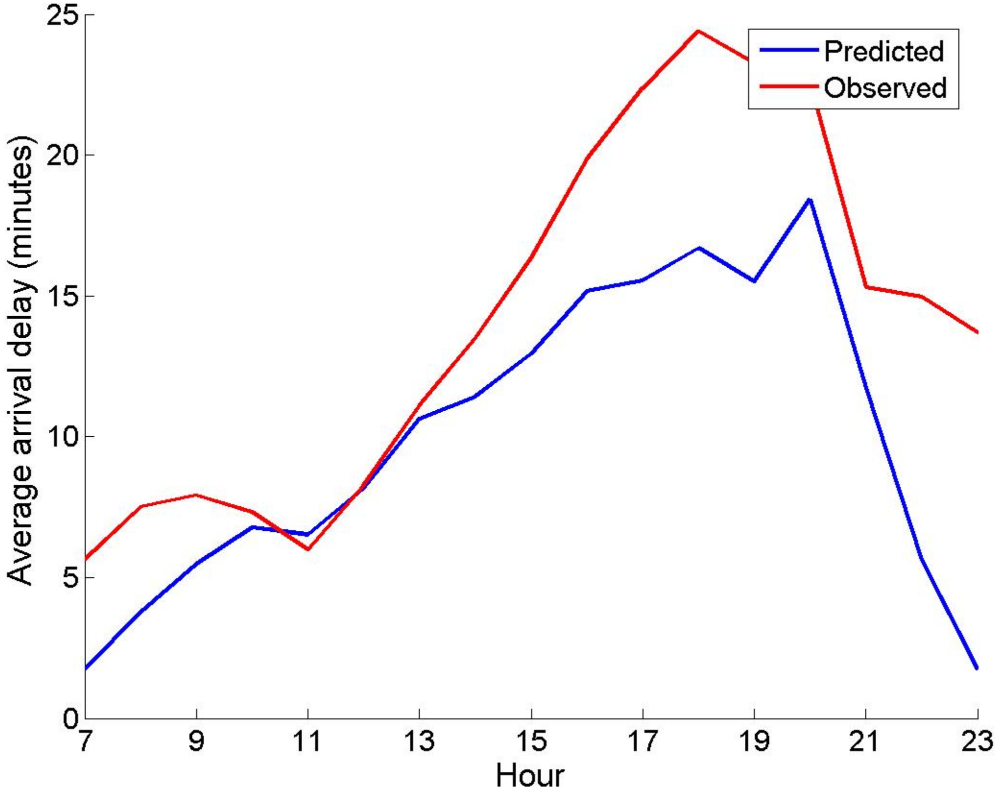

ORD, September 2004
Shift = 1.0
Residuals = 209.6

# II.5. Hourly profile plots




ORD, October 2004
Shift = 1.3
Residuals = 275.7

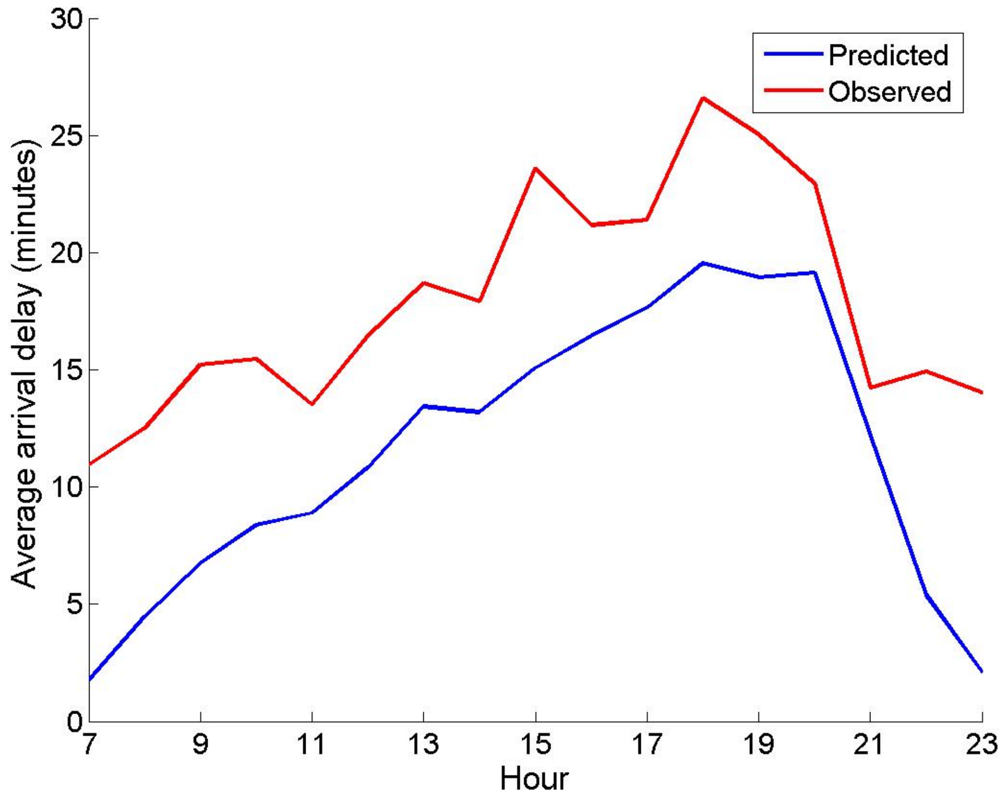
# II.5. Hourly profile plots



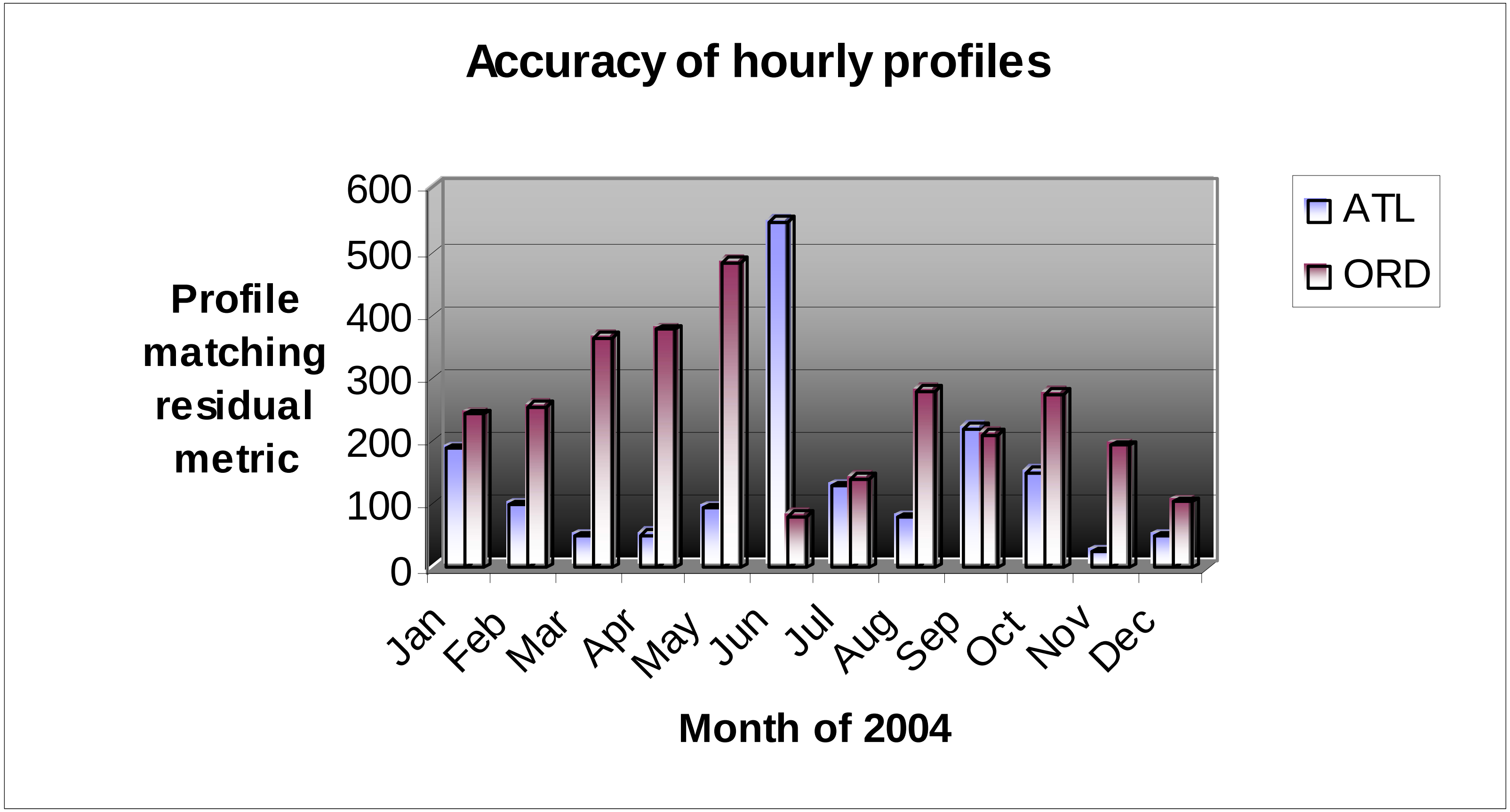

ORD, November 2004
Shift = 4.2
Residuals = 193.7

# II.5. Hourly profile plots

ORD, December 2004
Shift = 6.5
Residuals = 103.6



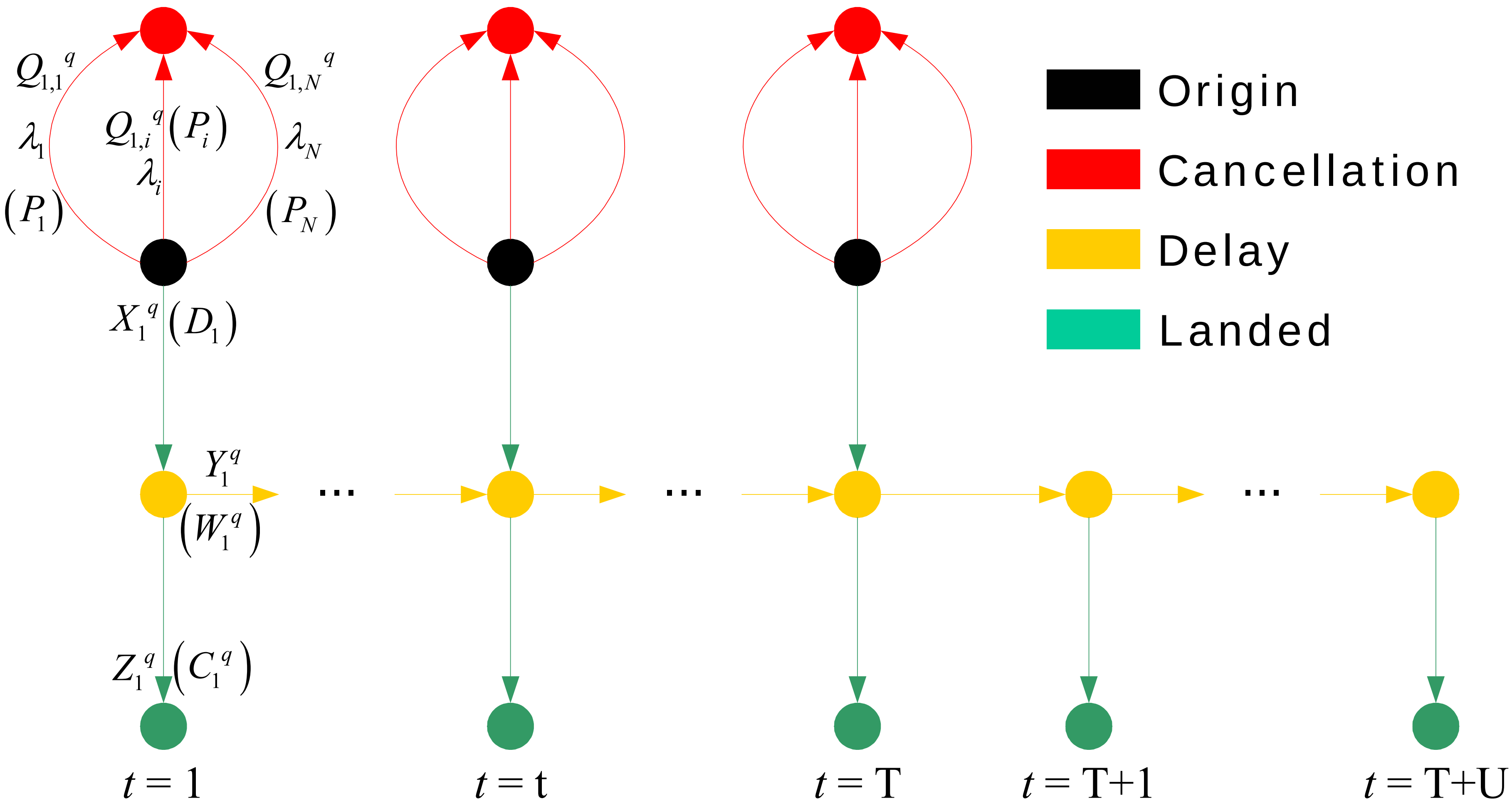
# II.5. Hourly profile plots



# III. Cancellation model

- Cast as a minimum cost network flow model (LP)
- Penalties in the objective function for delay arcs and for cancellations
- A maximum delay is imposed exogenously
- Calibration via known schedules, AARs, and cancellations from ASPM data

# III.1. Network flow model



# III.1. Model structure

- Minimum cost network flow problem

*Decision variables:*

$X_t$ = Flights accepted for landing
$Y_t$ = Delayed flights
$Z_t$ = Landed flights
$Q_{t,i}$ = Flights cancelled at cost $\lambda_i$

*Arc capacities:*

$D_t$ = Scheduled demand
$W_t$ = Transfer capacity
$C_t$ = Landing capacity
$P_i$ = Cancellations at cost $\lambda_i$

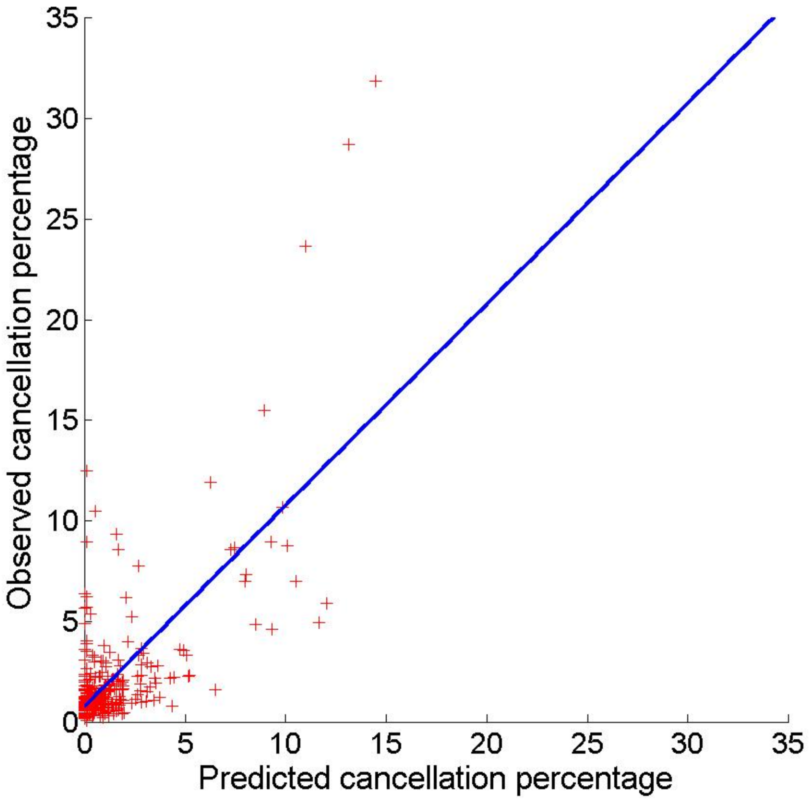
*Constants:*

$U$ = Maximum number of time slices a flight can be delayed
$\xi$ = Delay cost for one time slice, taken to be 1
$\lambda_i$ = Cancellation costs using cancellation arc $i$, relative to $\xi$

*Notes:*

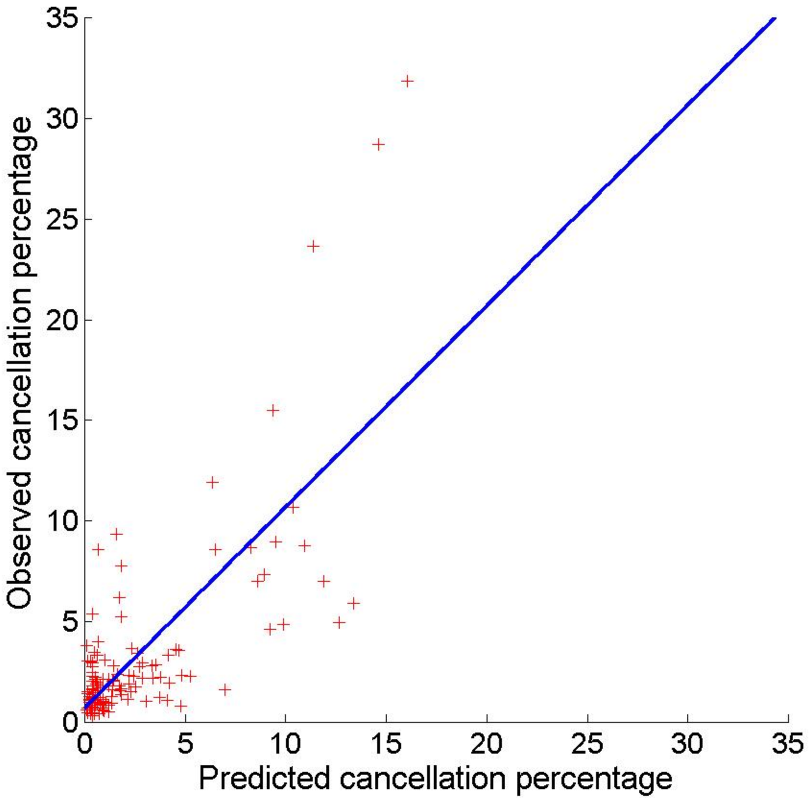
- Have $N$ cancellation arcs for each $t$
- No demand after time $T$
- No cancellation arcs after time $T$

# III.2. Daily plots



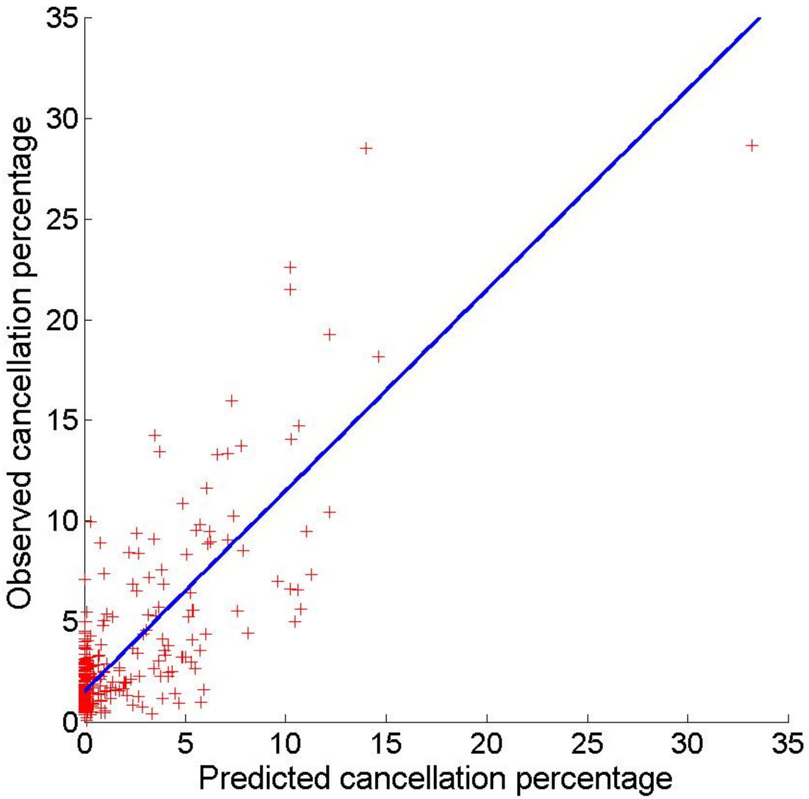

ATL2004, U=6, [9 18 36]
unfiltered
$y = 0.782 + 1.000x$
$R^2 = 0.497$

# III.2. Daily plots



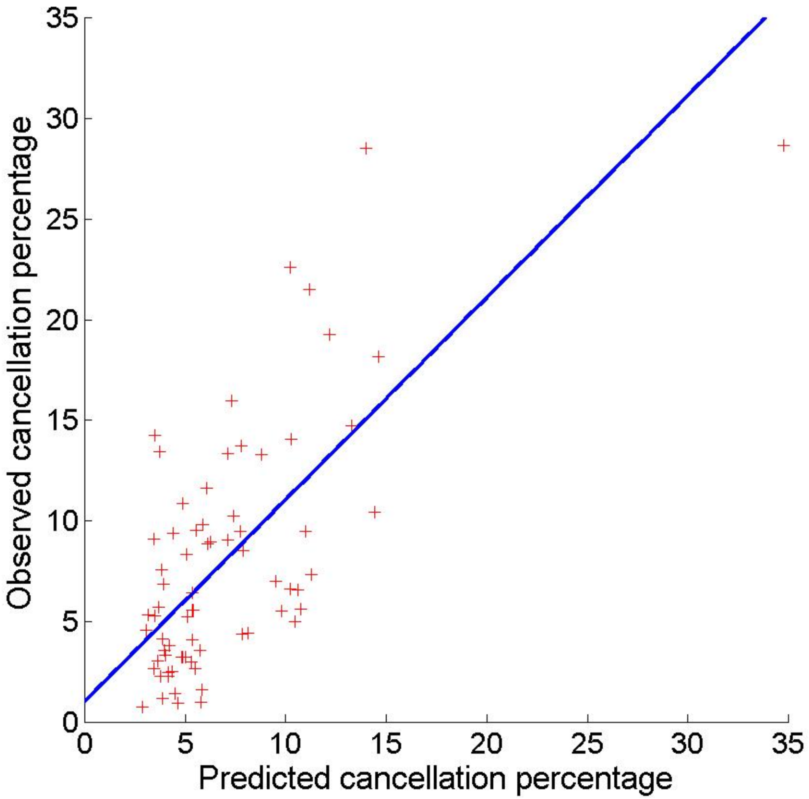

ATL2004, U=2, [15 30 60]
Filtered (>25th %ile DQ)
$y = 0.698 + 1.000x$
$R^2 = 0.561$

# III.2. Daily plots



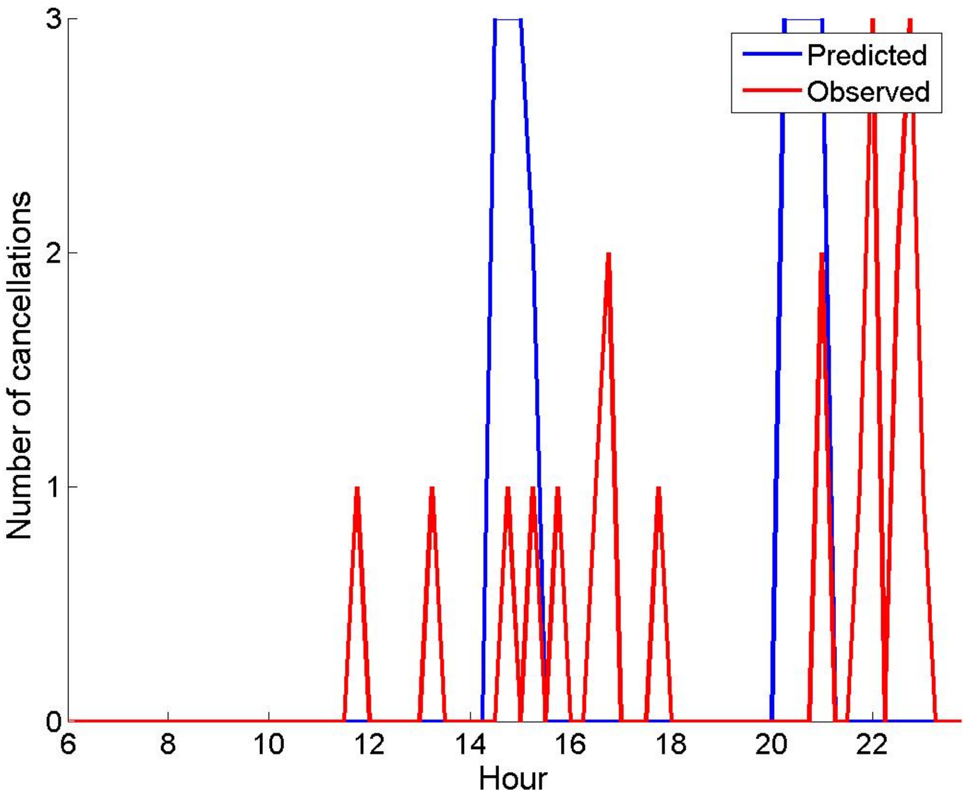

ORD2004, U=8, [18 36 72]
unfiltered
$y = 1.53 + 0.998x$
$R^2 = 0.678$

# III.2. Daily plots




ORD2004, U=6, [18 36 72]
Filtered (>25th %ile DQ)
$y = 1.02 + 1.005x$
$R^2 = 0.569$

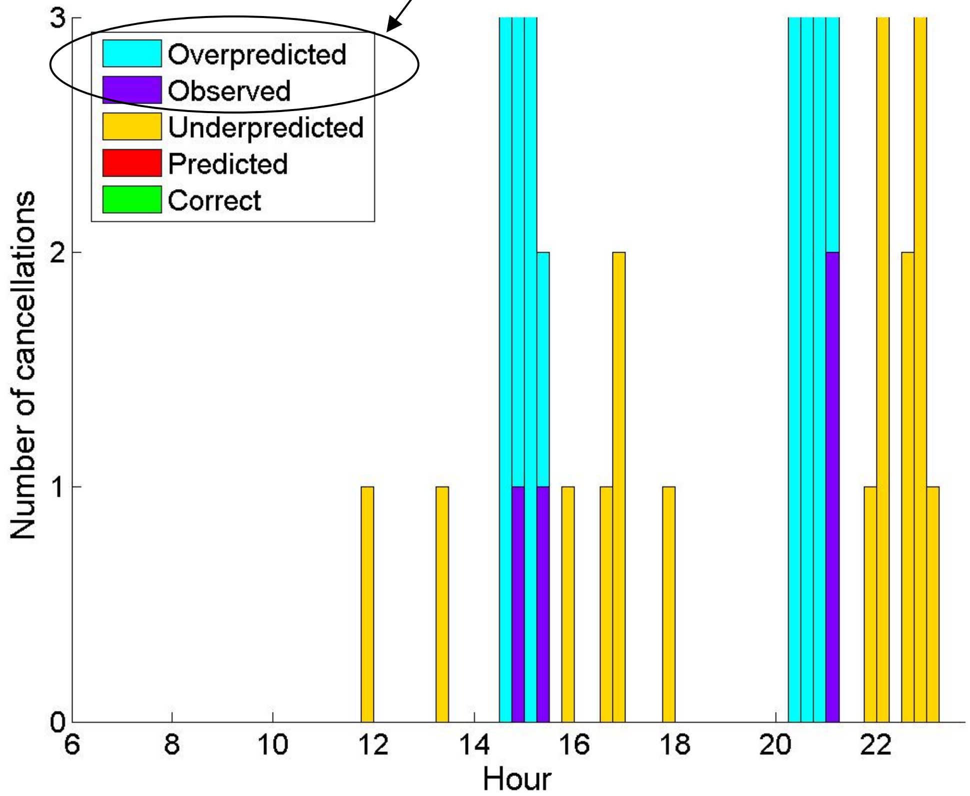
# III.3. Hourly profile plots




ATL2004, U=6, [9 18 36]
July 25, 2004
Predicted = 24
Observed = 21
Shift = -0.0312
Residuals = 84.9063

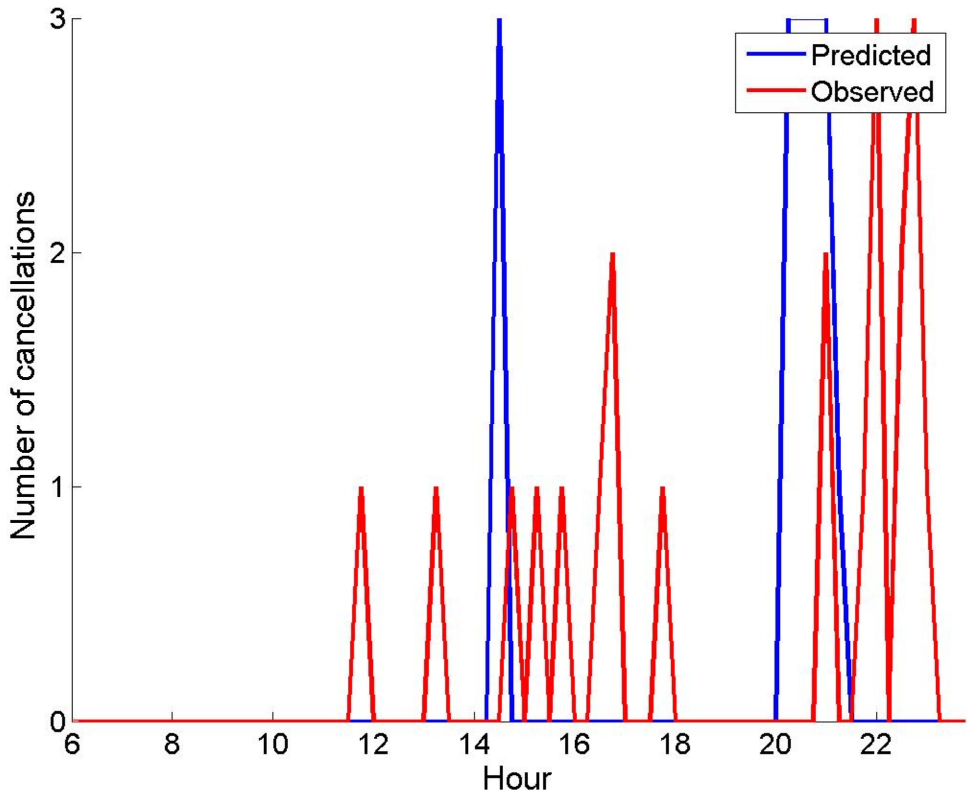
# III.3. Hourly profile plots

Prediction exceeds observed



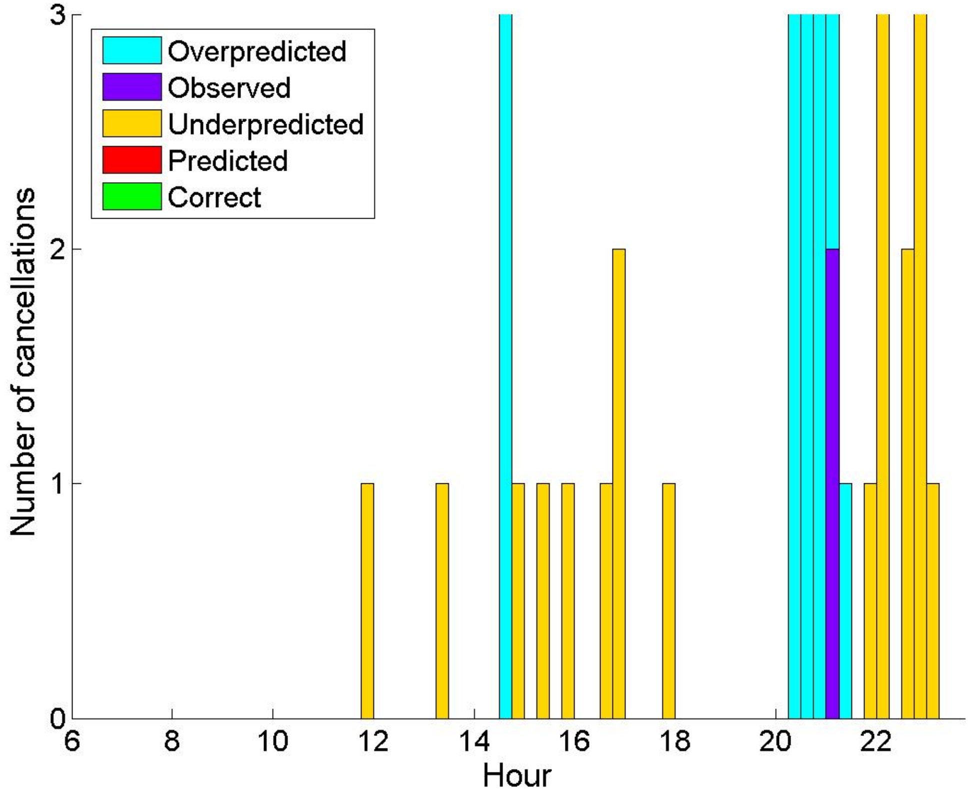

ATL2004, U=6, [9 18 36]
July 25, 2004
Predicted = 24
Observed = 21
Shift = -0.0312
Residuals = 84.9063

# III.3. Hourly profile plots



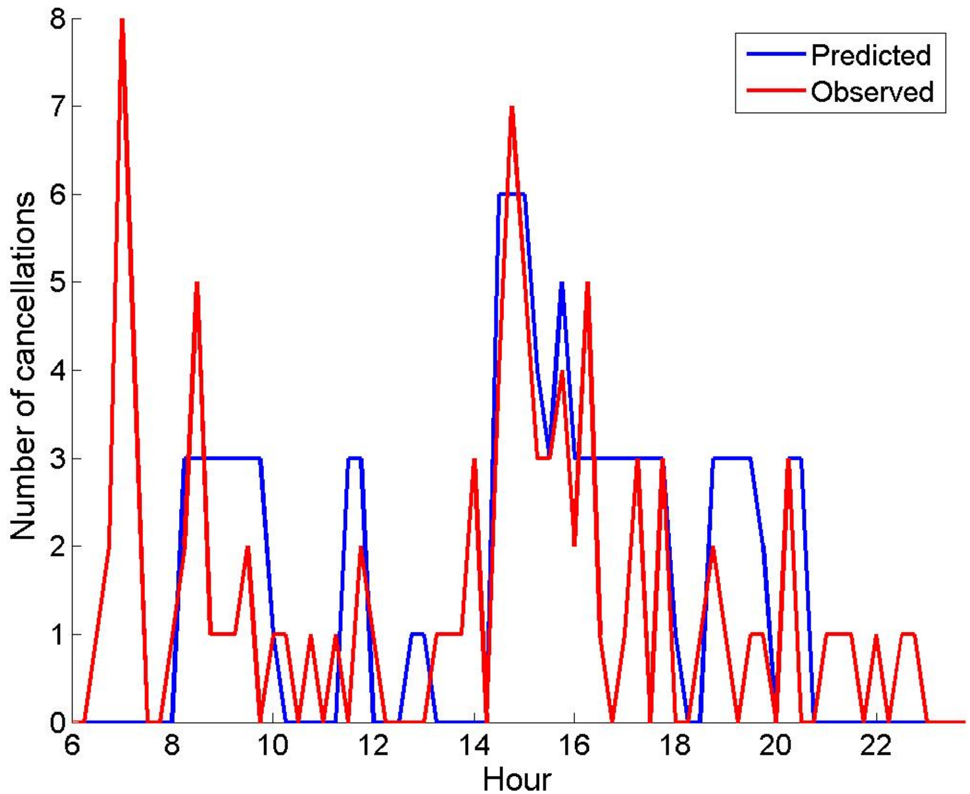

ATL2004, U=2, [15 30 60]
July 25, 2004
Predicted = 18
Observed = 21
Shift = 0.0313
Residuals = 76.9063

# III.3. Hourly profile plots




ATL2004, U=2, [15 30 60]
July 25, 2004
Predicted = 18
Observed = 21
Shift = 0.0313
Residuals = 76.9063

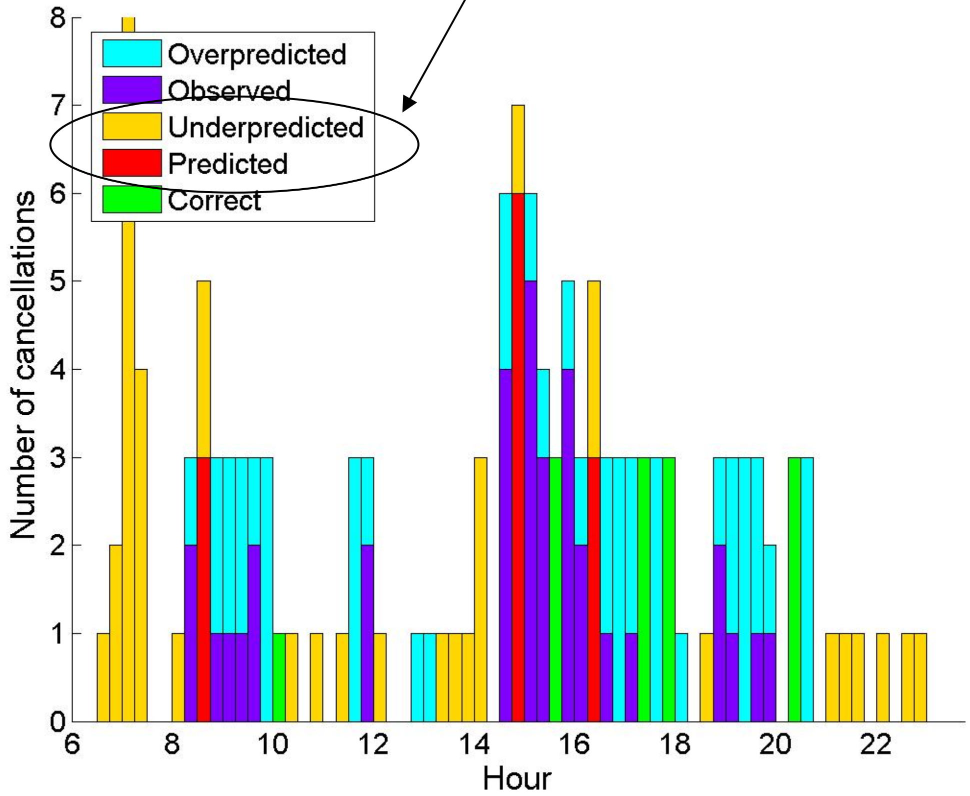
# III.3. Hourly profile plots




ATL2004, U=6, [9 18 36]
Sept 7, 2004
Predicted = 106
Observed = 97
Shift = -0.0938
Residuals = 216.1563

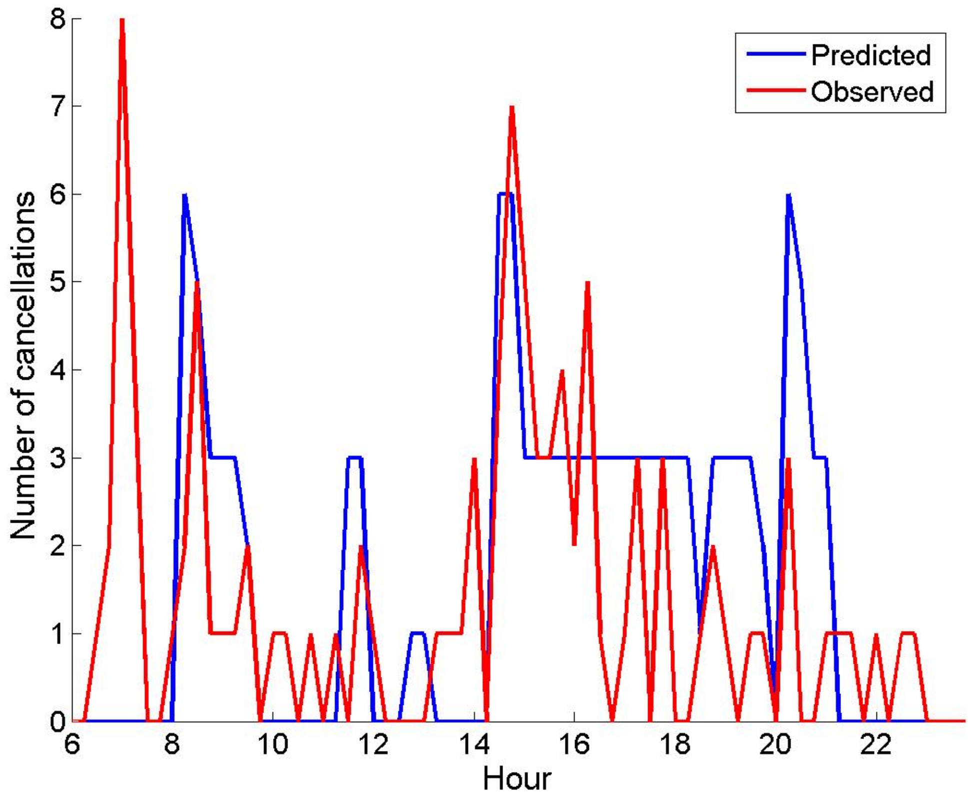
# III.3. Hourly profile plots

Observed exceeds prediction



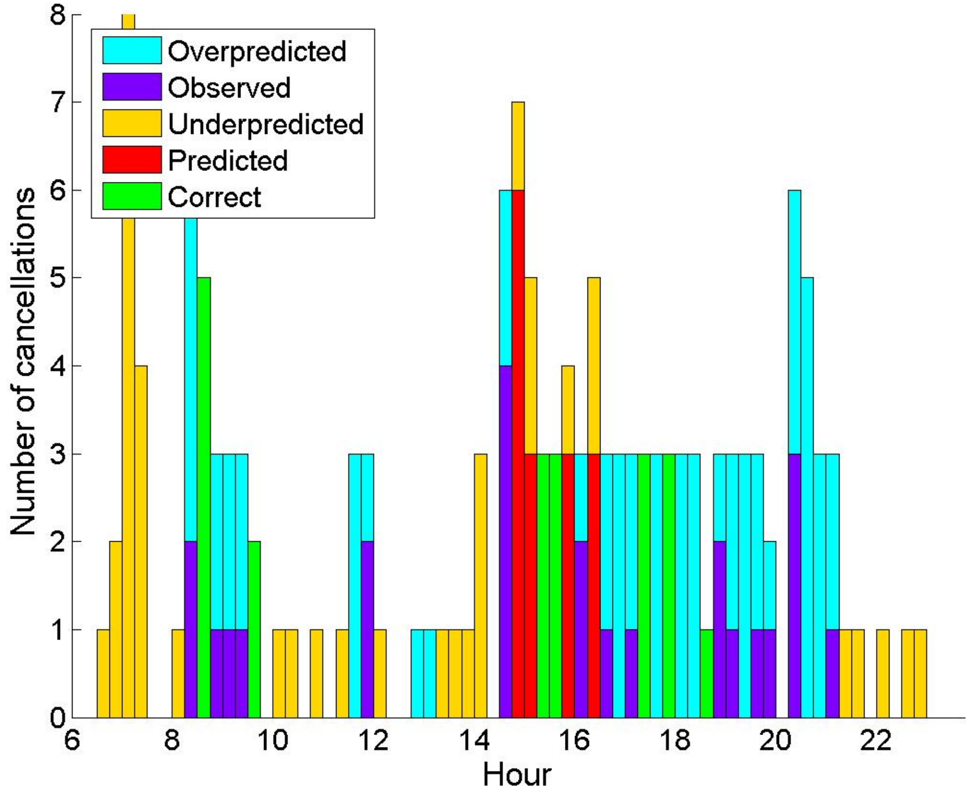

ATL2004, U=6, [9 18 36]
Sept 7, 2004
Predicted = 106
Observed = 97
Shift = -0.0938
Residuals = 216.1563

# III.3. Hourly profile plots



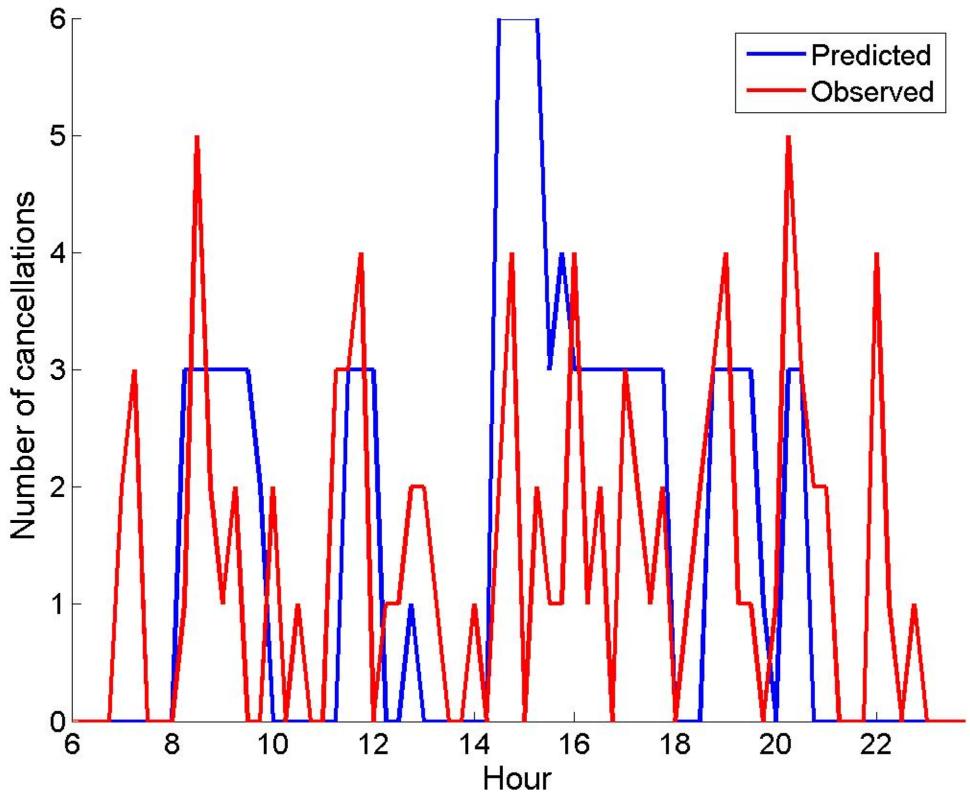

ATL2004, U=2, [15 30 60]
Sept 7, 2004
Predicted = 118
Observed = 97
Shift = -0.2188
Residuals = 272.4063

# III.3. Hourly profile plots



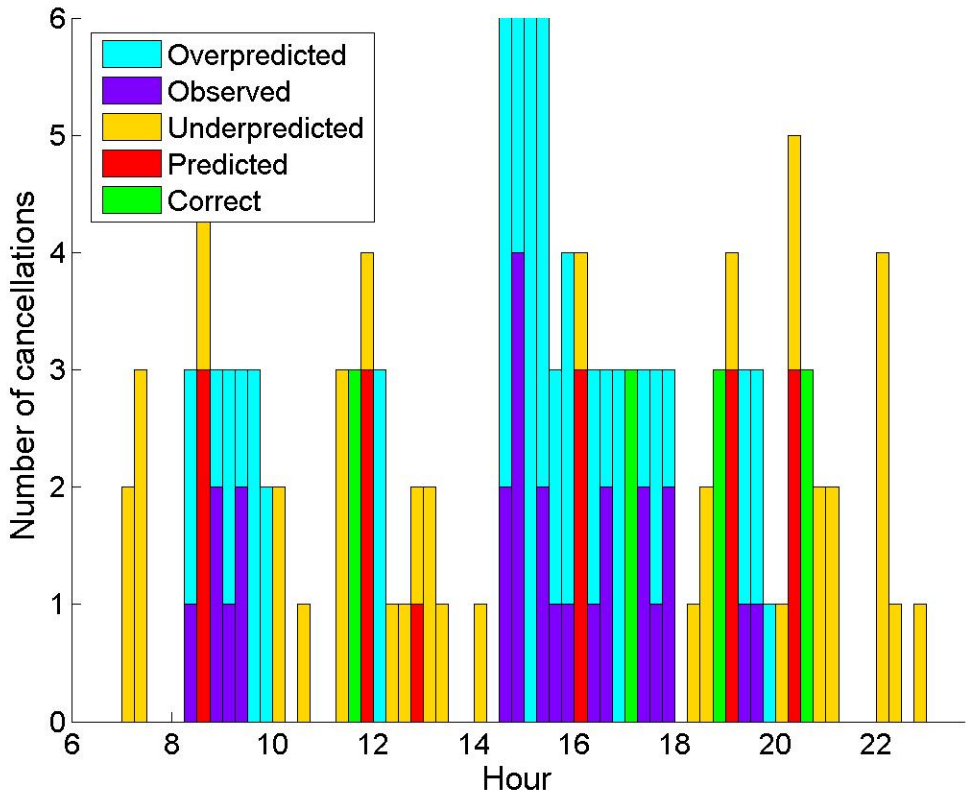

ATL2004, U=2, [15 30 60]
Sept 7, 2004
Predicted = 118
Observed = 97
Shift = -0.2188
Residuals = 272.4063

# III.3. Hourly profile plots



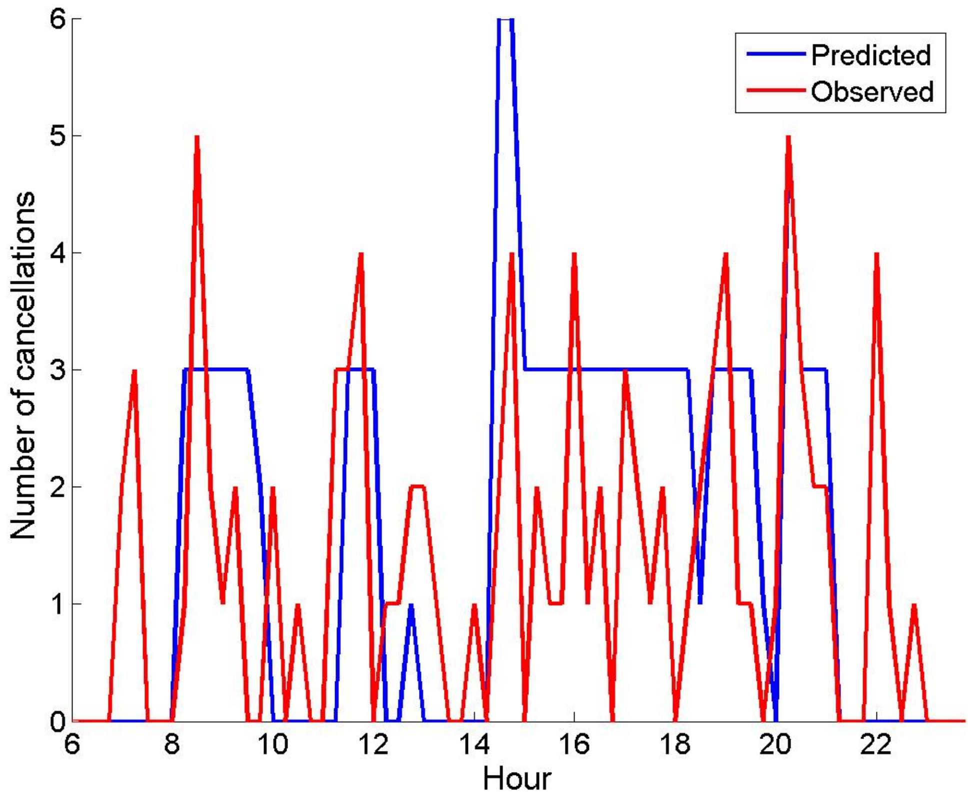

ATL2004, U=6, [9 18 36]
Sept 15, 2004
Predicted = 106
Observed = 93
Shift = -0.13542
Residuals = 227.2396

# III.3. Hourly profile plots



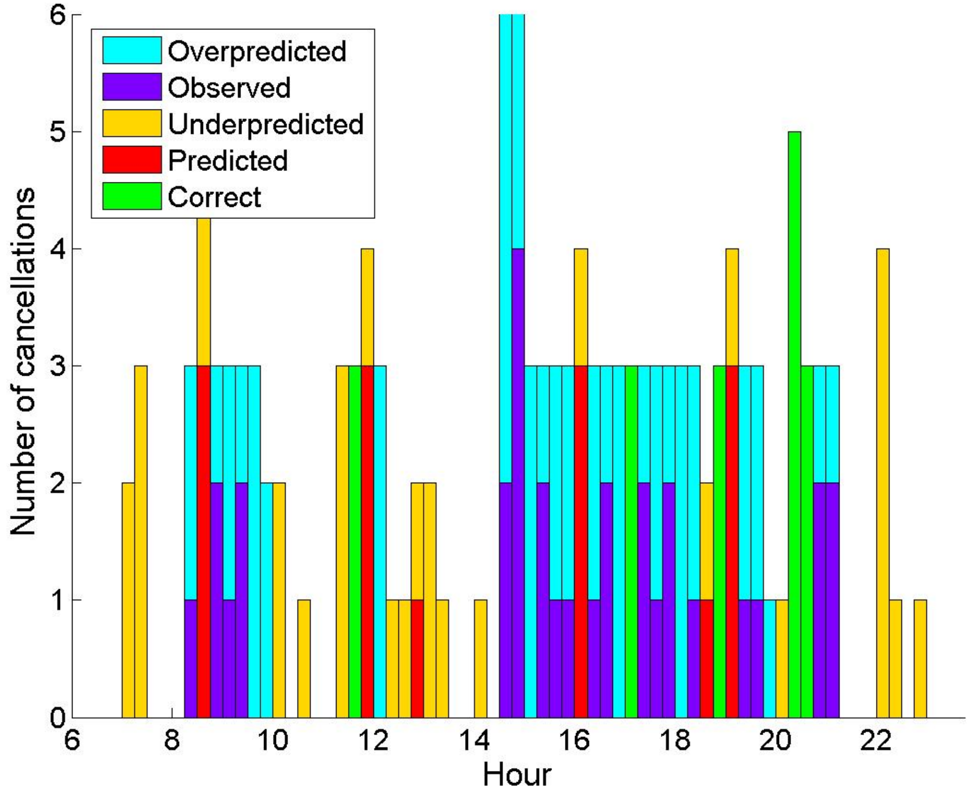

ATL2004, U=6, [9 18 36]
Sept 15, 2004
Predicted = 106
Observed = 93
Shift = -0.13542
Residuals = 227.2396

# III.3. Hourly profile plots




ATL2004, U=2, [15 30 60]
Sept 15, 2004
Predicted = 114
Observed = 93
Shift = -0.2188
Residuals = 176.4063

# III.3. Hourly profile plots




ATL2004, U=2, [15 30 60]
Sept 15, 2004
Predicted = 114
Observed = 93
Shift = -0.2188
Residuals = 176.4063

# IV. Conclusions

- Simple and expedient models
- Useful for iterative strategic planning exercises with multiple airlines:
  - Low levels of airline-specific competitive and/or proprietary information
  - Fast run times (on the order of seconds) to facilitate multiple scenarios and quick response
- Useful for setting preliminary values of parameters for new resource allocation regimes without a strong economic history
- The best predictions of delays and cancellations with minimal inputs that we are aware of