# Reservations, Forecasting, Yield Management and Railroads

**Michael F. Gorman**
**University of Dayton**
**INFORMS**
**November, 2005**

# Presentation Outline

- Introductory thoughts and comments – yield management and railroads
- Demand Forecast Project overview and considerations
- Statistical Forecast Methodologies and Assumptions
- Forecast results: Summary and Detailed
- Conclusions and Recommendations

# Yield Management and Railroads

- What is yield management?
  - The process of allocating the right type of capacity to the right customer at the right price in order to maximize revenue or 'yield' (Kimes 1989**)**

- Do railroads "do" yield management?
- Examples of yield management-related activities in rail:
  - Market segmentation:
    - Differential pricing by commodity, customer, and service level
    - Train type and block space allocation, track allocation, locomotive allocation…
  - Car allocation priorities (and BNSF's car allocation pricing: COTs, LOGs)
  - Price bundling and unbundling
  - But, in general… no.

# When does YM work?

- Yield Management suits industries where:
  - Demand is unstable or highly time-varied
  - The service/product is perishable
  - Market can be segmented

Railroads have these

  - Capacity is fixed and well-known
  - Inventory of capacity is homogenous

Railroads don’t have these

Do Railroads have these?

  - Selling of service to their customers occurs well in advance of consumption
    - In effect, price is determined well in advance of consumption, but volume is not determined until the point of consumption
    - (no reservations)
  - Low marginal costs of service
    - Railroads have the blessing, and the curse, of flexible capacity, which creates significant marginal costs of service (extra sections, extra locomotives, etc.)
    - And makes pricing to maximize yield a challenge (due to a created cost focus)

# Yield Management Ingredients

- **Demand forecast** –
  - What is coming, at a detailed level, that is to be priced
- **Capacity identification** –
  - What capacity is available, at a detailed level, that is to be sold
- **Demand management** –
  - What component of demand that can be shifted to better match capacity; how can customers be differentiated
- **Pricing –**
  - What price may drive the desired customer behavior (demand shifting);
  - What price roughly equates known revenue with an order now against expected foregone revenue of future reservations
- **Overbooking/Order acceptance** –
  - When do you accept a customer car order (reservation), not knowing what orders are yet to come, or even which train capacities the customer may use?

## Railroads Fall Short on many of these critical ingredients:

- **Demand forecast** –
  Very little advanced knowledge of customer behavior
  - Generally, no reservations to speak of
    (BNSF has tried with "RSVP", COTS, LOGS; CN may be an exception)
  - <u>Many</u> different products (origin, destination, train types, service levels, equipment types, time of day, day of week) to predict statistically
- **Capacity identification** –
  Rail capacity is complex, not well-known, and not fixed
  - 5 different physical capacity constraints (car, yard, line, crew, loco);
  - Train is the sixth capacity type that uses all of these physical capacities and ties them all together
  - But, the train capacity changes daily with second sections, annulments and consolidations --- effectively negating the fixed capacity assumptions that are bedrock in most yield management applications.
  - Contrary to airlines, hotels, etc., rail capacity is much more flexible than pricing, so supply adjusts to demand. Despite a growing "run the plan" philosophy, we still reroute 10% or so (or more?) of all traffic... well beyond airlines or hotels.

# Railroads fall short on these ingredients – cont.

- **Demand Management** –
  Shipping patterns are not easily swayed by price
  - Not much of demand is easily ported between days of week or equipment type. Production and delivery schedules aren't flexible enough to be swayed by production costs.
  - Demand management principally takes the form of pushing off accepted loads to non-peak days, which often results in service failures.
- **Pricing –**
  Systems and commercial relationships limit ability of railroads to use price
  - Rail pricing isn't nimble enough to take advantage of perceived mismatch of demand and supply conditions.
  - Rail pricing systems are not made for quickly changing prices, and the majority of historical customer relationships and longer-term contracts are not geared towards fluctuating prices.
- **Load Acceptance** –
  The load acceptance function is not executed in rail
  - Typically, railroads accept all orders, then fail to provide a car or on time service
  - Akin to overbooking;  But the "cost" of failure is not as well-known as in airlines, where an explicit cost is paid to "volunteers"

# MFG Consulting – BNSF
# Network Predictability Project

- **Objective:**

  **To evaluate predictability of network at various levels of aggregation to support operational decision-making.**
  **(Conversely, the "unpredictability" the limits operational decision-making.)**
  - **Predictability:** the "signal" or explainable portion of shipper behavior relative to the "noise" or unexplained fluctuations
  - **Levels of Aggregation:** level of detail of forecast; hub complex, hub, equipment type, equipment length, shipper, time interval, etc.
  - **Operational decision-making:** asset deployment to better match shipper demand patterns

- **Opportunity:**

  **Better visibility into "what is coming" will allow better asset deployment, improving both asset utilization and availability**

- **Project Scope:**
  - Three major intermodal ramps
  - Historical Data: Jan 1, 2003 - March 31, 2005 In-gate data (821 days)
  - Trailer and Container Loads only (not empties)
  - Forecast at a level which may be used to improve flat car and double stack car positioning, lift forecasting, train planning (not aggregate financial forecasting)

# Operational Forecast Requirements

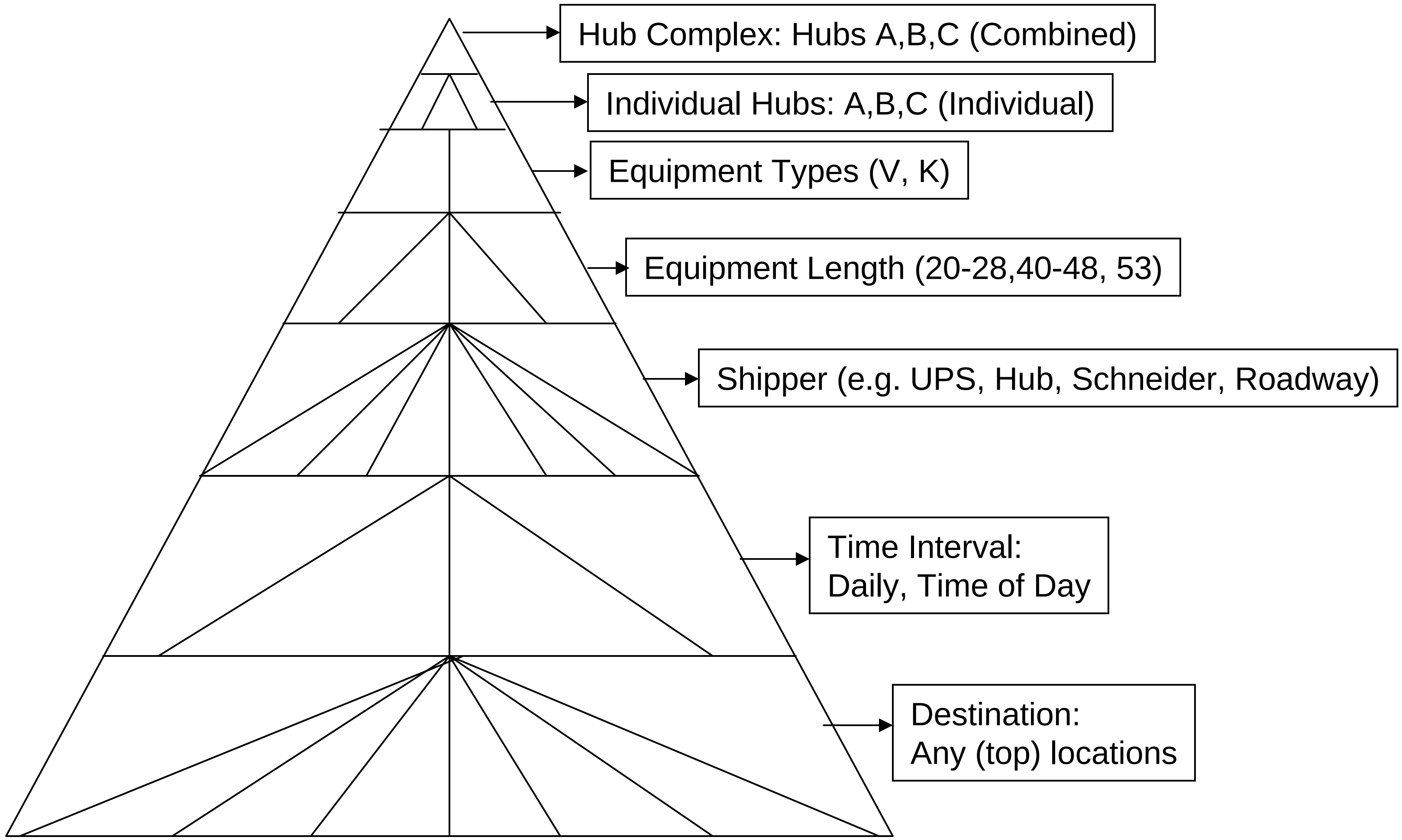
- **We evaluate the feasibility of forecasting shipper patterns as an alternative to shipper reservations for each shipment**
  - **"Reservation" level of detail:**
    - **Shipper, Date, Origin, Destination, Equipment, Length (and Service Level) (not time of day, but, would be nice)**

- **However, the desired level of detail for a forecast depends on its intended operational use**
  - **Very different from financial forecasting**

- **We can examine in gates at various levels of detail:**
  - **Ramp Management -- Total in-gates**
  - **Train Planning – Total vans and containers by destination**
  - **Car management – Total vans and containers (and lengths)**

- **If we can forecast accurately at these levels of detail, the value derived from customer reservations is reduced to some degree**
  - **If shipper-level forecast helps improve forecast accuracy, then forecasts can be made a more detailed level, then aggregated up**

## Pictorial of Forecasting Aggregation
## How far down the "pyramid" can we go and maintain accuracy?




We evaluate forecast accuracy at various levels of operational detail.

## Forecasting Parameters, Measures, Methods and Assumptions Summary

- **Forecasting Parameters:**
  - One day ahead forecast horizon
  - Static forecast
  - In-sample forecast evaluation
- **Measure of forecast quality:**
  - Mean Absolute Error (MAE)
  - Mean Absolute Percent Error (MAPE)
- **Methods Explored:**
  - Exponential Smoothing
  - Holt-Winters (with daily/monthly smoothing)
  - Stepwise autoregressive (by Day of Week)

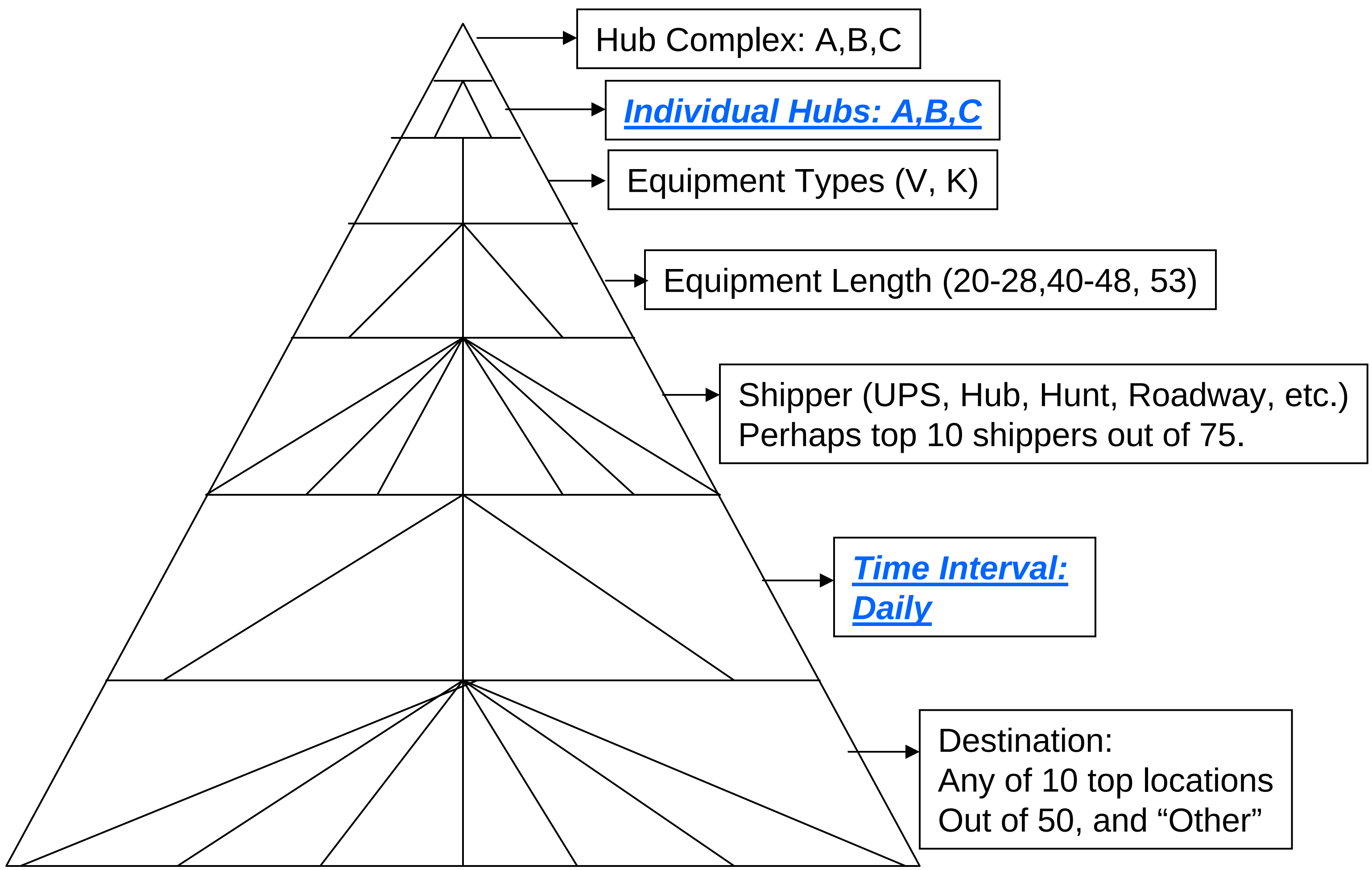
# Forecast Evaluation Overview

- **<u>Summary Forecast:</u>**
  - **Holt/Winters, Stepwise Autoregressive and Exponential Smoothing are evaluated**
  - **At the Hub level of aggregation**

- **<u>Dive Down detailed forecast</u>:**
- **One day ahead forecasts – Holt-Winters Method Used**
  - **Holt/Winters – Seasonal Multiplicative adjustment**

- **Various forecast levels of detail (Dive Downs)**
  - **Shipper Focused**
    - **Dive Down 1: Hub/Shipper**
    - **Dive Down 2: Hub/Shipper/Equipment Type**
    - **Dive Down 3: Hub/Shipper/Equipment Type/Equipment Length**
    - **Dive Down 4: Hub/Shipper/Equipment Type/Equipment Length Destination (reservation level of detail, excluding service level)**
  - **Operations Focus (Exclude shipper disaggregation)**
    - **Dive Down 1: Hub/Equipment Type**
    - **Dive Down 2: Hub/Destination**
    - **Dive Down 3: Hub/Destination/Equipment type**

## Pictorial of Forecasting Aggregation
## Summary Forecast: Hub Level, Day of Week



# Stepwise Autoregressive vs. Holt-Winters

**Winters with Weekly and Monthly Seas**

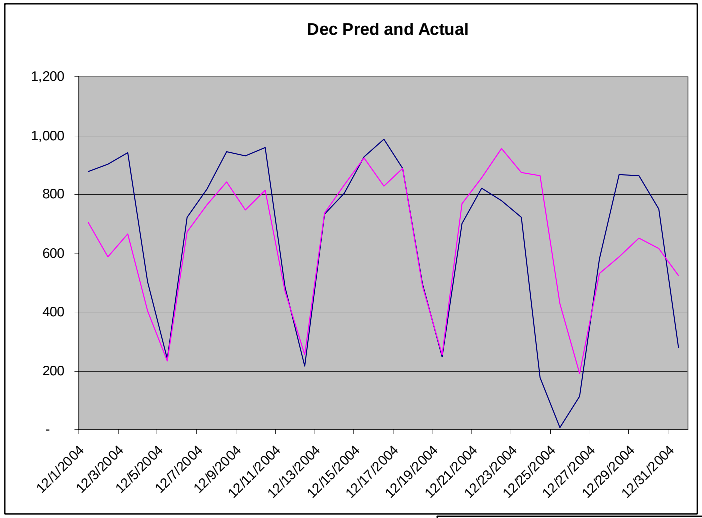
**MAE**

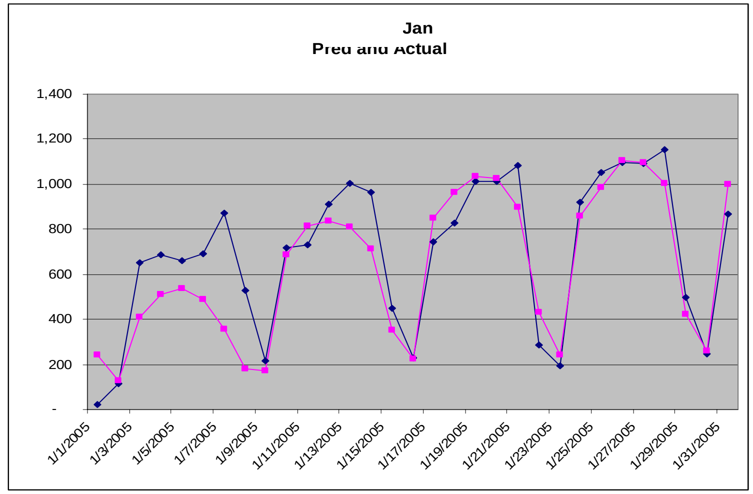
| | Winters-Seas | Winters | Stepar |
|---|---|---|---|
| A | 69.76 | 86.69 | 62.15 |
| B | 46.20 | 53.06 | 40.24 |
| C | 77.30 | 91.26 | 67.63 |

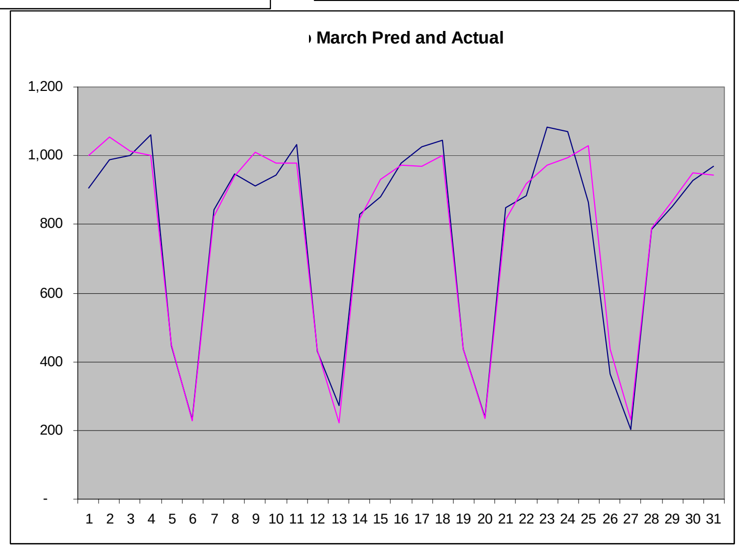
Stepwise Autoregressive has lowest MAE,
Holt-Winters with (daily) Seasonality performs comparably.

- **Holt-Winters will be used for "dive-down" forecasts –**
- **Single equation with daily adjustment versus seven different day-of-week equations**
  - **More economical for automation**
  - **Leverages Uses "same time last year" information – more intuitive**
  - **Leverages More recent information - One day ahead forecast (uses last three days' information to better adjust to within-month variations)**

# Holt-Winters with Monthly/Daily Seasonality



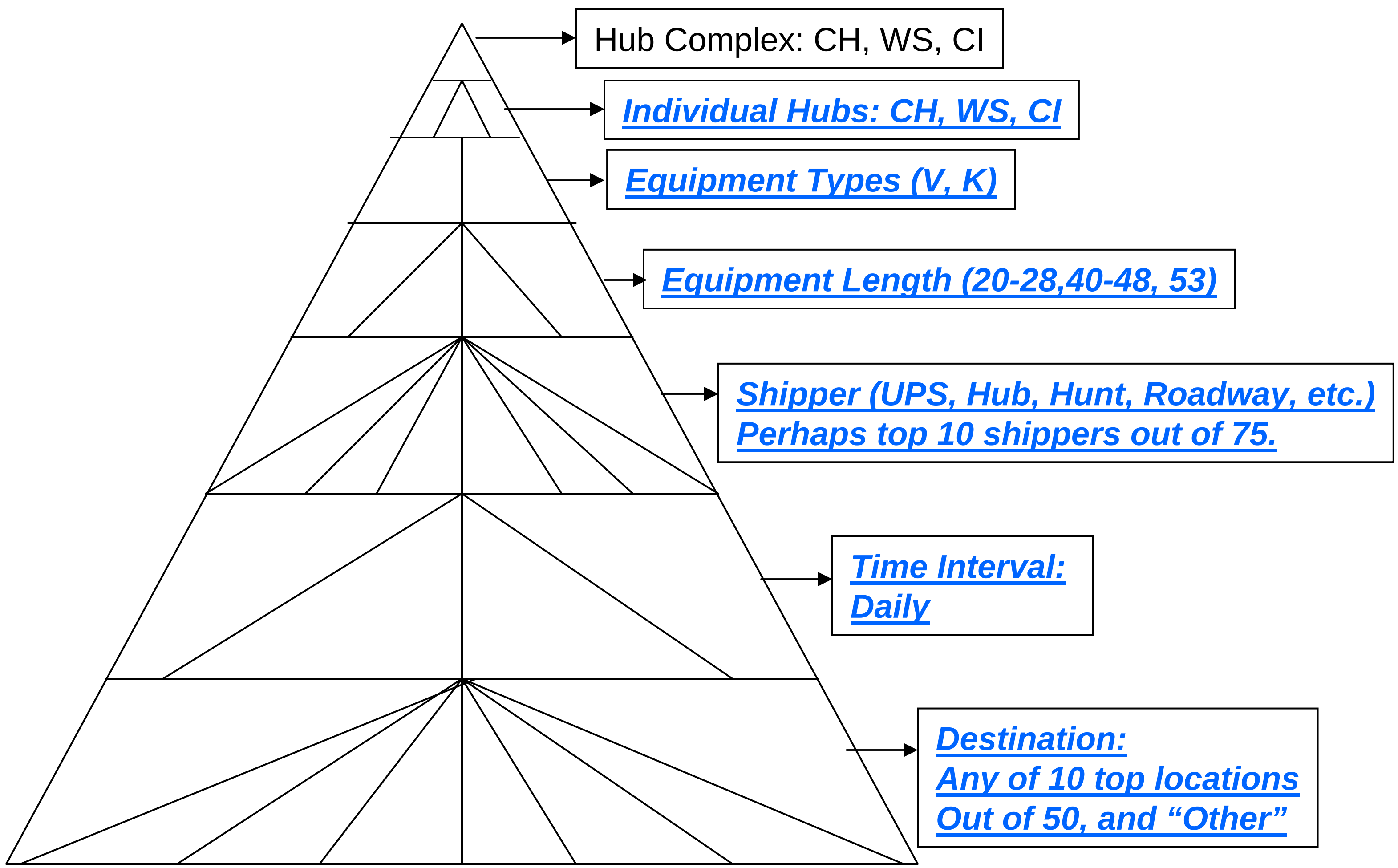

Holt-Winters generally tracks well when seasonal adjustments are included.

# Confidence Interval on Forecast: Hub Level

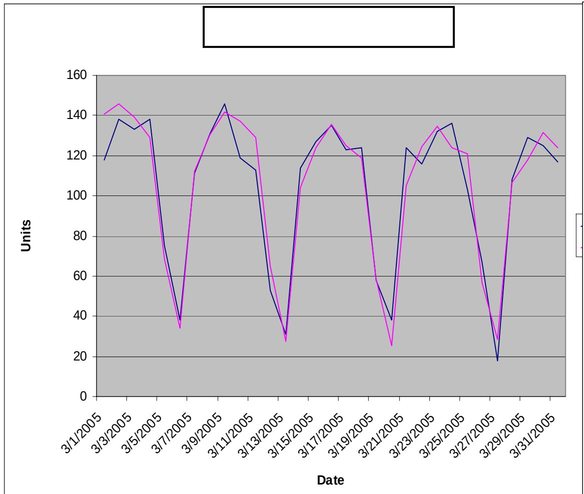
| Hub-Level Forecast<br>Method: Additive Holt/Winters | | |
|---|---|---|
| | | |
| Origin Ramp | Lower 95% Confidence Interval Boundary | Upper 95% Confidence Interval Boundary |
| A | 704 | 1,177 |
| B | 588 | 890 |
| C | 1,107 | 1,607 |

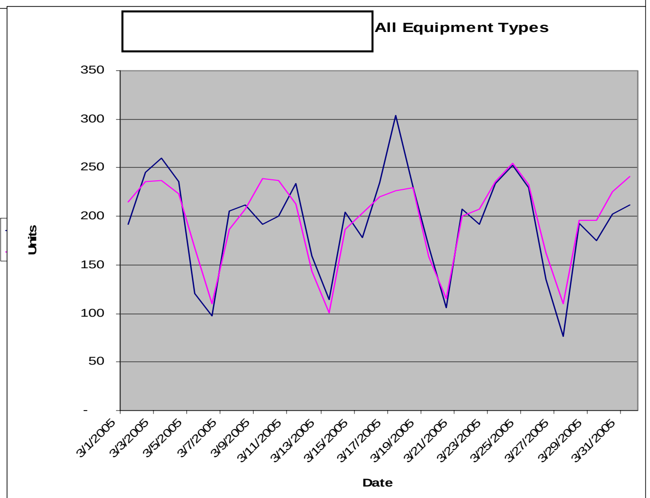
Even at the highest (Hub) level of aggregation,
the 95% confidence interval around the forecasted value is 300-500 units.

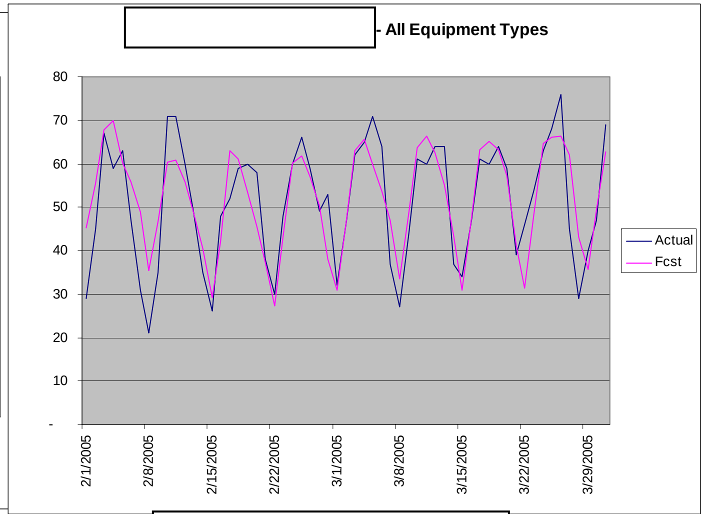
## Shipper Dive Down 4: (Reservation Level of Detail)
## Hub, Shipper, Equipment Type, Length and Destination

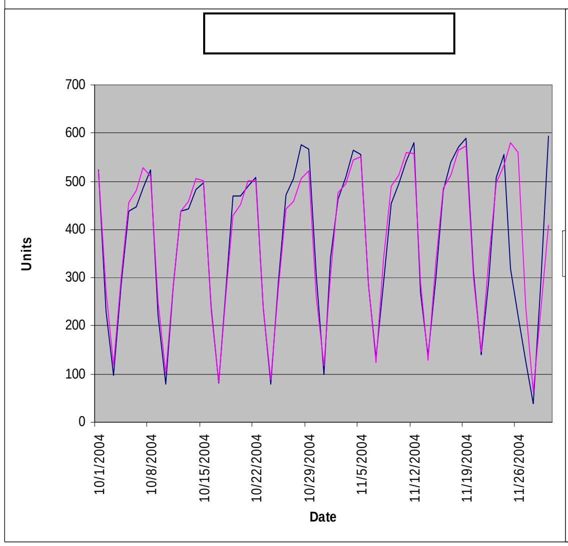


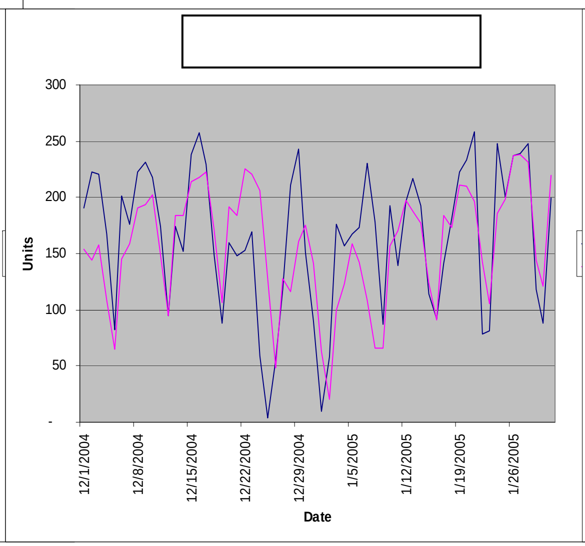
# Illustrative Graphs
# Shipper-Hub Level Forecast Accuracy (all Equipment)

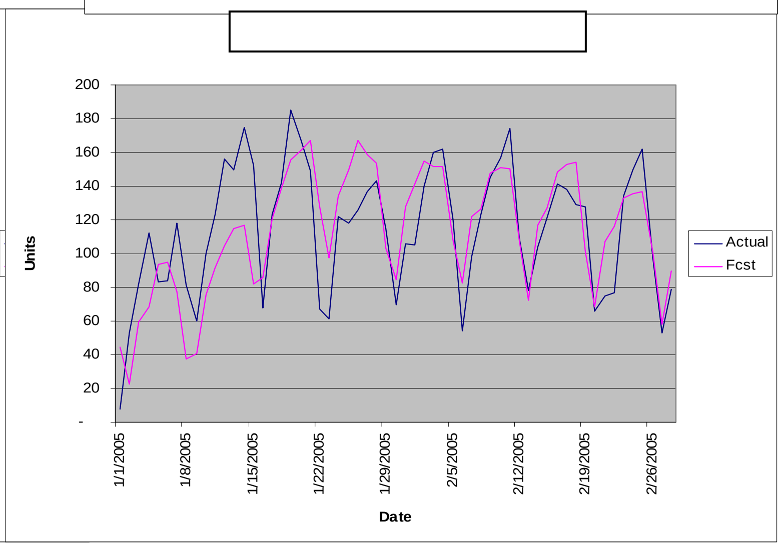


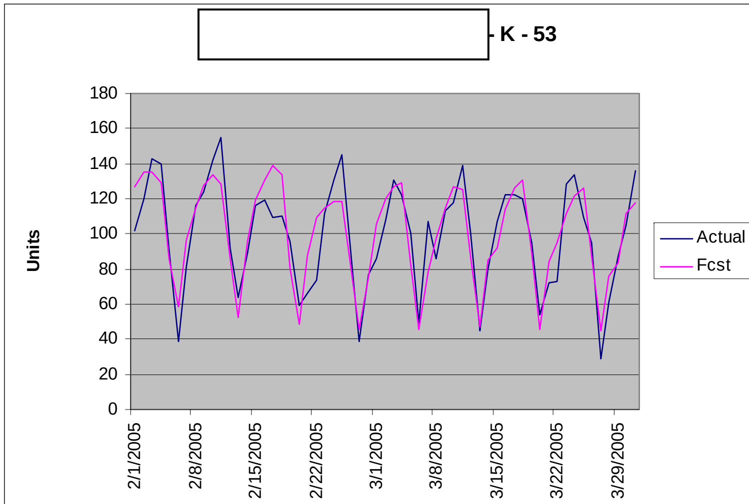



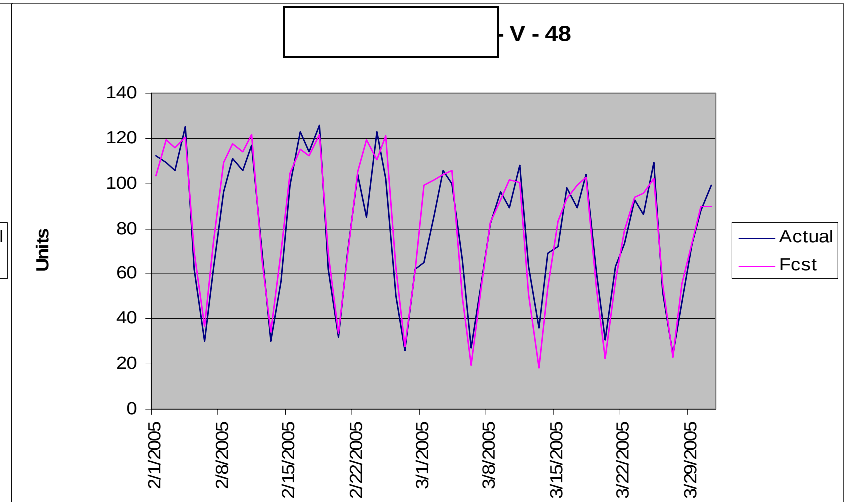



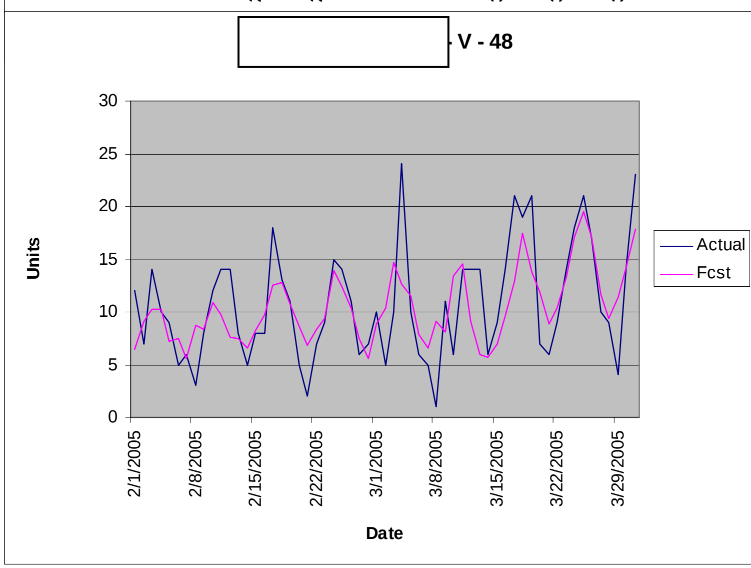



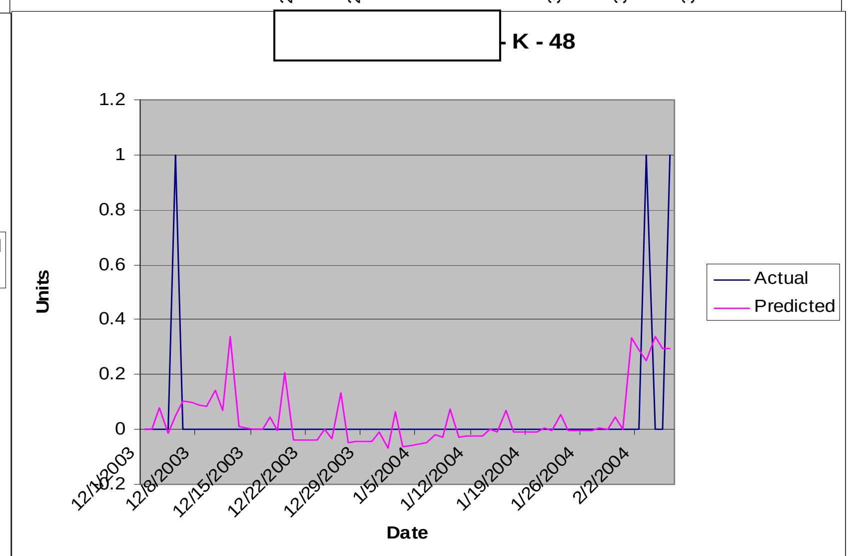



# Illustrative Graphs
# Shipper-Hub-Equipment-Length Level of Detail



# Comparison of Forecast Accuracy: Shipper-Specific Illustration Dive Down Levels 1-4

**Customer** **Forecast Dive Down:**

**Level 1**

| Ramp | CONSTANT | MAE | MAPE | N |
|---|---|---|---|---|
| | 4.59 | 2.15 | 16.83 | |
| | 29.99 | 6.30 | 34.14 | 820.00 |
| | 85.17 | 13.46 | 28.42 | 821.00 |

All units forecast.

**Level 2**

| Ramp | Equipment | CONSTANT | MAE | MAPE | N |
|---|---|---|---|---|---|
| | K | 3.60 | 5.27 | 26.99 | 800.00 |
| | K | 2.31 | 2.40 | 25.96 | 373.00 |
| | V | 26.62 | 6.23 | 34.24 | 820.00 |
| | V | 85.11 | 13.45 | 28.42 | 821.00 |

Four Groups not forecast

**Level 3**

| Ramp | Equipment | Length | CONSTANT | MAE | MAPE | N |
|---|---|---|---|---|---|---|
| | K | 53 | 3.51 | 3.26 | 10.52 | |
| | K | 53 | 2.50 | 1.89 | 52.02 | 176 |
| | V | 48 | 0.86 | 0.96 | 2.77 | 778 |
| | | 53 | 26.44 | 6.21 | 34.27 | 820 |
| | V | 48 | (0.20) | 1.62 | 10.91 | 818 |
| | | 53 | 84.55 | 13.41 | 28.42 | 821 |

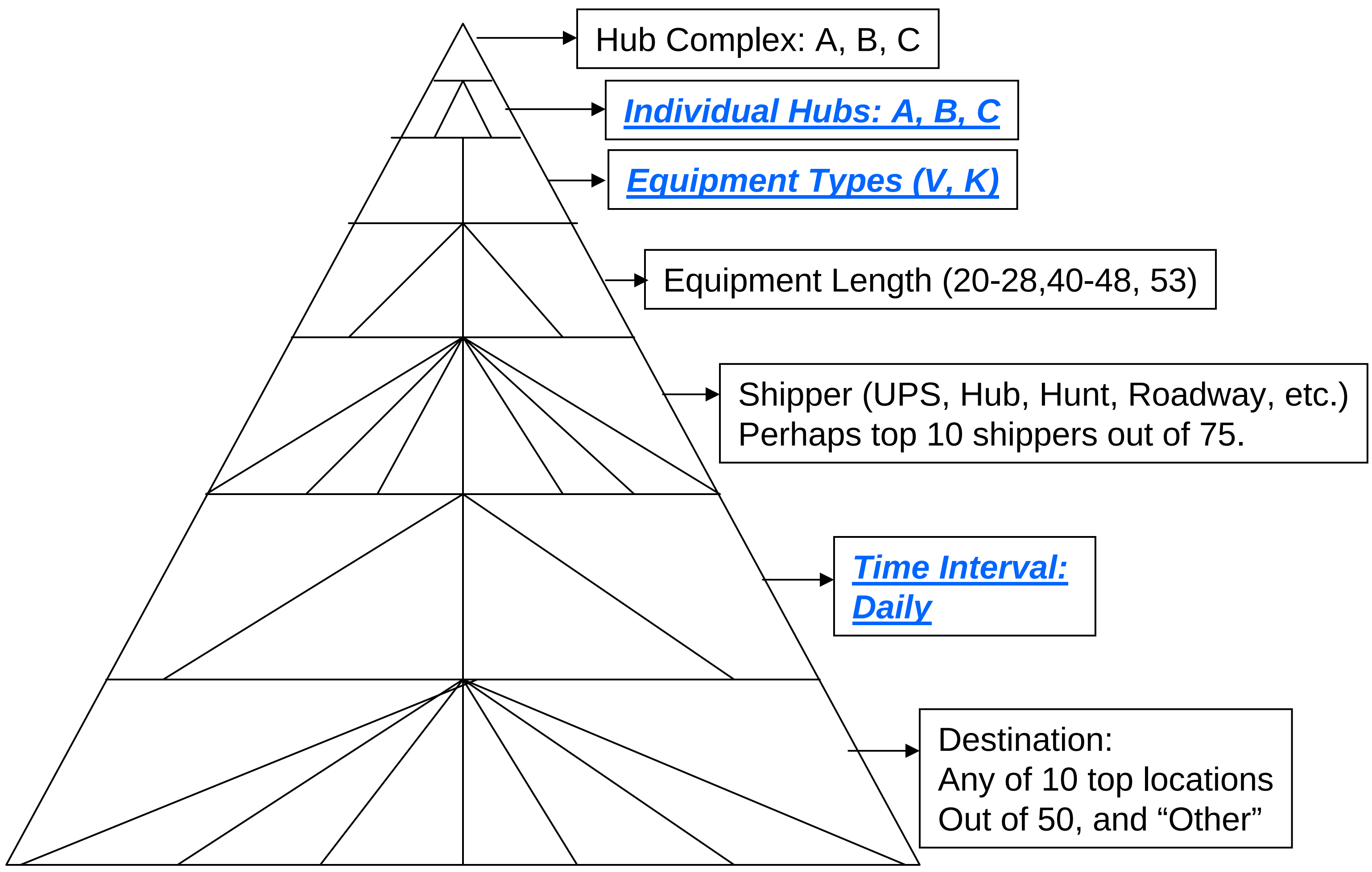
Seven by groups only one observation;
15 groups
Not forecast (22 total)

**Level 4**

| Ramp | Dest | Equip | Length | CONSTANT | MAE | MAPE | N |
|---|---|---|---|---|---|---|---|
| | | K | 53 | 3.33 | 13.78 | 19.58 | |
| | | V | 53 | 10.76 | 3.66 | 45.62 | 820 |
| | | V | 53 | 2.39 | 1.02 | 22.64 | 817 |
| | | K | 53 | 2.54 | 1.76 | 49.79 | 176 |
| | | V | 53 | 15.30 | 4.14 | 44.65 | 820 |
| | | V | 53 | 1.01 | 1.85 | 56.92 | 396 |
| | | V | 53 | 0.76 | 1.09 | 3.44 | 733 |
| | | V | 53 | 10.23 | 2.99 | 51.69 | 821 |
| | | V | 53 | 3.21 | 1.68 | 66.84 | 819 |
| | | V | 53 | 5.10 | 2.39 | 64.07 | 820 |
| | | V | 53 | 24.61 | 4.20 | 40.81 | 820 |
| | | V | 53 | 5.61 | 2.10 | 73.32 | 820 |
| | | V | 53 | 34.29 | 8.21 | 28.01 | 821 |
| | | V | 53 | 7.30 | 3.97 | 54.14 | 820 |

60 groups one obs;
127 not forecast
(187 total)

# Operational Dive Down 1: Hub and Equipment Type

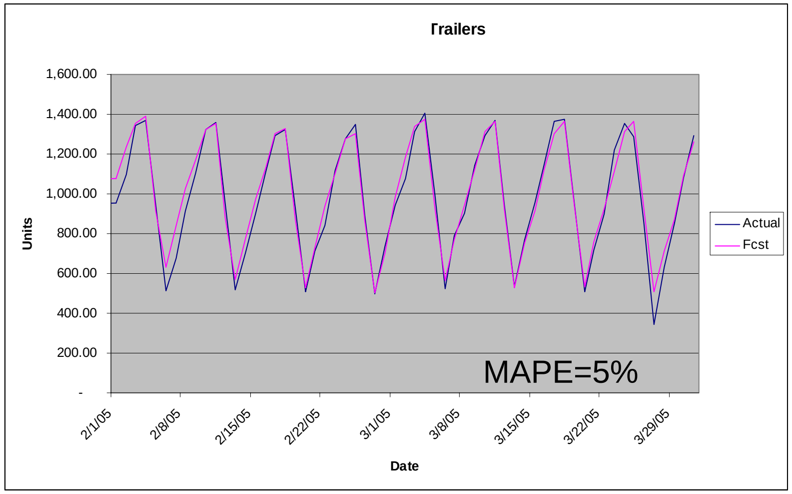


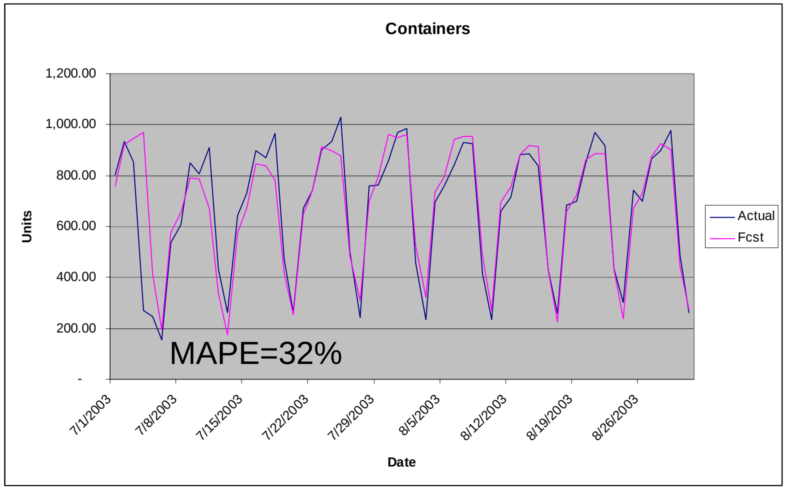
# Operational Dive Down 1: Hub and Equipment Type Forecast Accuracy Summary

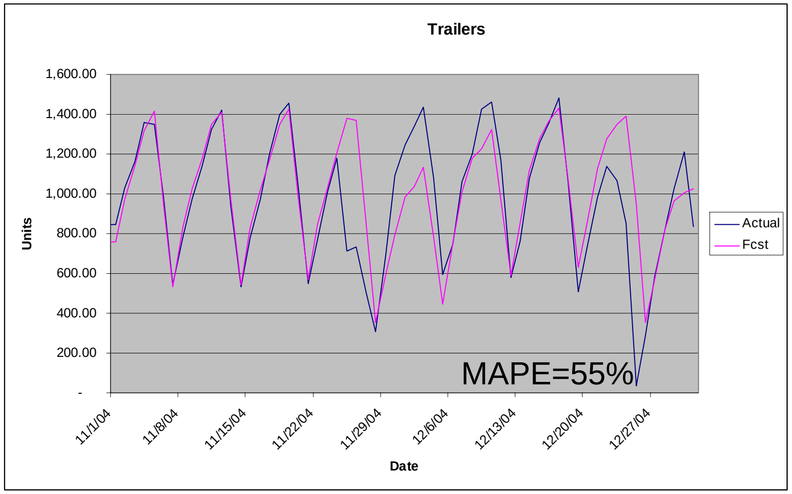
| Sum of Value | | Stat | | | |
|---|---|---|---|---|---|
| Ramp | Equip | CONSTANT | MAE | MAPE | N |
| | K | 710.73 | 69.30 | 42.07 | 821 |
| | V | 0.23 | 0.50 | 13.93 | 761 |
| | K | 259.86 | 28.47 | 32.69 | 821 |
| | V | 293.77 | 24.40 | 20.09 | 821 |
| | K | 63.79 | 11.42 | 28.71 | 821 |
| | V | 914.05 | 72.43 | 21.33 | 821 |

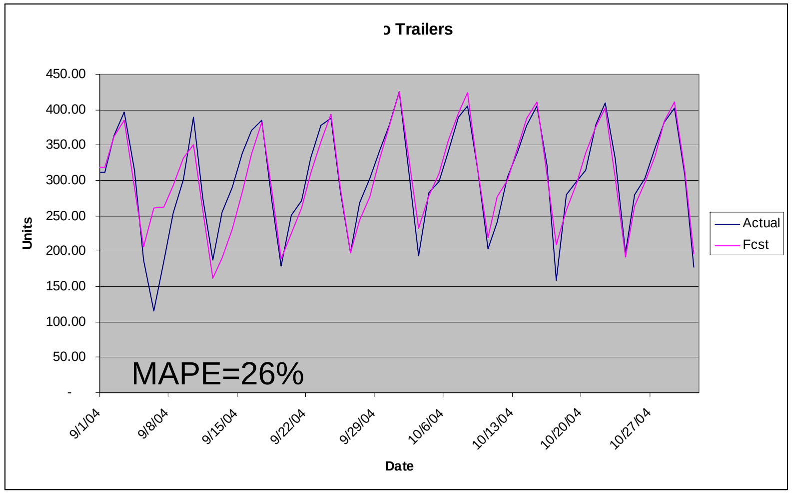
| Sum of Value | | | Stat | | | |
|---|---|---|---|---|---|---|
| Ramp | Equip | Date | L95 | U95 | Fcst | 95% Error Band |
| | K | 1-Apr-05 | 703.90 | 1,177.01 | 940.45 | 25% |
| | V | 23-Feb-05 | (2.31) | 2.46 | 0.07 | 3217% |
| | K | 1-Apr-05 | 251.91 | 425.21 | 338.56 | 26% |
| | V | 1-Apr-05 | 322.60 | 478.37 | 400.49 | 19% |
| | K | 1-Apr-05 | 50.92 | 111.36 | 81.14 | 37% |
| | V | 1-Apr-05 | 1,036.75 | 1,515.57 | 1,276.16 | 19% |

# Operational Dive Down 1: Hub and Equipment Type Illustrative Examples



# Operational Dive Down Forecast Accuracy Comparison
# Hub, Destination, Equipment Type

| Sum of Value | | Stat | | | |
|---|---|---|---|---|---|
| Ramp | Equip | CONSTANT | MAE | MAPE | N |
| | K | 710.73 | 69.30 | 42.07 | 821 |
| | V | 0.23 | 0.50 | 13.93 | 761 |
| | K | 259.86 | 28.47 | 32.69 | 821 |
| | V | 293.77 | 24.40 | 20.09 | 821 |
| | K | 63.79 | 11.42 | 28.71 | 821 |
| | V | 914.05 | 72.43 | 21.33 | 821 |

As operational forecast detail increases, forecast accuracy decreases at a slower rate than customer-specific forecasts. further, more “full” data sets are available for forecasting.

| Sum of N | Ramp | Stat | | |
|---|---|---|---|---|
| Dest | CONSTANT | MAE | MAPE | N |
| | 209.35 | 18.77 | 19.02 | 821 |
| | 202.39 | 21.54 | 18.49 | 821 |
| | 89.92 | 11.30 | 18.49 | 821 |
| | 89.17 | 8.73 | 22.56 | 821 |
| | 64.79 | 9.17 | 37.51 | 821 |
| | 58.70 | 6.74 | 16.67 | 821 |
| | 58.45 | 5.97 | 15.33 | 820 |
| | 32.39 | 4.04 | 27.79 | 820 |
| | 27.26 | 4.42 | 23.99 | 821 |
| | 26.77 | 4.07 | 23.42 | 820 |
| | 26.55 | 4.56 | 23.77 | 820 |
| | 26.52 | 4.24 | 23.29 | 820 |
| | 25.35 | 3.77 | 43.81 | 817 |
| | 17.89 | 2.96 | 27.99 | 820 |
| | 13.65 | 2.20 | 36.07 | 820 |
| | 10.23 | 3.19 | 55.36 | 820 |
| | 5.78 | 1.13 | 19.12 | 488 |

| Sum of Value | | Ramp | Stat | | |
|---|---|---|---|---|---|
| Dest | Equip | CONSTANT | MAE | MAPE | N |
| | V | 190.98 | 17.11 | 20.82 | 821.00 |
| | K | 18.37 | 5.75 | 43.94 | 821.00 |
| | V | 186.25 | 19.54 | 18.43 | 821.00 |
| | K | 16.15 | 6.04 | 70.09 | 820.00 |
| | V | 82.88 | 10.96 | 20.82 | 821.00 |
| | K | 7.05 | 2.40 | 67.69 | 820.00 |
| | V | 63.42 | 8.91 | 36.58 | 821.00 |
| | K | 1.42 | 1.50 | 39.33 | 805.00 |
| | V | 11.47 | 3.83 | 27.43 | 821.00 |
| | K | 15.80 | 1.85 | 41.08 | 821.00 |
| | V | 55.72 | 6.51 | 16.83 | 821.00 |
| | K | 2.98 | 1.56 | 49.43 | 815.00 |
| | V | 88.79 | 8.62 | 22.23 | 821.00 |
| | K | 0.61 | 0.94 | 28.08 | 806.00 |
| | V | 0.33 | 0.23 | 10.76 | 239.00 |
| | V | 8.69 | 2.82 | 56.07 | 820.00 |
| | K | 1.54 | 1.17 | 39.86 | 818.00 |

# General Statements
# Forecasting Accuracy and Feasibility

- **Generally, the "finer" or more detailed the shipper-specific forecast, the less accurate it becomes**
  - **Noise begins to dominate trend**
  - **Small deviations and fluctuations in shipper behavior have larger impact on error level**
  - **Many forecast groups have miniscule expected values; making forecasting in any meaningful way a challenge**
  - **However, forecasting at a disaggregate level, and then aggregating, is often more accurate at the aggregate level than an aggregate forecast**
    - **There seem to be strong shipper-specific patterns that are somewhat more easily forecast than the more aggregate level**
    - **These patterns are lost when aggregated before forecasting**

- **As we dive down the pyramid, fewer and fewer observations per forecast group, which hampers forecasting**
  - **Many dates with zero units for small forecast groups**
    - **Fewer days over which to build forecast**
    - **Missing (zero) observations are "gapped"**
    - **MAPE is only calculated for non-zero observations**
  - **In many cases, not enough observations exist to generate any forecast equation**
    - **Hardest groups to forecast, but aren't included in our forecast error estimates**
    - **Many of the forecast groups (even for large customers) are too small to forecast at all – would require "All Other Groups" aggregation to forecast them**

# Observations

- **We might not need to forecast at a "Reservation" level of detail (Customer, origin, destination, equipment type, equipment length) in order to have more proactive operations**
  - **We just aggregate the reservation-level forecast into an operational level, for examples:**
    - **Origin – In gate planning, hub management**
    - **Origin-Destination – train planning**
    - **Origin-Equipment Type – lift planning**
    - **(etc.)**
  - **Forecasting at this level generally has lower forecast error, and more detail does not significantly decrease forecast accuracy or ability**
- **Forecast is a second best to actual reservations**
  - **Accuracy is reasonable, but nothing close to a reservation**
  - **Anecdotal evidence of "regular, predictable" traffic is seemingly confirmed by graphical representation, but clearly contradicted by numerical summaries**
    - **Aggregate volumes are predictable; marginal loads are nearly impossible to forecast with any level of detail**
  - **Knowing what is coming (via forecast) has different management implications than a reservation**
    - **Operations makes adjustments due to <u>forecasted</u> volume**
    - **"Load Acceptance" (no such current function) can make a determination on whether a reservation should be accepted, or at what price and with what service level promise, as opposed to preparing for a forecasted volume**