NRM - Scientific Corner

Revenue optimization workshop

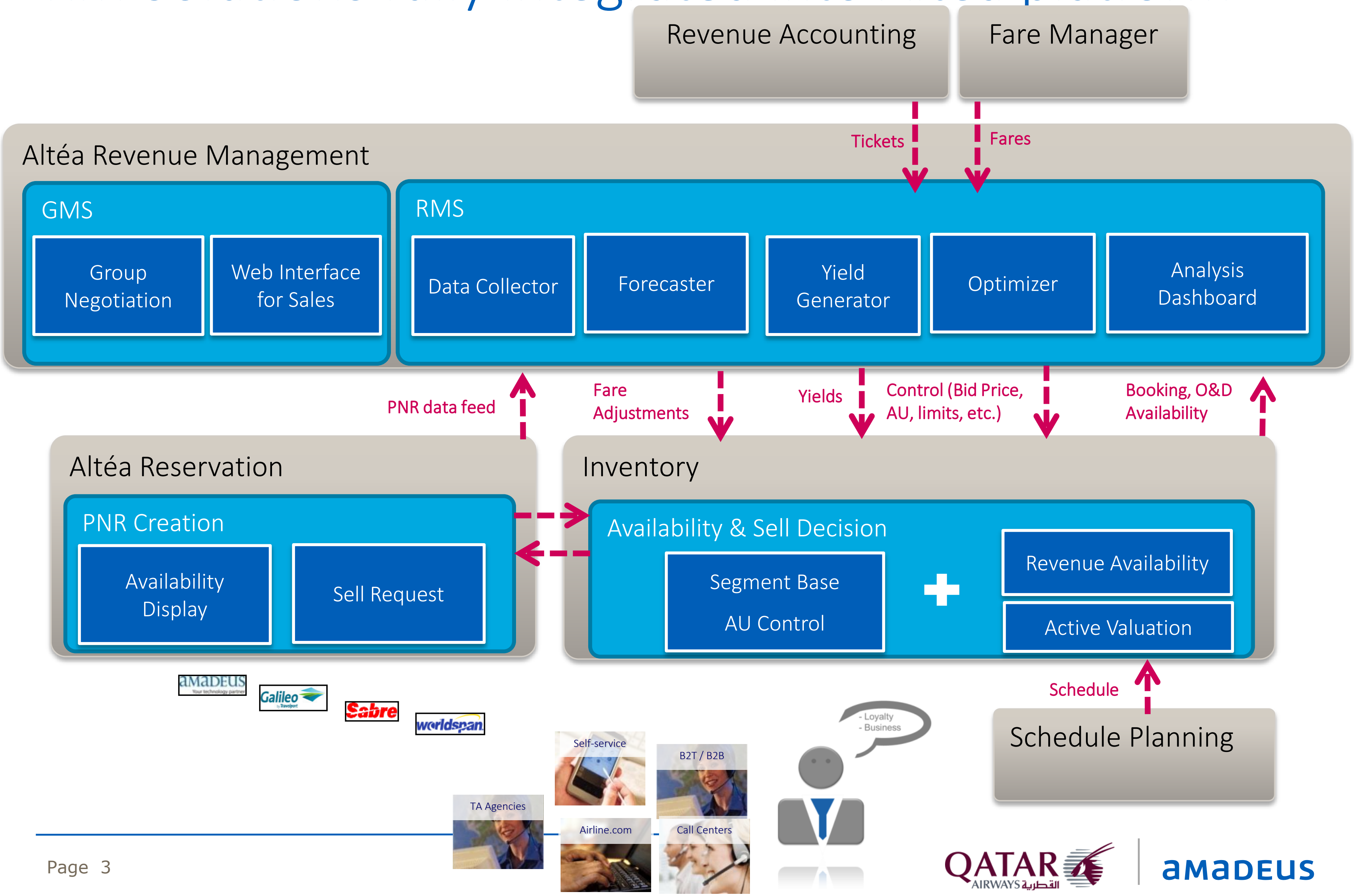
Thomas Fiig
Chief Scientist
December 5th 2016, Doha

Agenda

1. Forecasting overview
2. Virtualization
3. Estimation
4. Back testing - Validation
5. Forecasting - Projection
6. Transformation of dependent demand to independent demand
7. Bayesian forecasting methodologies
8. Conclusion

Confidential
RESTRICTED

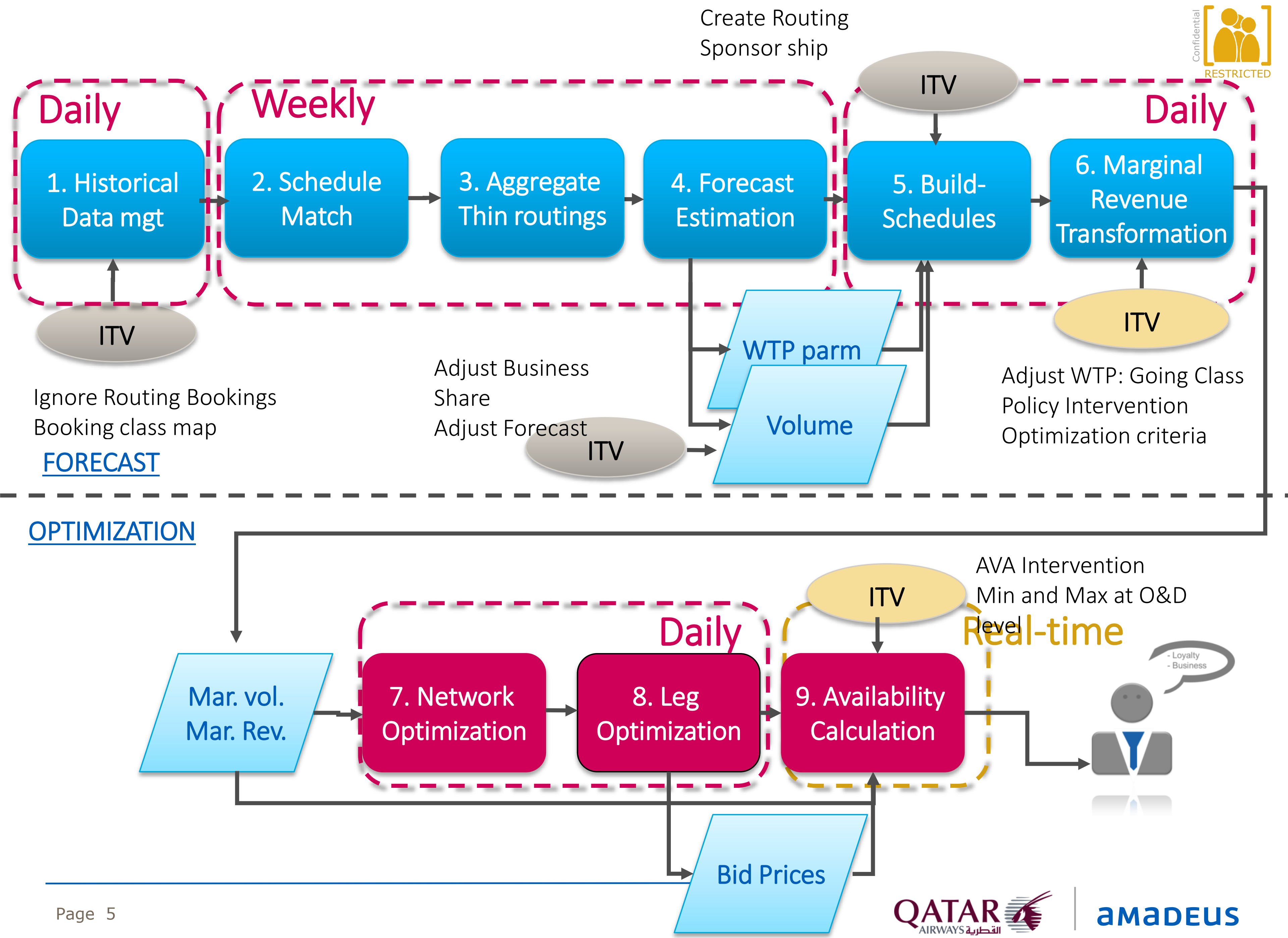
RM Solutions fully integrated into Altéa platform



QATAR AIRWAYS القطرية
amadeus

Strictly CONFIDENTIAL

Forecasting Overview


Confidential
RESTRICTED
Create Routing
Sponsor ship
Daily
Weekly
Daily
ITV
1. Historical Data mgt
2. Schedule Match
3. Aggregate Thin routings
4. Forecast Estimation
5. Build-Schedules
6. Marginal Revenue Transformation
ITV
ITV
WTP parm
Volume
ITV
Ignore Routing Bookings
Booking class map
FORECAST
Adjust Business
Share
Adjust Forecast
Adjust WTP: Going Class
Policy Intervention
Optimization criteria
OPTIMIZATION
AVA Intervention
Min and Max at O&D
level
ITV
Daily
Real-time
Mar. vol.
Mar. Rev.
7. Network Optimization
8. Leg Optimization
9. Availability Calculation
- Loyalty
- Business
Bid Prices
QATAR AIRWAYS القطرية
amadeus

Definition of the traffic flow

Confidential RESTRICTED

Schedule

Legs	OSLCPH CPHSFO
Departure time	13:25 15:35
Point of Sale	DK
Point of Origin	DK
Direction	Outbound
Travel purpose	Leisure
Passenger type	I
Previous airlines	DY
Next airlines	UA
Previous cities	BGO
Next cities	LAX

Routing

Confidential RESTRICTED

Historical Data Management

Input Data

Reservations
- Historic PNR

Inventory (traffic flow level)
- Interventions (Segment limits)
- Current Inventory counters
- Availability per depdate, dbd.
- Schedule information

DCS
- No-show

ETS
- Ancillary revenue

Revenue Accounting
- Coupons (Tickets information)

Fare Quote
- Fares (ATPCO), Tax, Airline specific tax

Airline Data
- Carrying Cost (variable cost)

Market Data
- Competitor prices (QL2/Infare)
- Total Demand
- Schedules

Schedule data are shared instantaneously to all RM modules.

Taxes and fare rules interpretation requires unique knowledge of pricing.

Strictly CONFIDENTIAL

2.

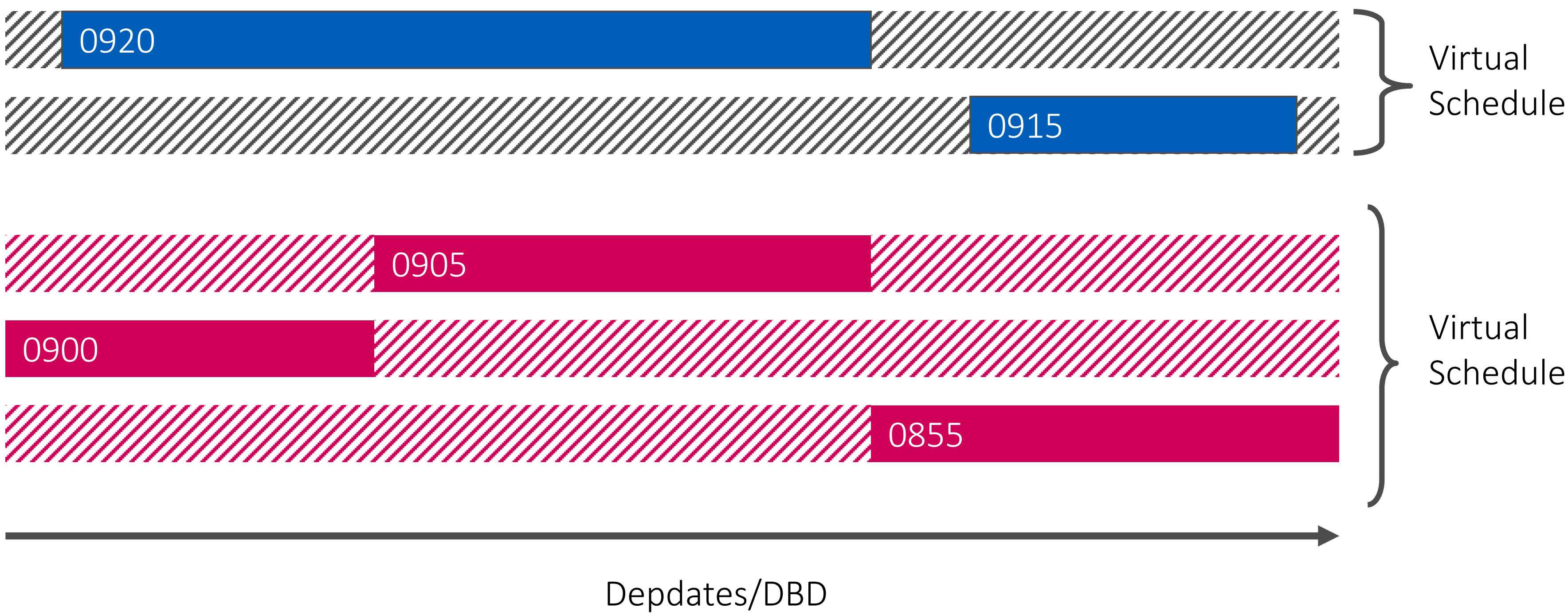
Forecasting - Virtualization

Historical Data Management

Determine Significant Historical Flows

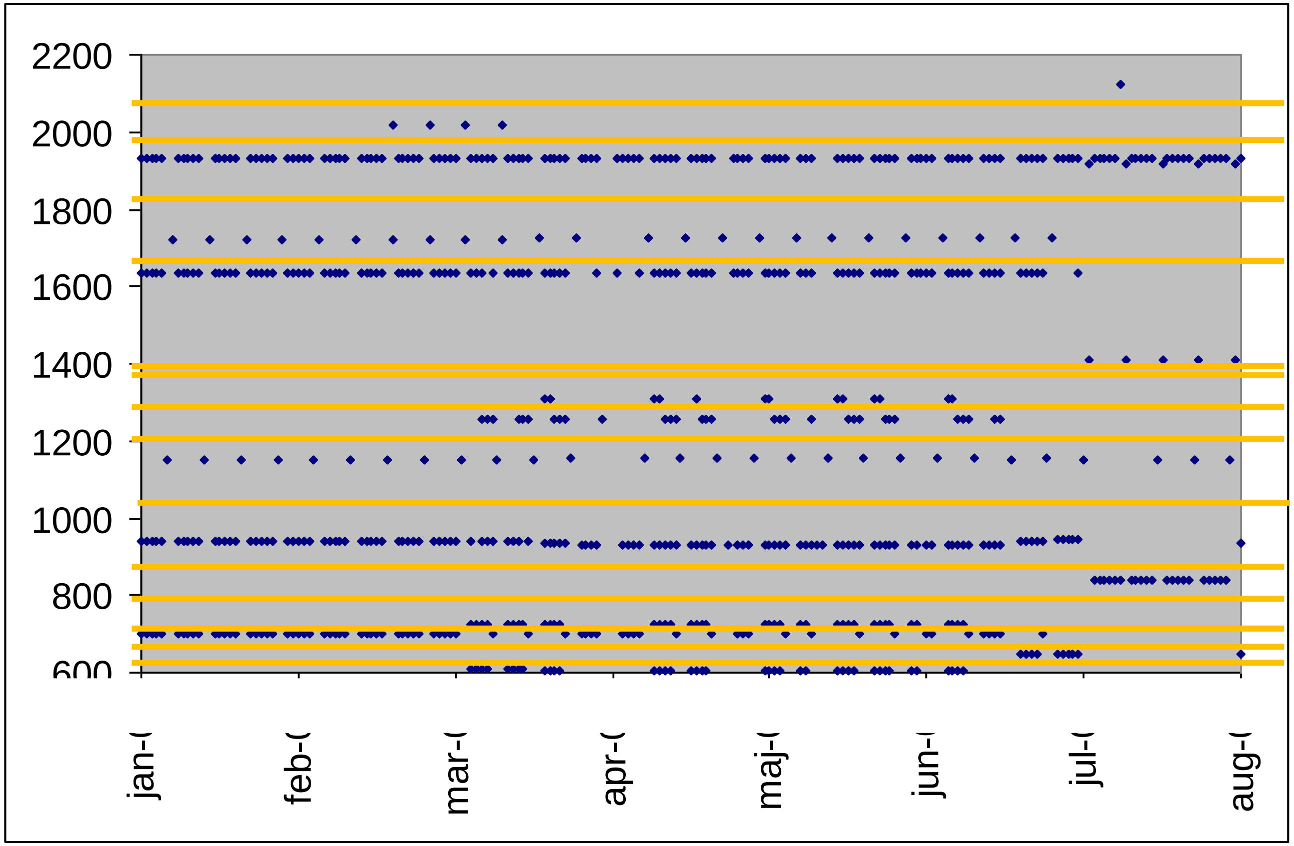
- Virtual schedule is an sequence of flight schedules. (VSID)
- Schedules within a virtual schedule cannot be overlapping.
- Each routing is uniquely mapped to a virtual schedule.

Actual flight schedules



Schedule virtualization

Single leg



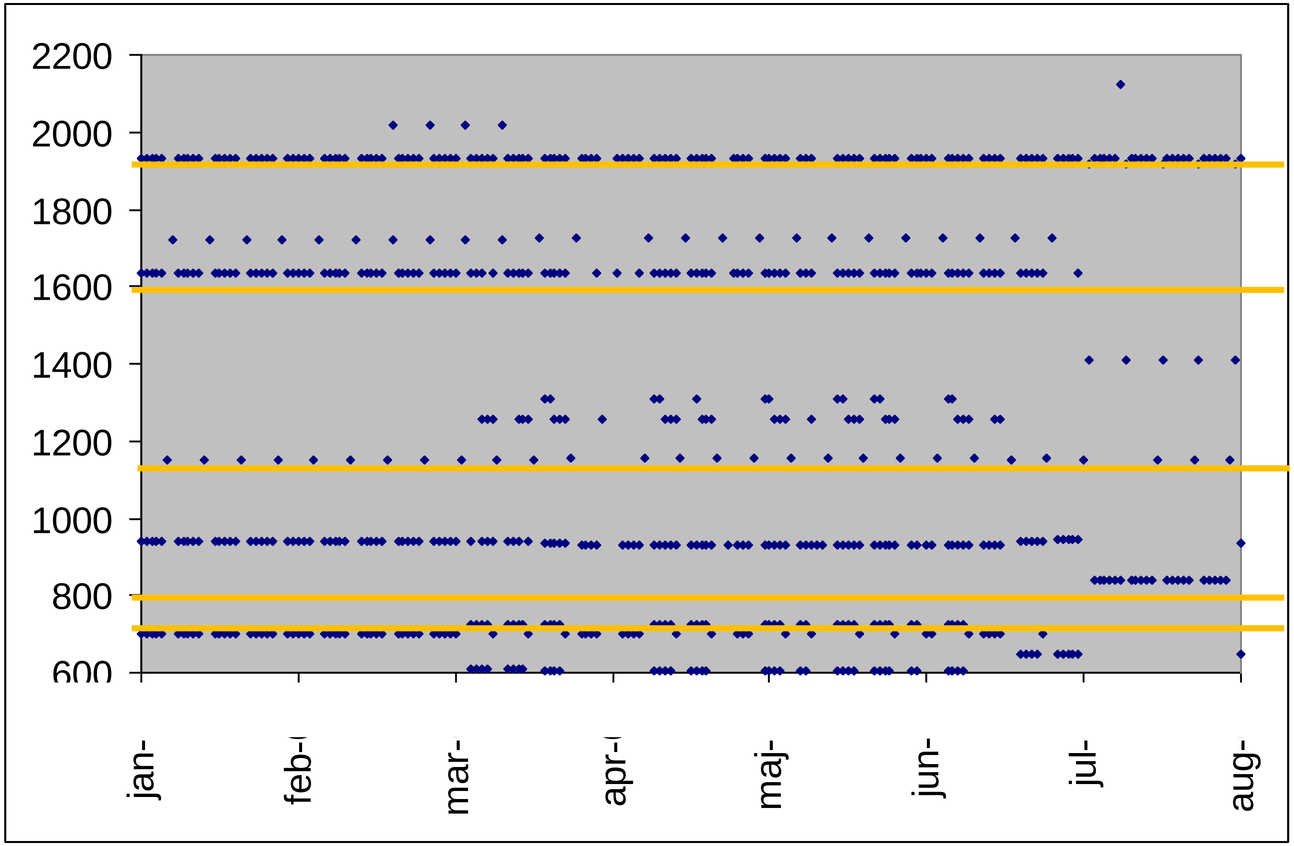
- Schedules are joined into virtual schedules.

Algorithm:

- Rank schedules according to sum of pax.
- Determine boundary as largest range where there are no overlap.
- Join schedules and repeat.

Schedule virtualization

Single leg

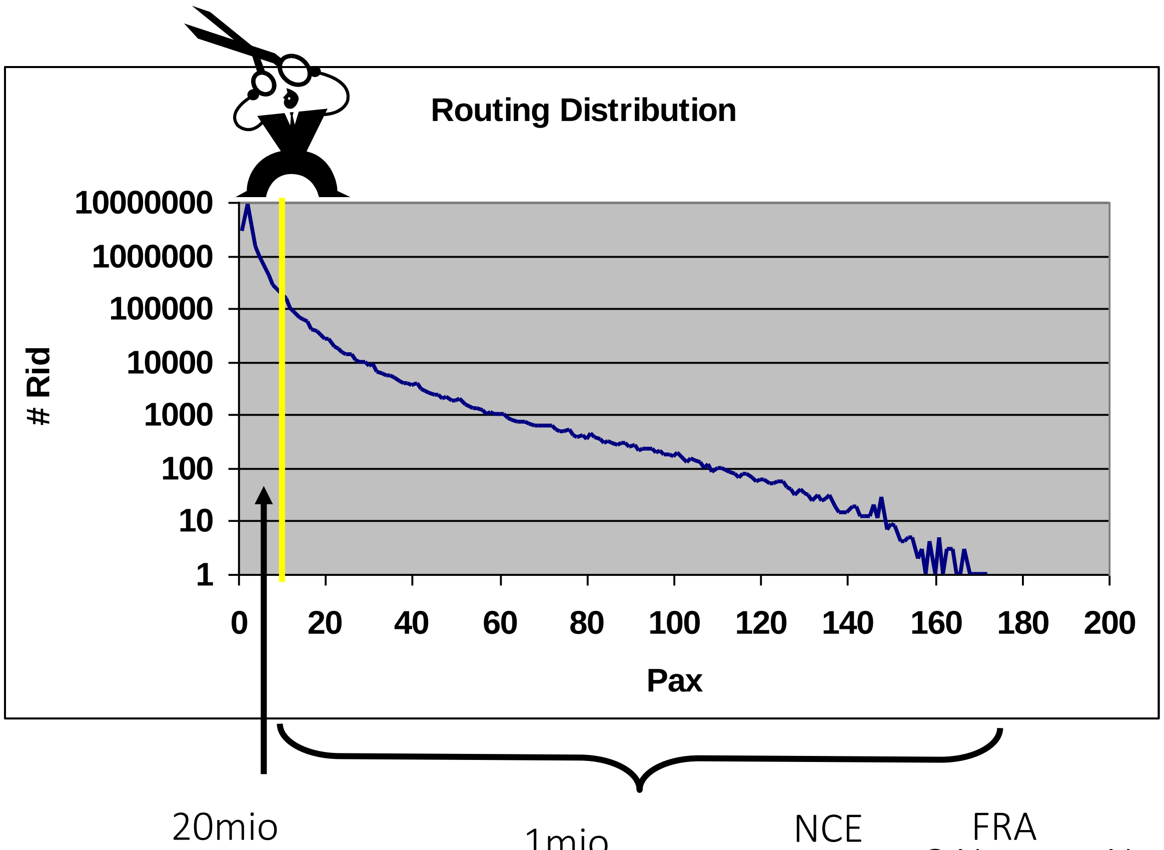


- Schedules are joined into virtual schedules.

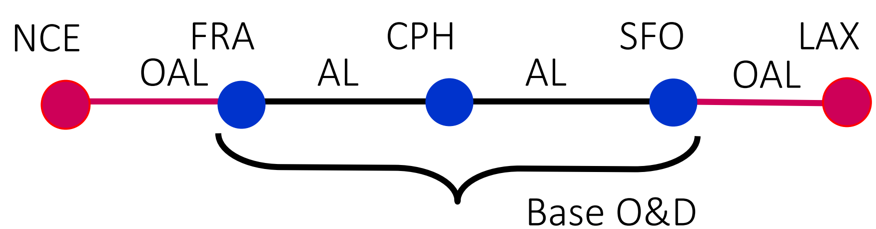
Algorithm:

- Rank schedules according to sum of pax.
- Determine boundary as largest range where there are no overlap.
- Join schedules and repeat.

Aggregating thin routings to base O&D



- Thin/Thick based on less/more than 5 pax per two years history.
- Cut thin routings to base O&D
- Remove OAL Prev/Next airline and city
- Point of Sale = "YY"
- POC = Control O&D On city

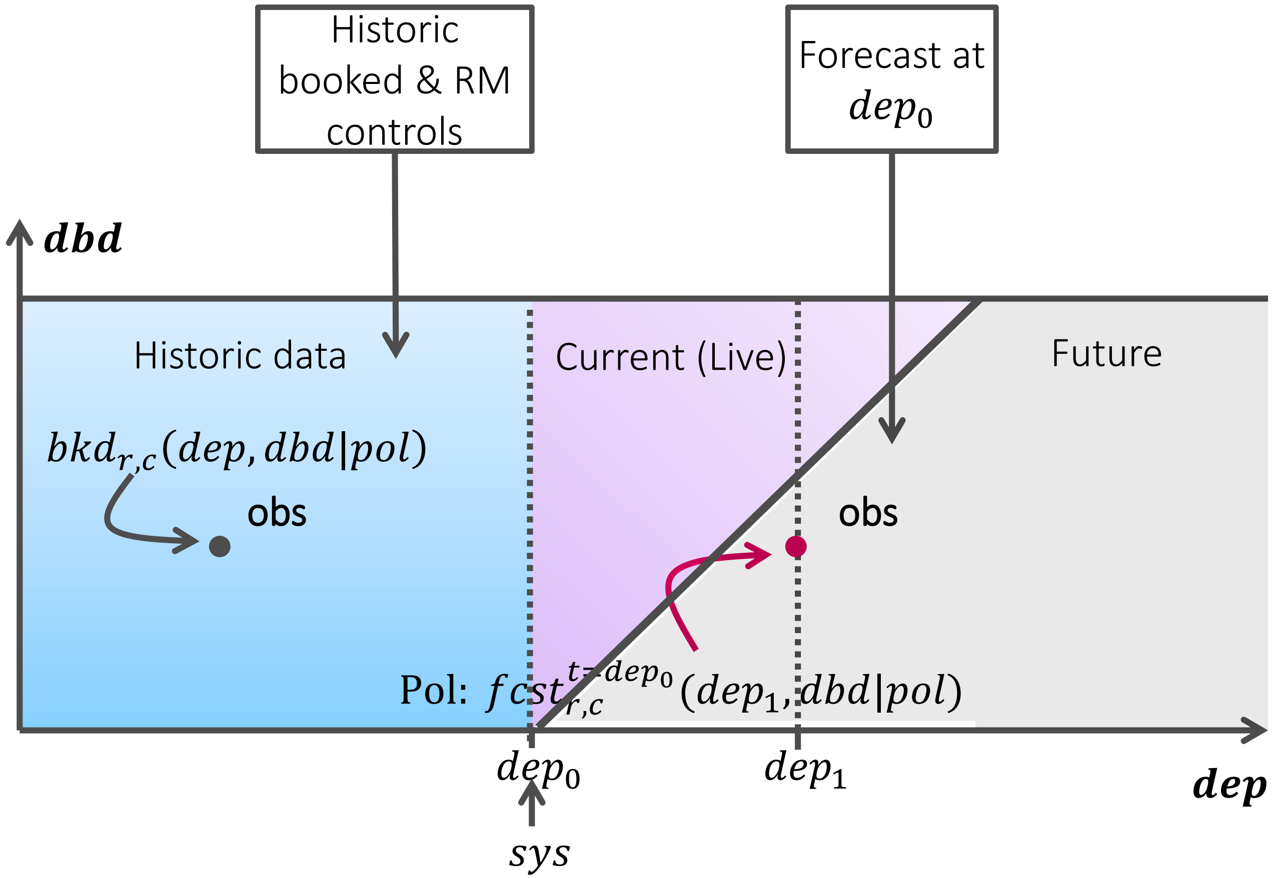


Note: Number of flows is scalable to airline size

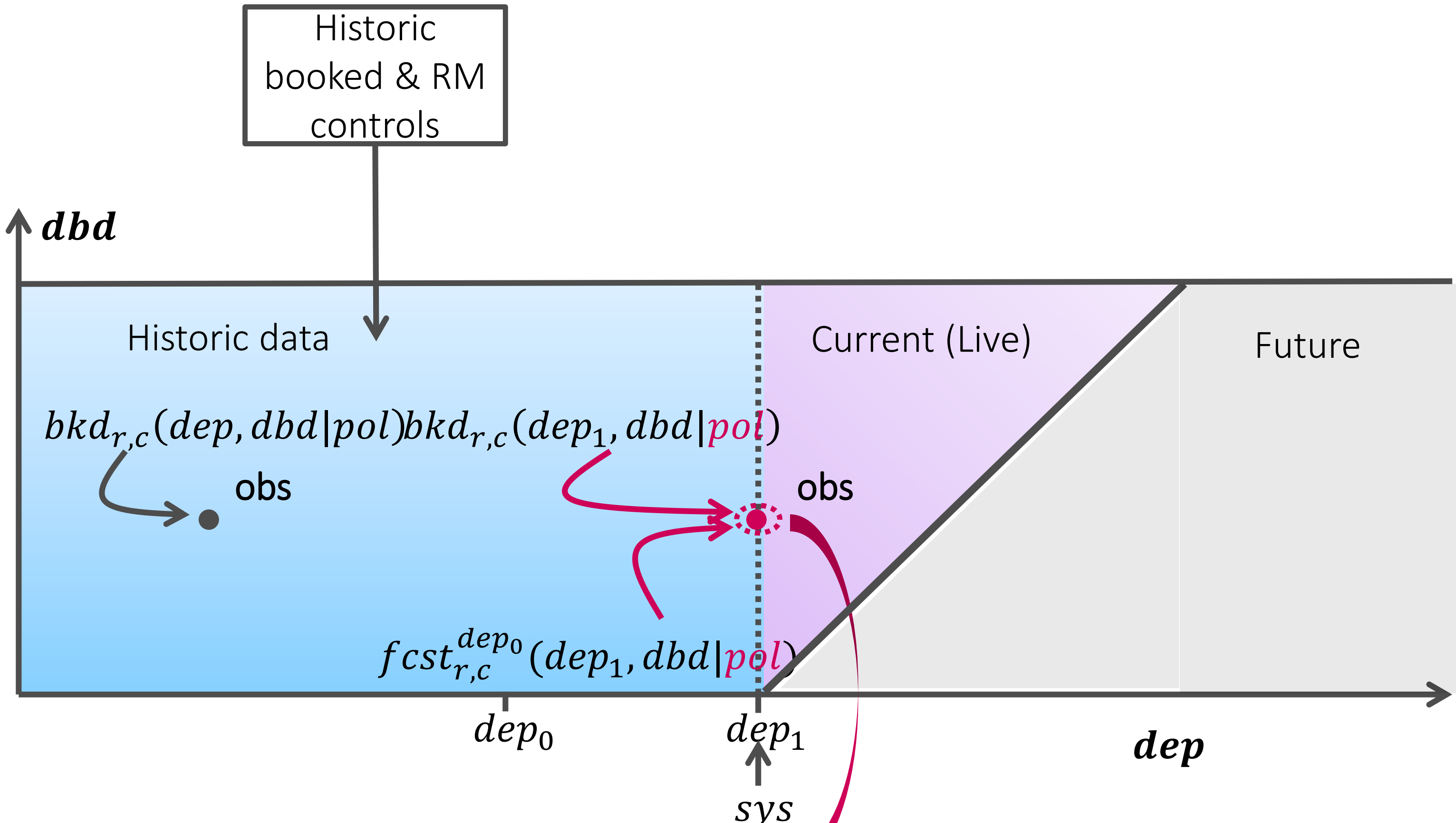
3.

Estimation

Forecast Quality Measurements



Forecast Quality Measurements



$$err_{r,c}^{dep_0} = fcst_{r,c}^{dep_0}(dep_1, dbd|pol) - bkd_{r,c}(dep_1, dbd|pol)$$

Parametric Forecasting Model

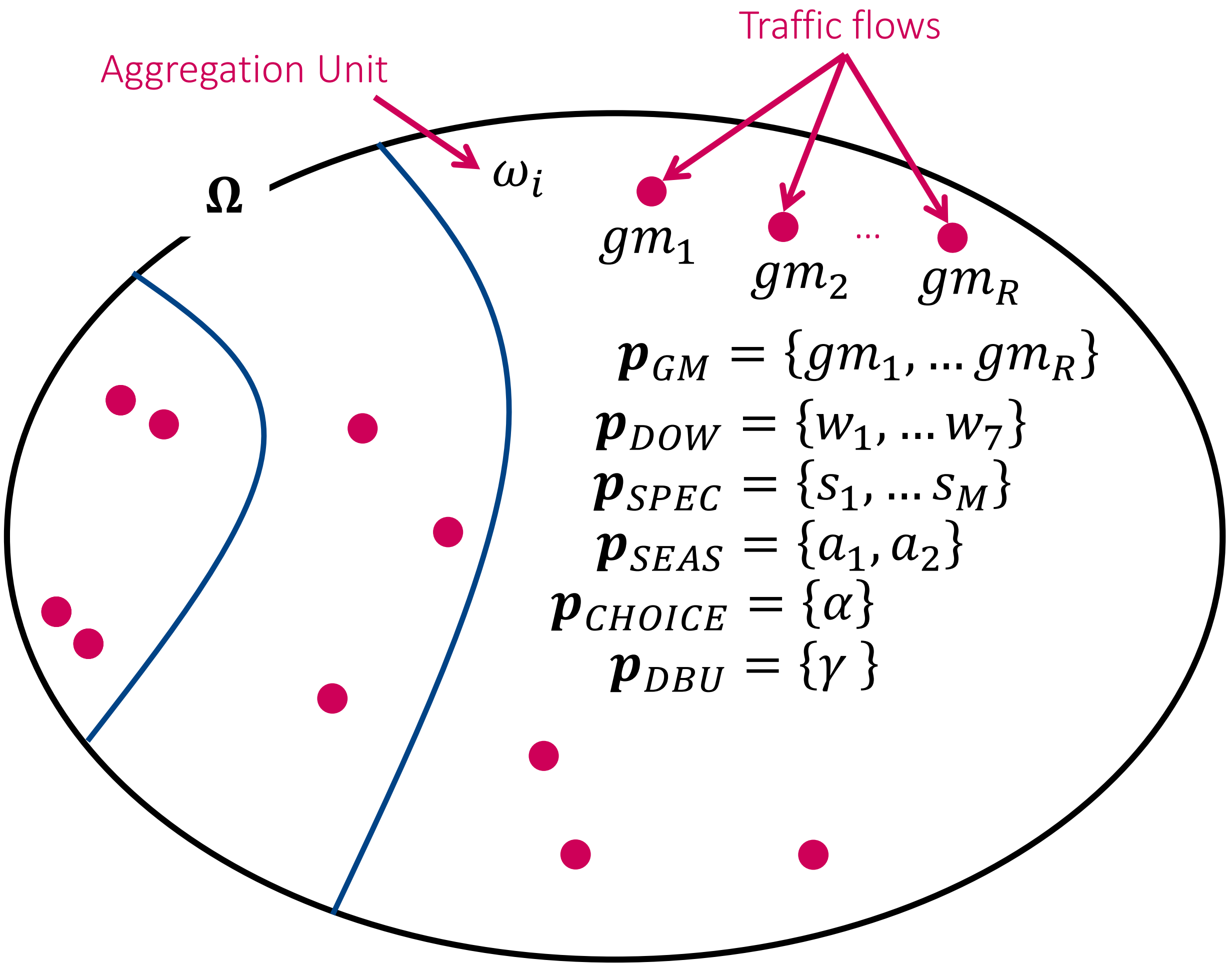
Single fare family

$$fcst_{r,c}(\boldsymbol{p}, dep, dbd | pol = \{c\}) = vol(\boldsymbol{p}, dep, dbd) \times MNL(\boldsymbol{f}, \boldsymbol{\beta}) =$$

$$\underbrace{\overset{(1)}{gm_r}\ \overset{(2)}{w_k}\overset{(3)}{s_m}\left(1 + a_1 \sin\left(2\pi \frac{doy}{365}\right) \overset{(4)}{+} a_2 \cos\left(2\pi \frac{doy}{365}\right)\right) \overset{(5)}{e^{-\gamma \cdot dbd}}(1 - e^{-\gamma})}_{\text{Volume}}\ \underbrace{\overset{(6)}{MNL(\boldsymbol{f}, \boldsymbol{\beta})}}_{\text{Choice prob.}}$$

	Description	Factor	Parameters	# Par.
(1)	Grand mean	$f_{GM,r}(\boldsymbol{p}_{GM}) = gm_r;\ r = 1, \dots, R$	$\boldsymbol{p}_{GM} = (gm_1, \dots, gm_R)$	R
(2)	Day of week	$f_{DOW}(\boldsymbol{p}_{DOW}, dep) = w_k;\ k = 1, \dots, 7$	$\boldsymbol{p}_{DOW} = (w_1, \dots, w_7)$	6
(3)	Special periods	$f_{SPEC}(\boldsymbol{p}_{SPEC}, dep) = s_m;\ m = 1, \dots, M$	$\boldsymbol{p}_{SPEC} = (s_1, \dots, s_M)$	$M - 1$
(4)	Seasonality	$f_{SEAS}(\boldsymbol{p}_{SEAS}, dep) = \left[1 + a_1 \sin\left(2\pi \frac{doy}{365}\right) + a_2 \cos\left(2\pi \frac{doy}{365}\right)\right]$	$\boldsymbol{p}_{SEAS} = (a_1, a_2)$	2
(5)	Differential build-up	$f_{DBU}(\boldsymbol{p}_{DBU}, dbd) = e^{-\gamma \cdot dbd}(1 - e^{-\gamma})$	$\boldsymbol{p}_{DBU} = \gamma$	1
(6)	Choice prob. (1 fare fam.)	$f_{PROB,c}(\boldsymbol{p}_{PROB} \| pol = \{c'\}) = e^{V(f,\boldsymbol{\beta})}/(1 + e^{V(f,\boldsymbol{\beta})})$	$\boldsymbol{\beta} = (\beta_{FF}, \beta_f)$	2

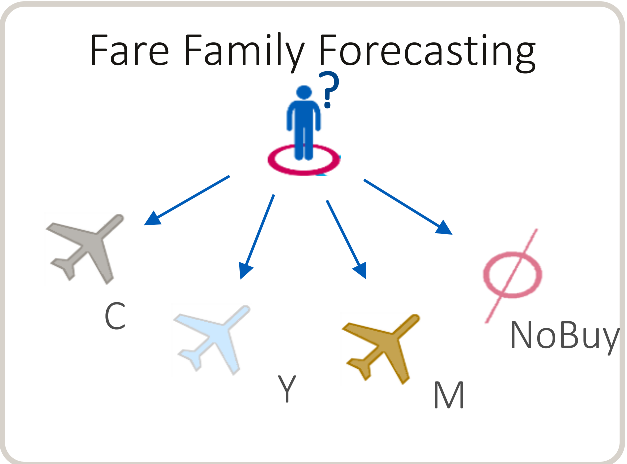
Parameters levels



- Traffic flows partitioned into aggregation units.
- Within agg. unit only one volume parameter per flow. All other parameters shared by the agg. unit.
- Number of free parameters: R+M+9.
- Robust: Parsimonious use of parameters, provides robustness for sparse data.

Choosing between multiple alternatives

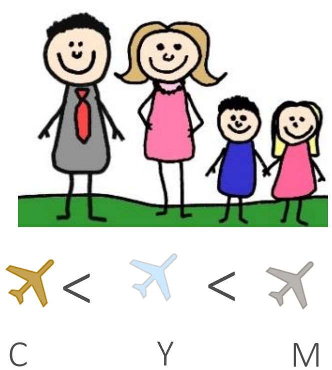
MNL model in Altéa RMS by customer type

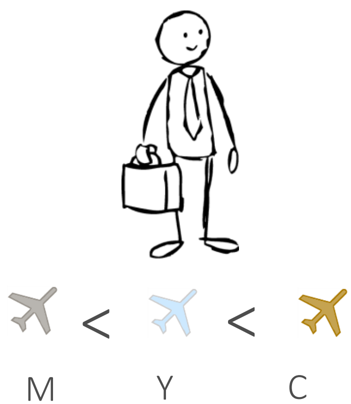


A potential customer can choose between different alternatives: buy product (in Fare Family CYM) or NoBuy.

$$Prob_C + Prob_Y + Prob_M + Prob_{NoBuy} = 1$$

Choice determined by the perceived relative utility of alternatives. The utility is different for everyone.
Depends partly on individual specific factors such as if travelling for work, or whether bringing children on the trip.





$$V_i = \beta_{FF,i} - \beta_f \cdot f_i$$

$$p_i = \frac{e^{V_i}}{\sum_j e^{V_j}}$$

amadeus

Estimation methodology

Log likelihood

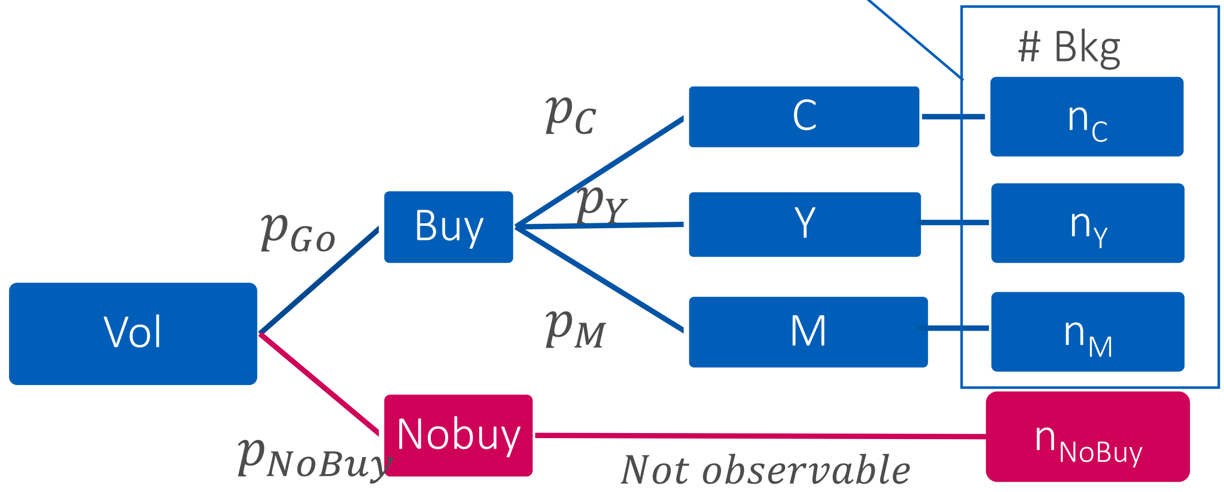
- Assume the arrival rate is Poisson distributed; and the choice prob. follow a multinomial distribution => Total likelihood of model parameters is the product of the two distributions, given the historical booking observations in C,Y, and M.

$$Likelihood\left(Vol,\boldsymbol{\beta} \mid n_C, n_Y, n_M, n_{NoBuy}\right) = Pois(Vol)\ \mathrm{Multinom}\left(\boldsymbol{\beta}, n_C, n_Y, n_M, n_{NoBuy}\right)$$

- It can be simplified into a log-likelihood function:

$$LL = \sum_{obs}\left(n_{tot}\ \ln(fcst) - fcst +\right.$$

p_i^* is the prob of buying the alternative i *given* that the customer buys something*)



*) Jeffrey P. Newman, Mark E. Ferguson, Laurie A. Garrow, Timothy L. Jacobs (2014) Manufacturing & Service Operations Management 16(2):184-197.

Solution $(\partial LL_\omega / \partial \boldsymbol{p})(\boldsymbol{p}^*) = \mathbf{0}$

Solve week index k=1,..,7 $\quad \dfrac{\partial LL}{\partial w_k} = 0$

↓

Solve special index m=1,..,M $\quad \dfrac{\partial LL}{\partial s_m} = 0$

↓

Solve WTP β $\quad \dfrac{\partial LL}{\partial \beta} = 0$

↓

Solve season a_1, a_2 $\quad \dfrac{\partial LL}{\partial a_1} = 0, \dfrac{\partial LL}{\partial a_2} = 0$

↓

Solve grand mean r=1,..,R $\quad \dfrac{\partial LL}{\partial gm_{vrid}} = 0$

Iterate until convergence

- **Consistent solution**: All forecast components solved simultaneously.
- **Separable equations**: Due to parametric model are decomposed into separate components.
- **Unique optimum**: Each forecast component is convex in the unknown parameters.
- **Analytical solution**: Properties of MSE and BIAS error measures enable an analytical solution.
- **Advantages**:
  - Robustness
  - Accuracy
  - Computational Speed

Strictly CONFIDENTIAL

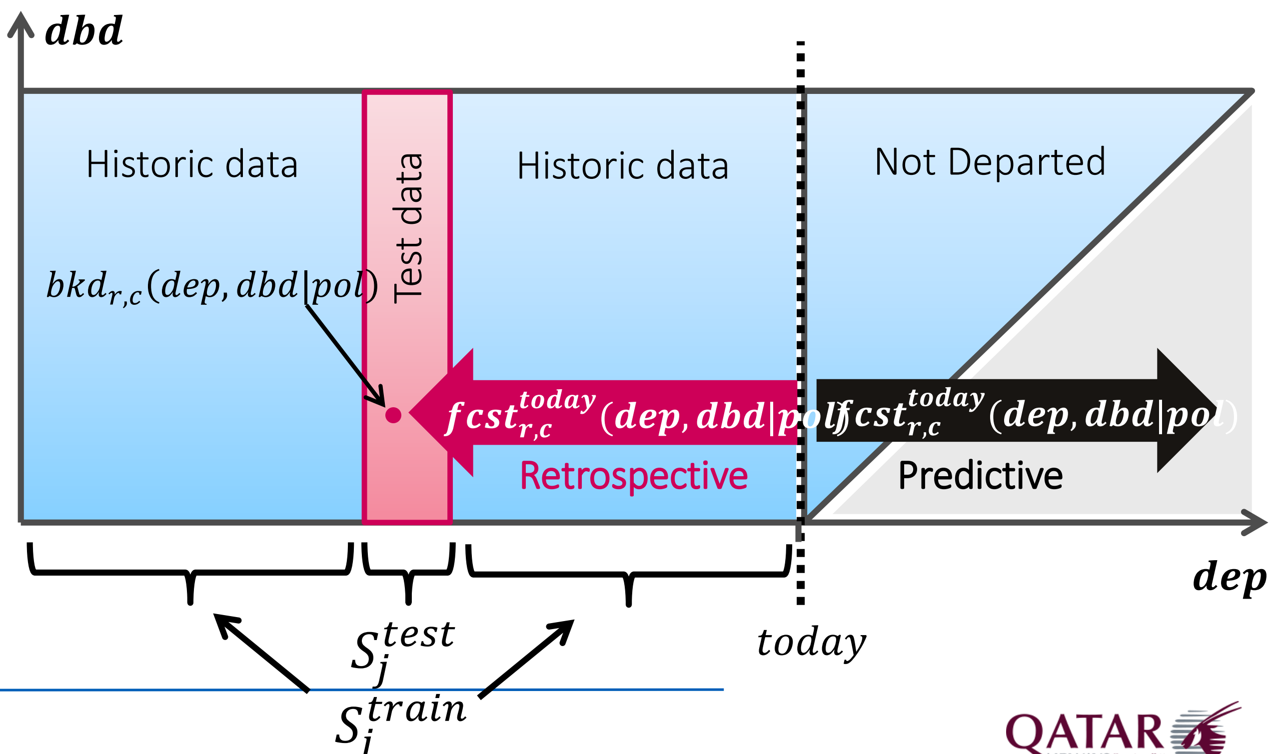
4.

Back testing - Validation

Real Airline Data

Jack-knife resampling techniques on retrospective forecast

- Jack-knife resampling techniques
  - Construct N resampled datasets from the original sample data, by repeatedly splitting into a training S_j^{train} and a test dataset S_j^{test} based on the departure dates.
- Compare retrospective forecast with historical observations
  - Compare at the lowest level. Aggregate in various dimensions.

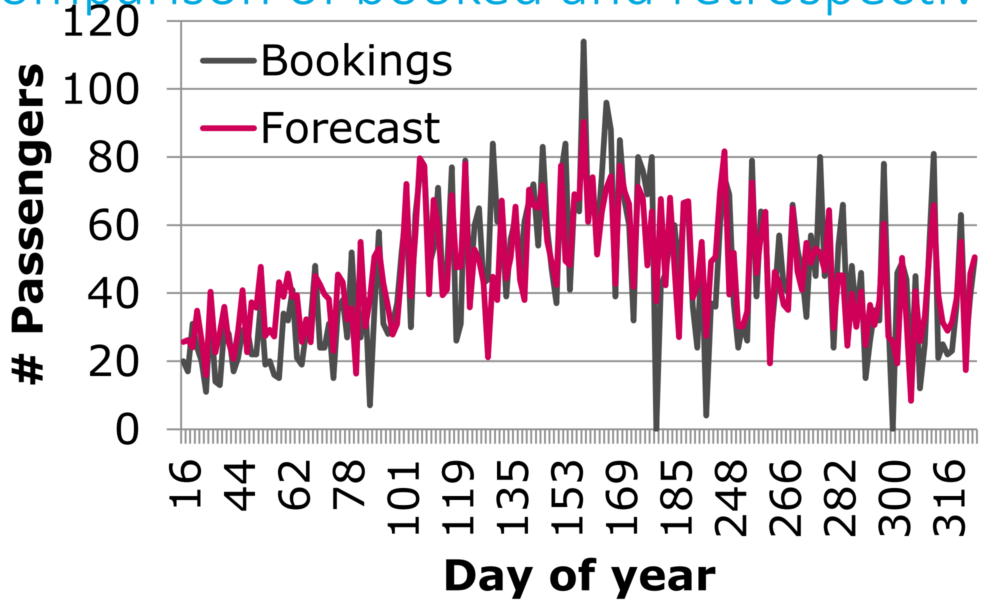


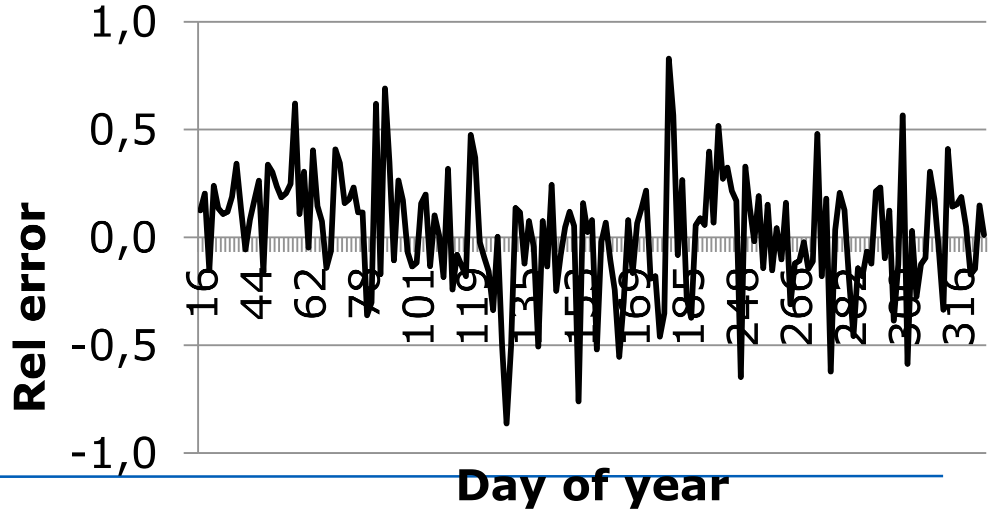
QATAR AIRWAYS القطرية

amadeus

Real Airline Data

Comparison of booked and retrospective forecast

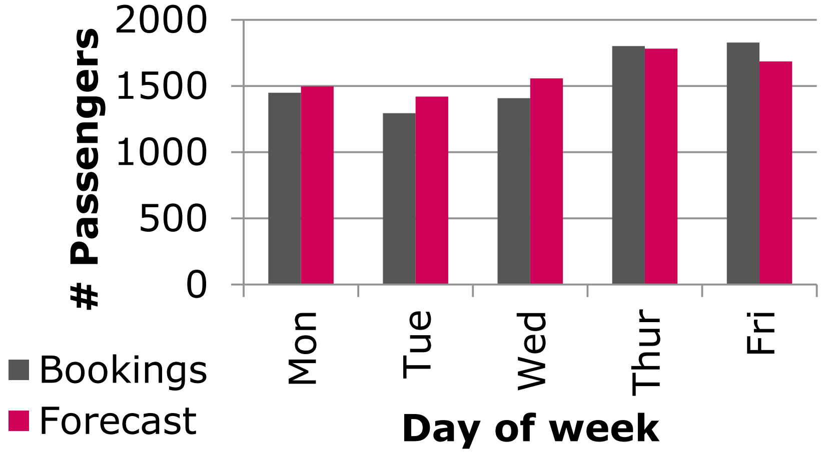


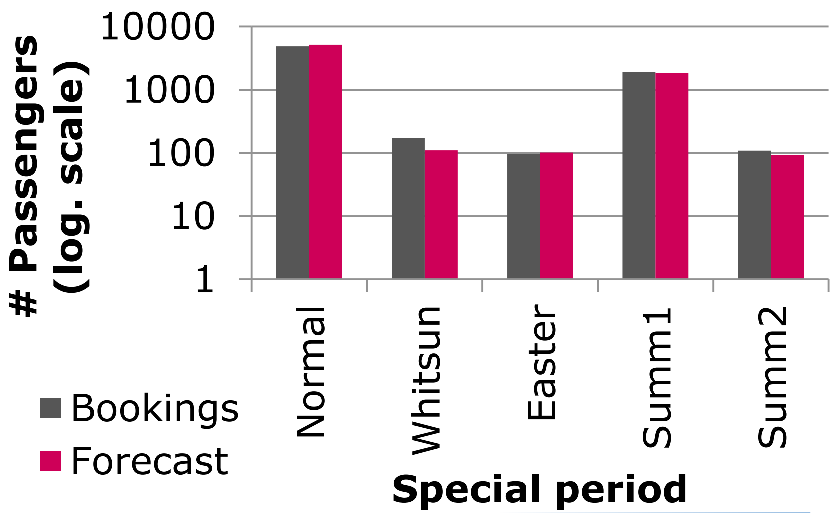


- Total no. of book: 7,780.
- Total no. of obs.: 1.4×10^6
- Total no. parameters: 33
- **No systematic errors**: The forecasting model accurately captures the underlying factors.

Real Airline Data

Comparison of booked and retrospective forecast (continued)

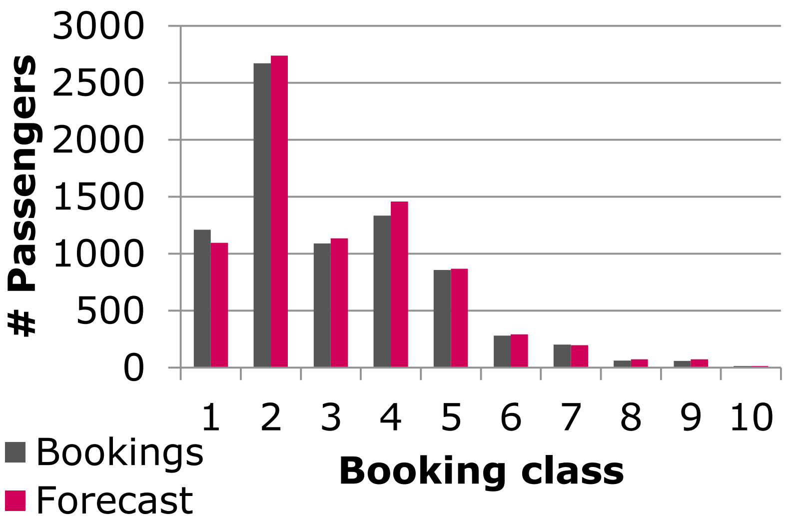


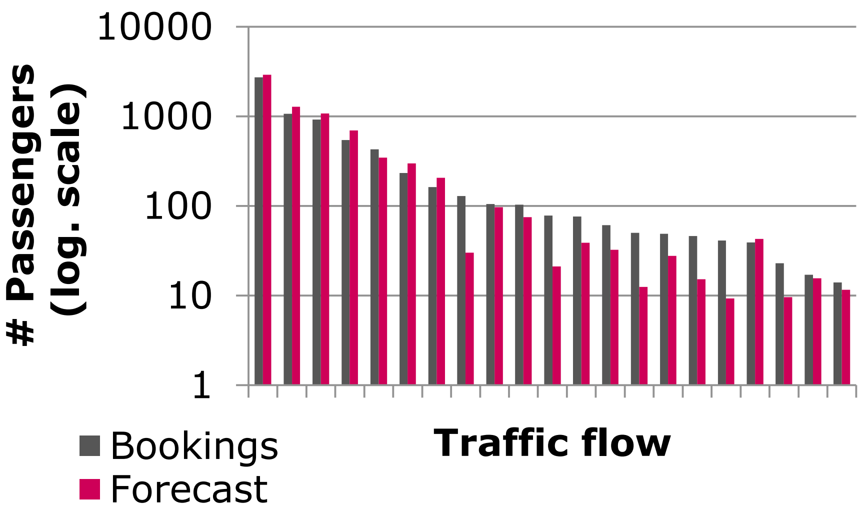


- **Excellent forecast quality for all forecast components:** Observed here for DOW and Special period.

Real Airline Data

Comparison of booked and retrospective forecast (continued)





- **Excellent forecast quality across class dimension:** Provides accurate estimates of the sell-up forecast parameters
- **Excellent forecast quality for flow volume**: Observed here over two decades.

5.

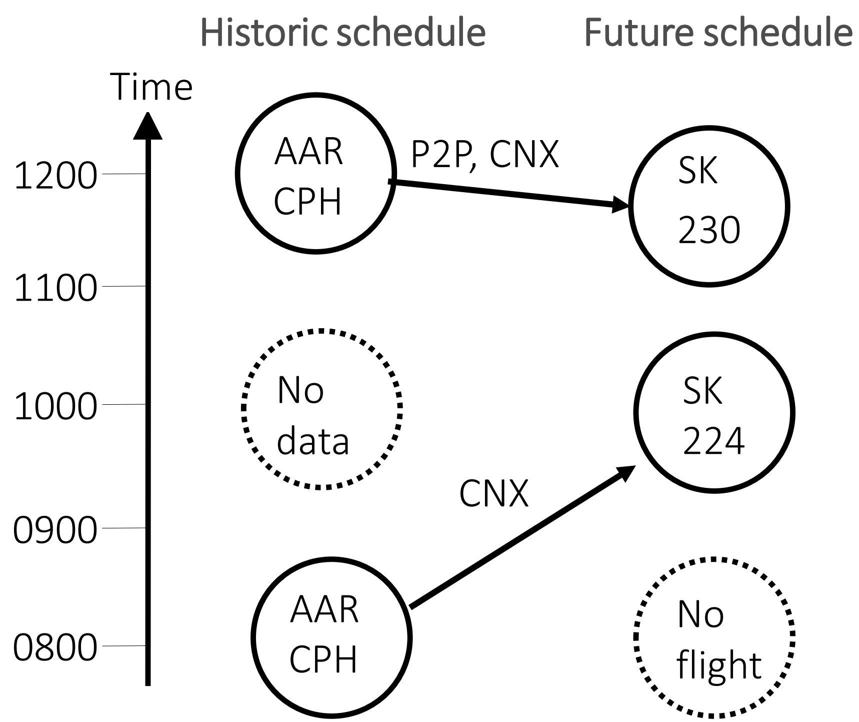
Forecasting - Projection

Build Schedules

5. Build-Schedules

Matching the future schedules with historical schedules.

- Builds (all) feasible O&D schedules incl. OAL schedules and selects closest match to historic departed.
- New flights (DOW) automatically handled by forecast parameters.
- New destinations/frequencies: Automatic sponsors based on proximity.
- Dead-ends. Re-route dead-ends to alternative if appropriate.
- Live PNRs used to create new traffic flows.
- Interventions used to create user defined CNX



Forecasting - Projection

- Different travel solutions possible.
- A potential customer can choose between different alternatives.
- Demand is split between the alternatives.
- Choice determined by the perceived relative utility of alternatives.

$p_1 = 80\%$
AMSNRT

$p_2 = 18\%$
AMSTPE
TPENRT

$p_3 = 2\%$
AMSBKK BKKTPE
TPENRT

Trip Duration (T)
All flights and Layovers

Number of Stops (S)
Or number of connections

Trigonometric Time Preferences
(depTime, arrTime)

Schedule Utility V

$$p_i = \frac{e^{V_i}}{\sum_j e^{V_j}}$$

$$V_i = \beta_s \cdot S_i + \beta_t \cdot T_i + \sum_{k=1}^{3} \beta_{sk} \cdot \sin(2k\pi t)_j + \beta_{ck} \cdot \cos(2k\pi t)_j$$

Strictly CONFIDENTIAL

6.

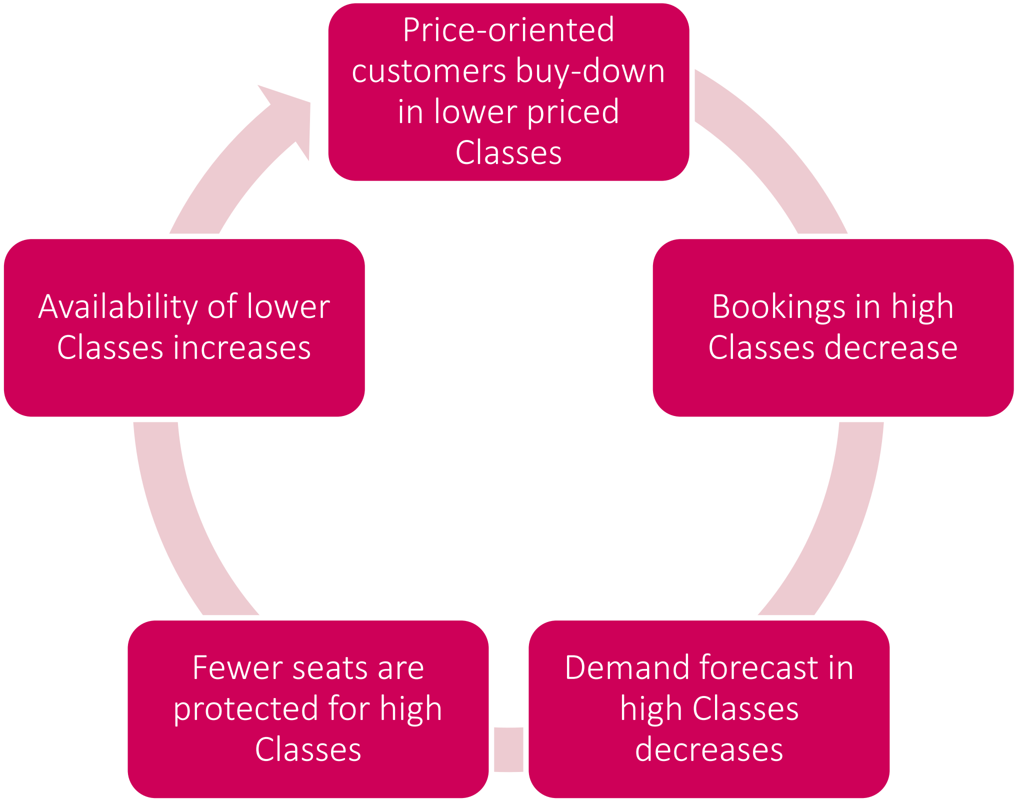
Transformation of dependent demand to independent demand

QATAR | amadeus

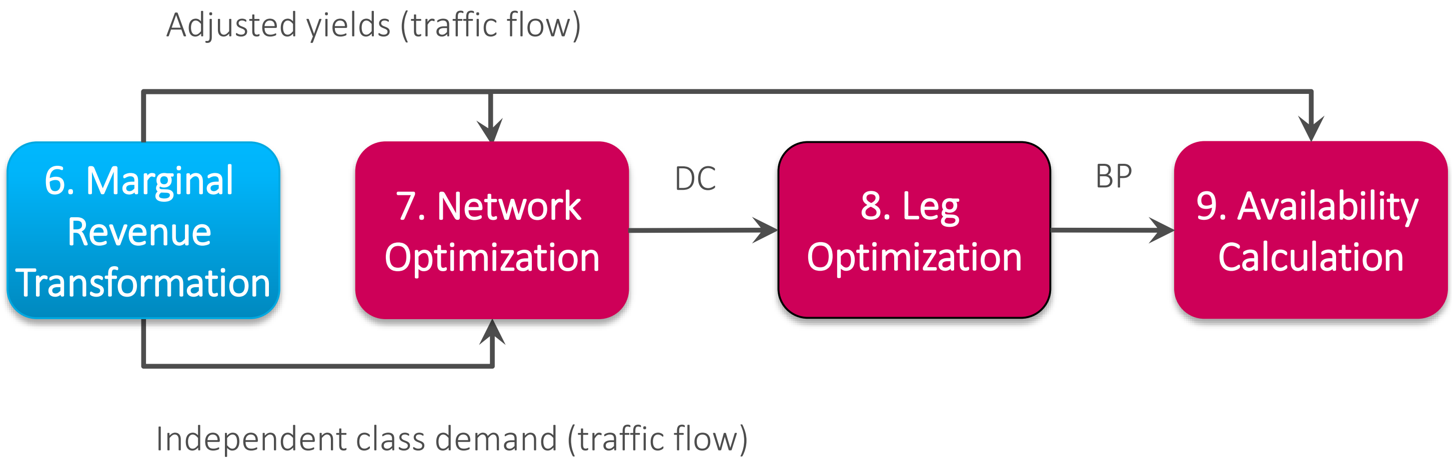
Simplified fare-structures → Spiralling Down

6. Marginal Revenue Transformation

Confidential
RESTRICTED



Marginal Revenue Transformation



- Marginal Revenue Transformation calculates the marginal revenue and demand of accepting a booking considering the risk of buy-down.
- Marginal Revenue are calculated daily for all traffic flows, depdates and days to departure.
- The adjusted fares are used in both optimization and availability calculation.
- This ensures the right controls apply consistently in the real-time availability decision.

Marginal Revenue Transformation

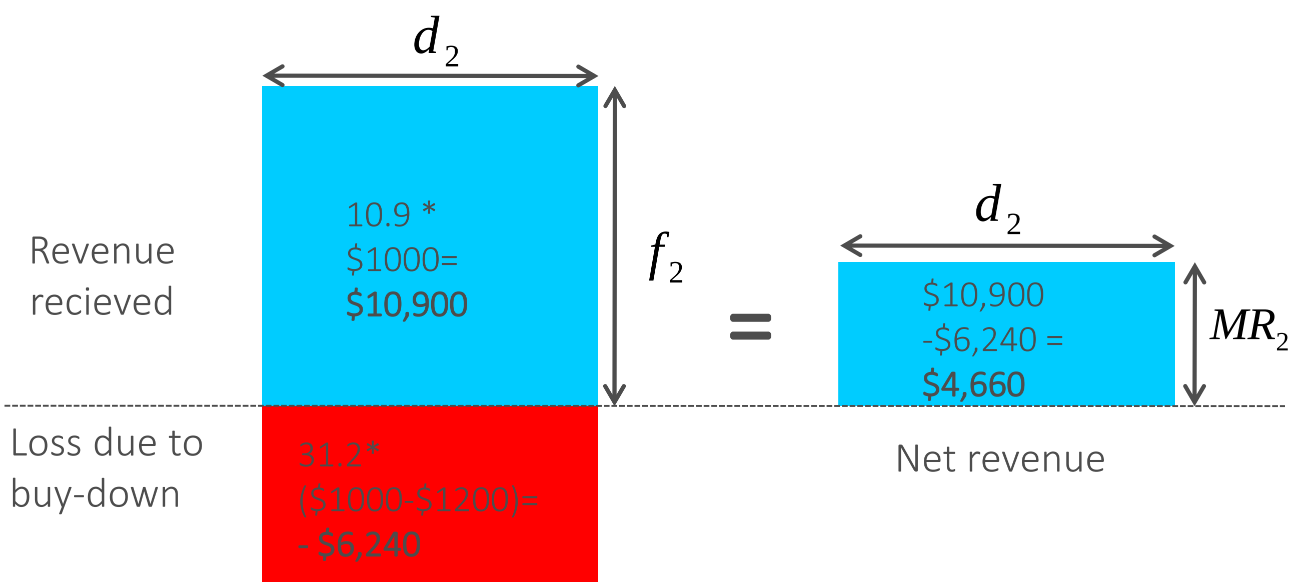
Single leg - Deterministic demand

$$Q_k = \sum_{j=1}^{k} d_j$$

$$TR_k = f_k Q_k$$

			Fully un-differentiated	
fi	di	Qi	TRi	MRi
\$1.200	31,2	31,2	\$37.486	\$1.200
\$1.000	10,9	42,2	\$42.167	\$428
\$800	14,8	56,9	\$45.536	\$228
\$600	19,9	76,8	\$46.100	\$28
\$400	26,9	103,7	\$41.486	-\$172
\$200	36,3	140,0	\$28.000	-\$372

$$MR_2 = \frac{TR_2 - TR_1}{Q_2 - Q_1} = \frac{\$42{,}167 - \$37{,}486}{42{,}2 - 31.2} = \$428$$

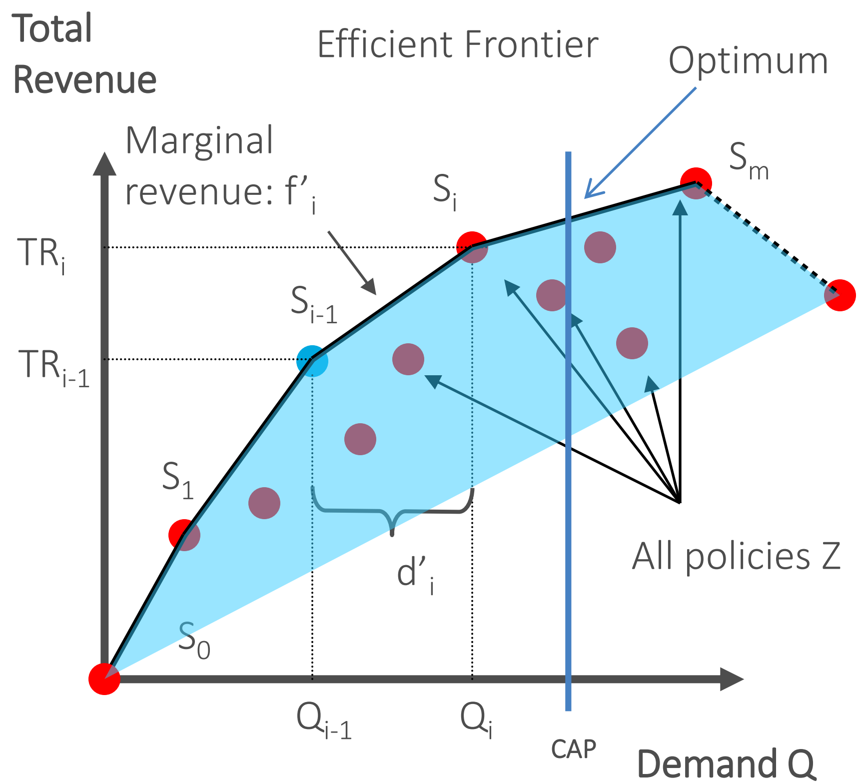


$$MR_2 = \frac{\$4{,}660}{10.9} = \$428$$

Optimization: General Formulation

Arbitrary fare structure

Fare products	$f_j,\ j=1,\ldots,n$
Policy $Z \subseteq N$ (any set of open classes)	$\{\},\{1\},\{1,3\},\ldots$
Demand	$d_j(Z)$
Accumulated Dem.	$Q(Z)=\sum_{j\in Z} d_j(Z)$
Total Revenue	$TR(Z)=\sum_{j\in Z} d_j(Z)\, f_j$
Objective	$\max TR(Z)$ s.t. $Q(Z)\le cap$



Marginal Revenue Transformation theorem

Policies on the convex hull → **Independent demand**

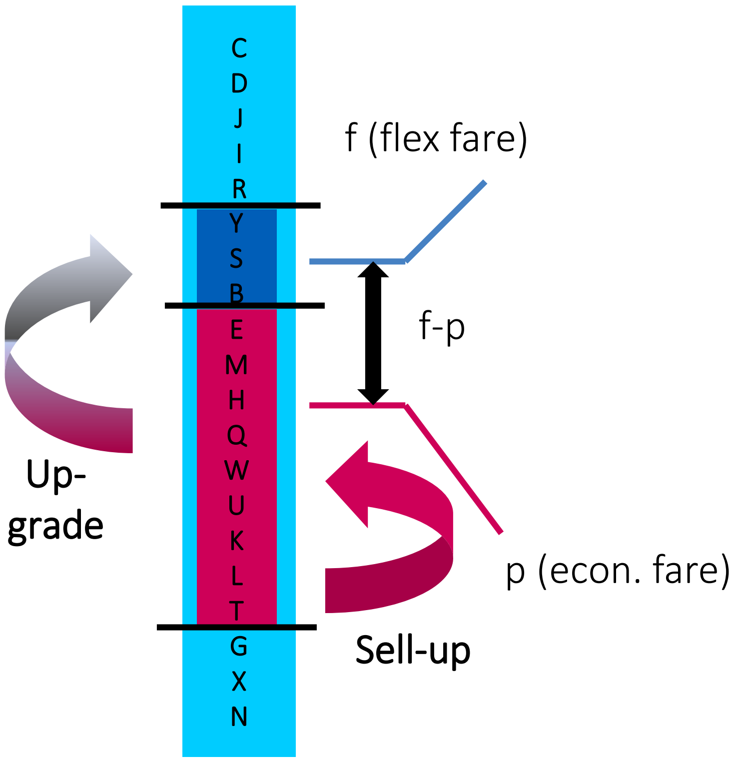
Policy	Dem.	TR
S_1	Q_1	TR_1
S_2	Q_2	TR_2
...	...	
S_m	Q_m	TR_m

Partition Dem.	Adj. Fare
$d_1^{'} = Q_1$	$f_1^{'} = f_1$
$d_2^{'} = Q_2 - Q_1$	$f_2^{'} = \left(TR_2 - TR_1\right) / d_2^{'}$
...	...
$d_m^{'} = Q_m - Q_{m-1}$	$f_m^{'} = \left(TR_m - TR_{m-1}\right) / d_m^{'}$

Marginal Revenue Transformation theorem

- The transformed policies are independent.
- Optimization using the original fare structure and the marginal revenue transformed in policy space gives identical results.

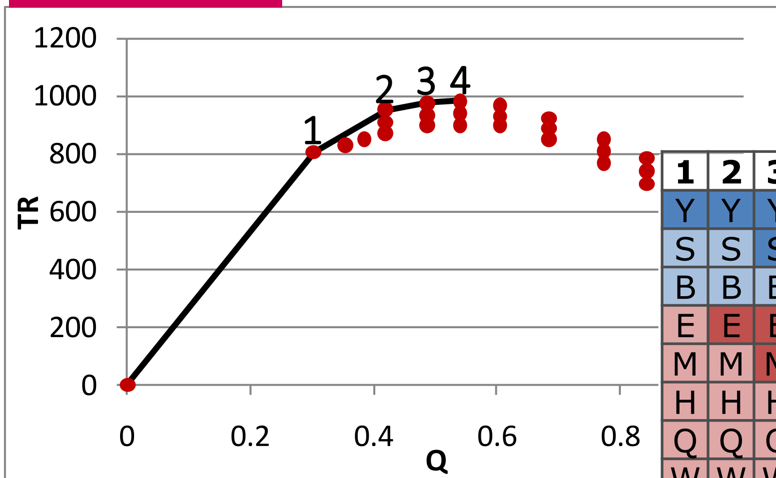
Fare Families from a RM perspective



- Definition:
  - The fare structure consists of two or more fare families. Each fare family has the same set of restrictions (within a family price is the only difference).
  - Shown are the open fares in flex (f) and the open fare in economy (p). f-p is the up-grade level.
- RM objective: max total revenue requires:
  - Forecasting: Demand model that for any policy (p,f) predicts the demand in economy and flex.
  - Optimization: Determine the set of efficient policies (p,f) by path. The ordering sequence of the policies. Bid-price for the legs.

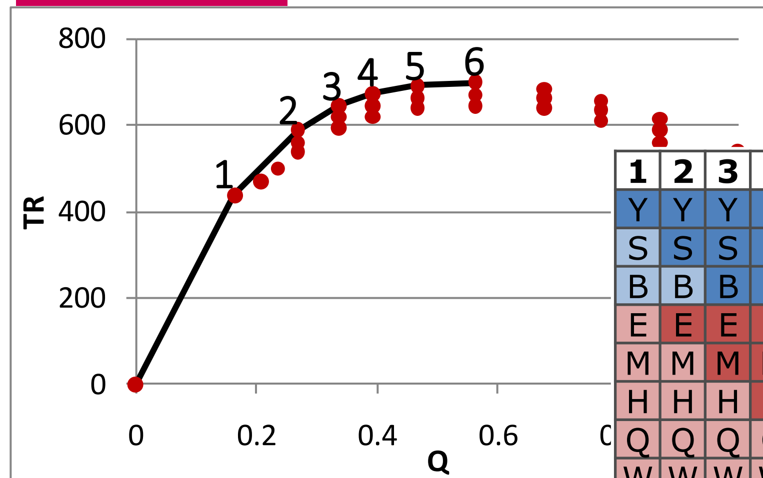
MRT applied to Fare Families

DBD=1

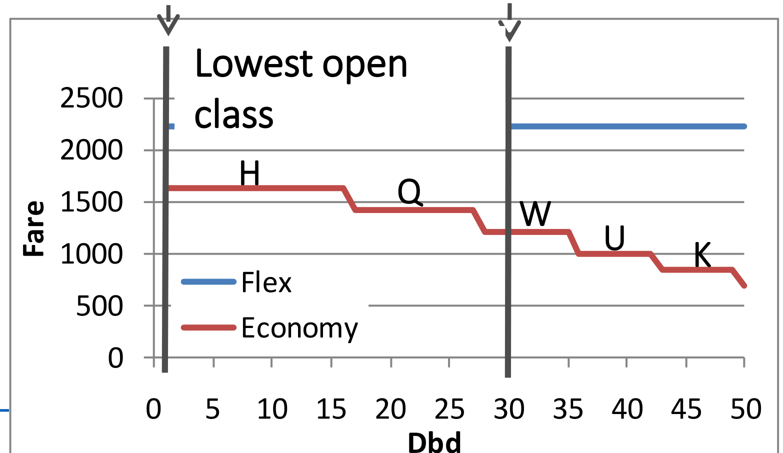


1	2	3	4
Y	Y	Y	Y
S	S	S	S
B	B	B	B
E	E	E	E
M	M	M	M
H	H	H	H
Q	Q	Q	Q
W	W	W	W

DBD=30



1	2	3	4	5	6
Y	Y	Y	Y	Y	Y
S	S	S	S	S	S
B	B	B	B	B	B
E	E	E	E	E	E
M	M	M	M	M	M
H	H	H	H	H	H
Q	Q	Q	Q	Q	Q
W	W	W	W	W	W



7.

Bayesian forecasting methodologies

Strictly CONFIDENTIAL

References

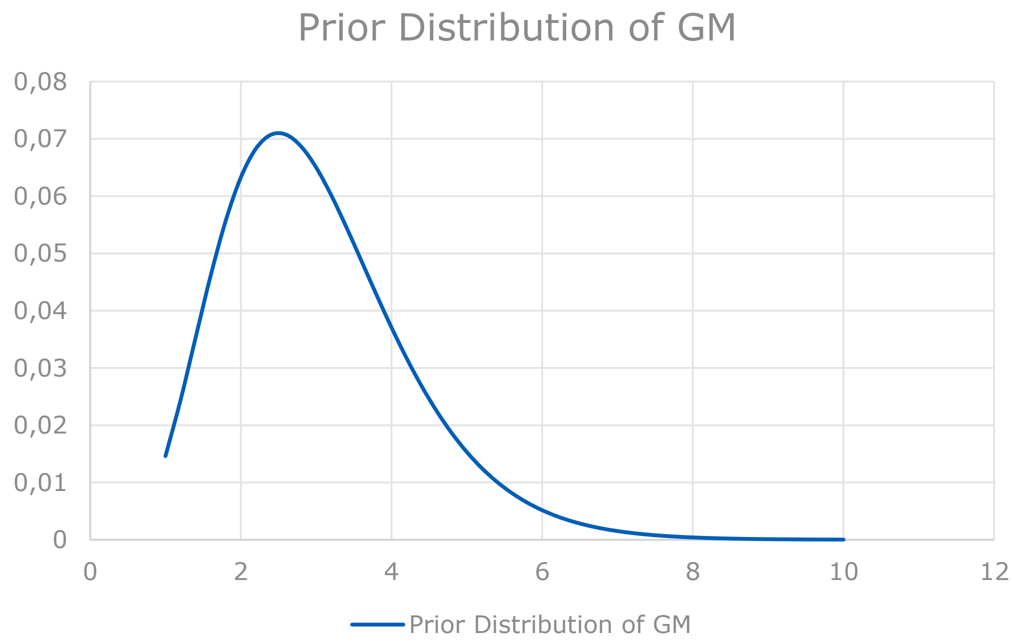
1. Walczak, D., Boyd,AE., Cramer, R., 2012. Revenue Management. In: Quantitative Problem Solving Methods in the Airline Industry. Springer.
2. Talluri, KT., Van Ryzin, GJ. Revenue Management. Springer
3. Walczak, D., Kambour, E., 2013. Revenue Management for fare families with price-sensitive demand. J. Revenue and Pricing Management. 13: 273-290.

Bayesian estimation example – average demand (GM)

- Estimate average demand from booking data

depdate	doy	weekday	pax
2013-01-01	1	2	5
2013-01-02	2	3	0
2013-01-03	3	4	2
2013-01-04	4	5	1
2013-01-05	5	6	4
2013-01-06	6	7	0
2013-01-07	7	1	2
2013-01-08	8	2	1
2013-01-09	9	3	1
2013-01-10	10	4	2
2013-01-11	11	5	6
2013-01-12	12	6	6
2013-01-13	13	7	6
2013-01-14	14	1	4
2013-01-15	15	2	4
2013-01-16	16	3	2

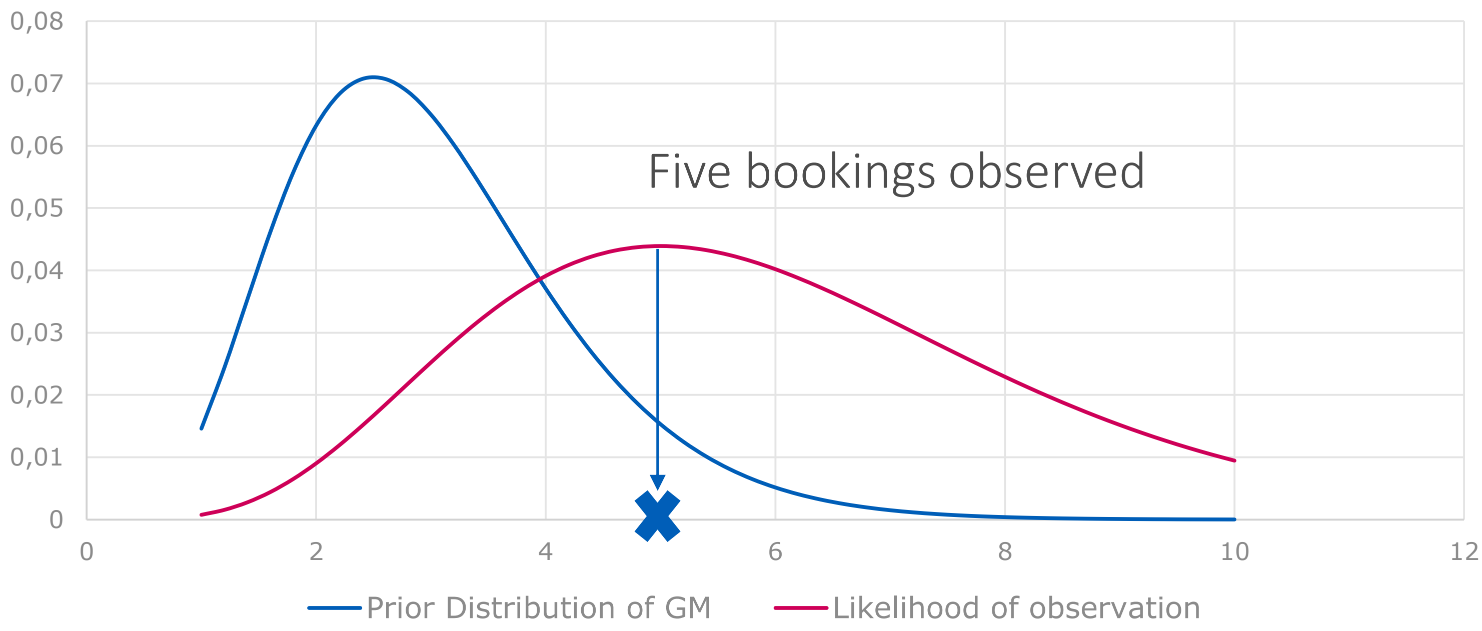
Original best idea of the average demand



The likelihood of each new observation is computed

One booking count observation

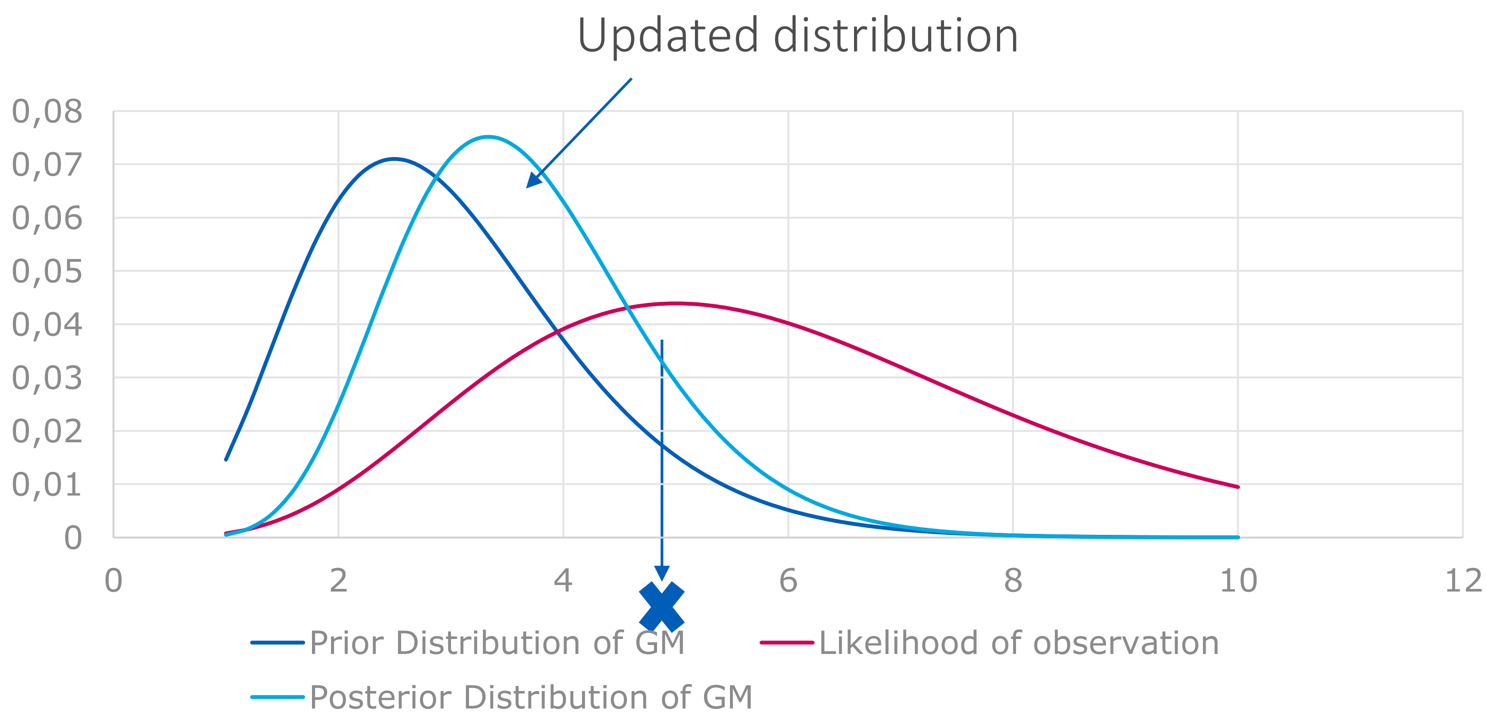
depdate	doy	weekday	pax
2013-01-01	1	2	5



The likelihood of each new observation is computed

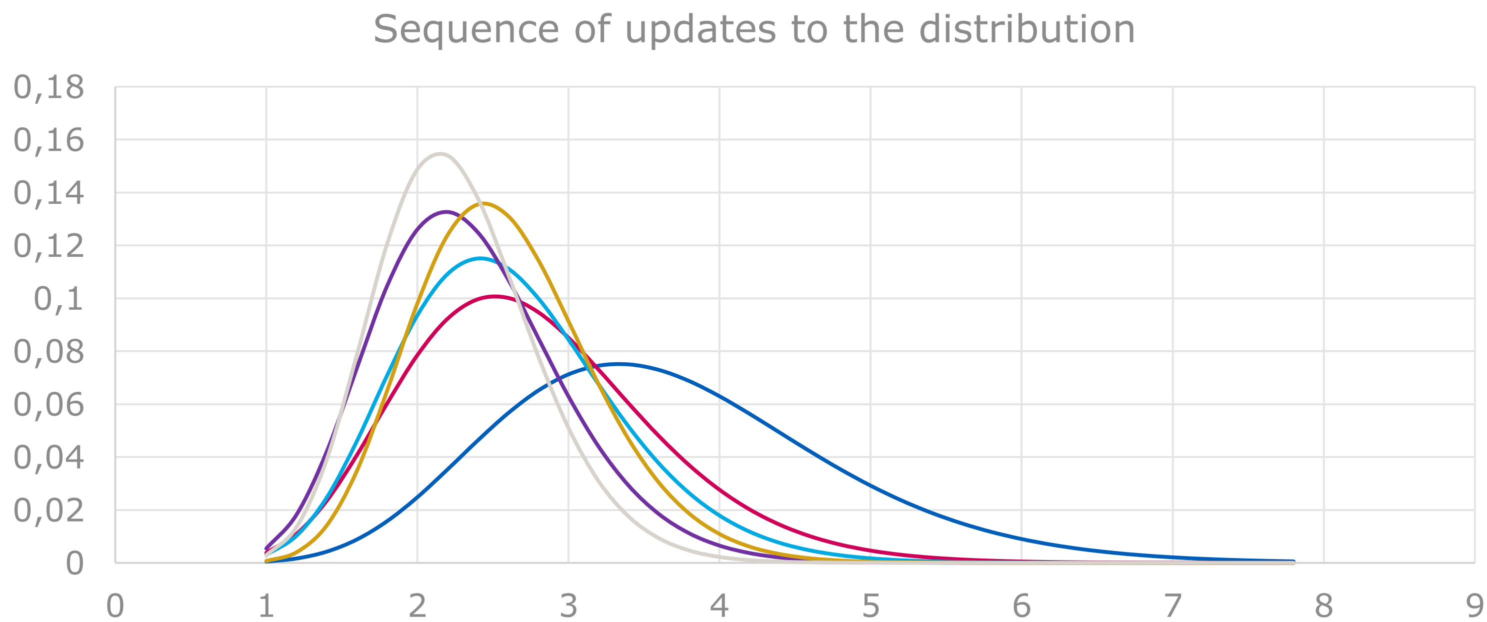
One booking count observation

depdate	doy	weekday	pax
2013-01-01	1	2	5



Each new observation leads to a new updated distribution

depdate	doy	weekday	pax
2013-01-01	1	2	5
2013-01-02	2	3	0
2013-01-03	3	4	2
2013-01-04	4	5	1
2013-01-05	5	6	4
2013-01-06	6	7	0
2013-01-07	7	1	2



Confidential RESTRICTED

Bayesian estimation methodology

Conclusion for Bayesian estimation

- Relative to static model, Bayesian estimation delivers:
  - Improved estimation via adaptation to "local" parameters, if autocorrelation of market conditions is important
  - Better knowledge about the uncertainty about regression parameters
  - Adaptability to (small) changes in market
- Concerns
  - Updates after each obseration creates noise
  - Difficult based on a single observation to decouple effects from multiple factors, dow, seasonality, special events, etc.

8.

Conclusion

	Altéa O&D forecast model	Traditional forecast models	PODS
Volume forecast by DOW, Seasonality, Special day, Trends	**Yes** (estimation done in consistently)	**Yes** (estimation done in multiple steps)	**No** (Only seasonality possible)
Un-constraining	**Yes** (Done implicitly by minimize constrained forecast error)	**Yes** (Done explicitly by un-constraining obs. Cannot be observed/verified	**Yes** (Done explicitly by un-constraining obs. Cannot be observed/verified
Forecasts type	**Net-forecasts** (True survived bookings from PNRs)	**Gross-forecasts** (Most advanced. Canc method 3 from PODS)	**Net and Gross-forecasts** (Currently testing)
Use live data in estimation	**No** (Only departed observations used)	**Yes/No** (Can use live data but issue with canc.)	**Yes/No** (Available, but currently not used)
Cancellations	**Yes** (Forecast for no-show and cancellations)	**Yes** (Forecast for no-show and cancellations)	**Yes** (Currently testing)
Estimation methodology	**Parametric** Analytical using all historical data	**Parametric/Non-parametric** Update forecast parameters based on historical obs.	**Non-parametric** Update forecast parameters based on historical obs.
Calibration	**Automated** Self-calibrating every week using historic data	**Manual** Done at installation	**Automated** During warm-up

	Altéa O&D forecast model	Traditional forecast models	PODS
O&D Traffic flow, O&D data collection	**Yes** (Booked, Yield, Taxes, Ava at the detailed level)	**(Yes)** (Limited access to Ava. and Tax at O&D level)	**(Yes)** Access to AVA. Tax excluded. All flows significant.
Supported fare structures	Fenced, Fenceless, **Fare Families**	Fenced, Fenceless, **Fare Families NOT supported**	Fenced, Fenceless, **Fare Families**
Monitoring forecast quality	**Yes** (Constructed to minimize constrained forecast error)	**No** (Fenced fare str. Or base demand for fenceless.)	**No** (Fenced fare str. Or base demand for fenceless.)
Estimation of the Volume at O&D traffic flow level	**Yes**	**Yes**	**Yes**
Estimation of the sell-up rate (Willingness to pay)	**Yes**	**No/(Yes)** (normally at segment level)	**No/(Yes)** (currently frat5 input, but can be estimated system-wide by product)
Robustness on sparse data	**Yes** (Due to all data being used in estimation; Parsimonious use of forecast parameters)	**No** (Unstable due to updates based on latest obs. to parms. at lowest level)	**No/Yes** Estimators for WTP by market tested were insufficient.
Use live inventory in forecast	**No** Correlation assumed to be zero	**No/Yes** (Live data can be included)	**No/Yes** (Available, but currently not used)

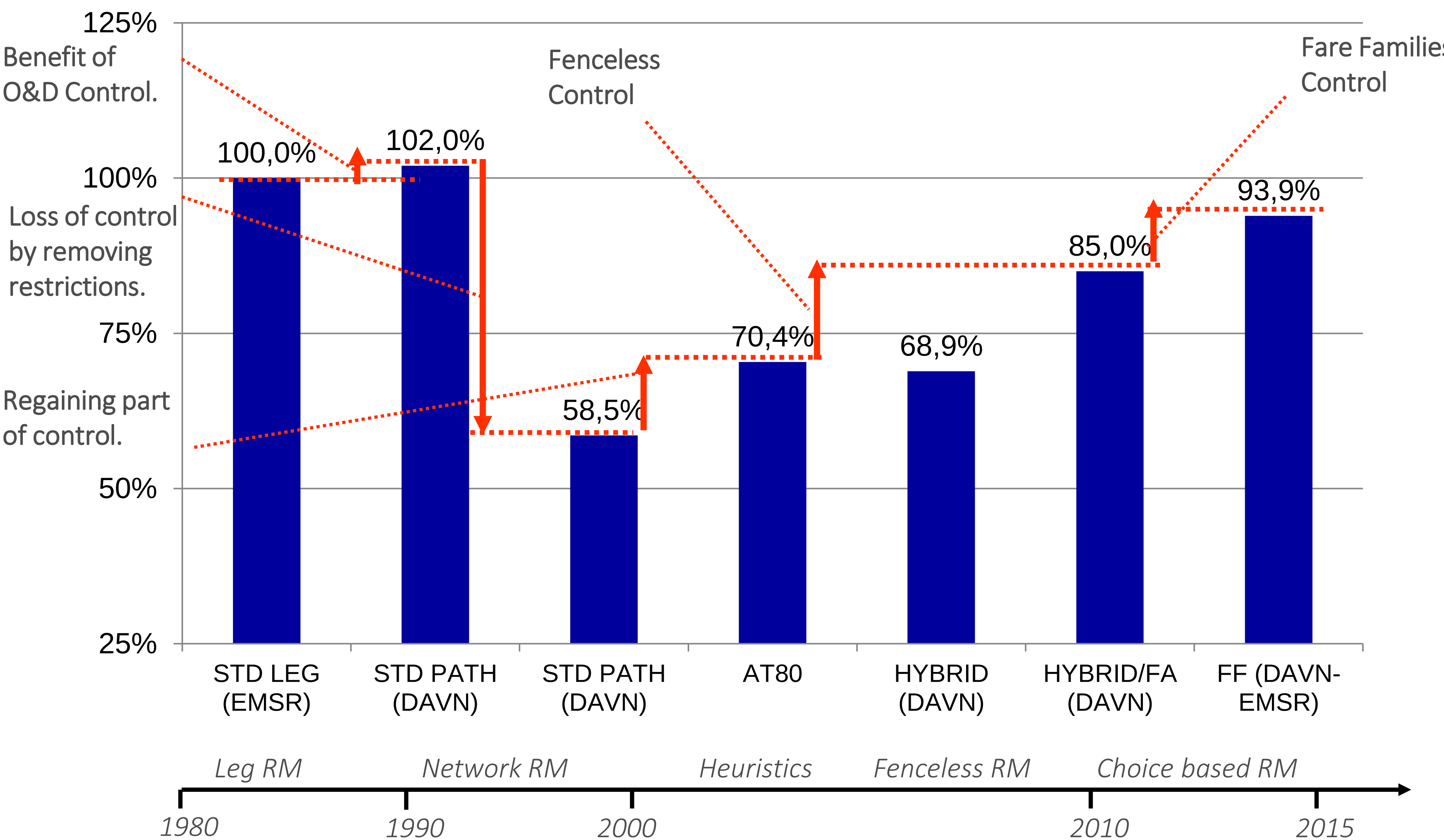
Confidential RESTRICTED

Key scientific elements of forecasting dependent demand

- **Data quality**. The Altéa platform provides high quality of data to support the forecasting and WTP estimation.
- **Joint estimation of all forecast parameters.** Forecaster optimizes forecast accuracy by estimating all forecast parameters (volume, WTP, product preference) in a single process.
- **No unconstraing process**.
- **Robust estimation** by parsimonious use of paramaters.
- **Weekly recalibration**.
- **Daily automatic management of schedule changes**
- **No manual inventory overrides required** to limit the spiral down:
- **Consistent evaluation of marginal revenue in NRM and inventory**: Fare modifiers are automatically computed and applied in both network optimization and in the availability processor.

PODS Benchmark studies*)

Relative revenue performance



*) Cf. Hopperstad Consulting Report Dec 2012.

Confidential
RESTRICTED

Thank you

You can follow us on:
AmadeusITGroup

amadeus.com/blog
amadeus.com

AMADEUS