Sales Forecasting: Time Series Analysis in a Nutshell

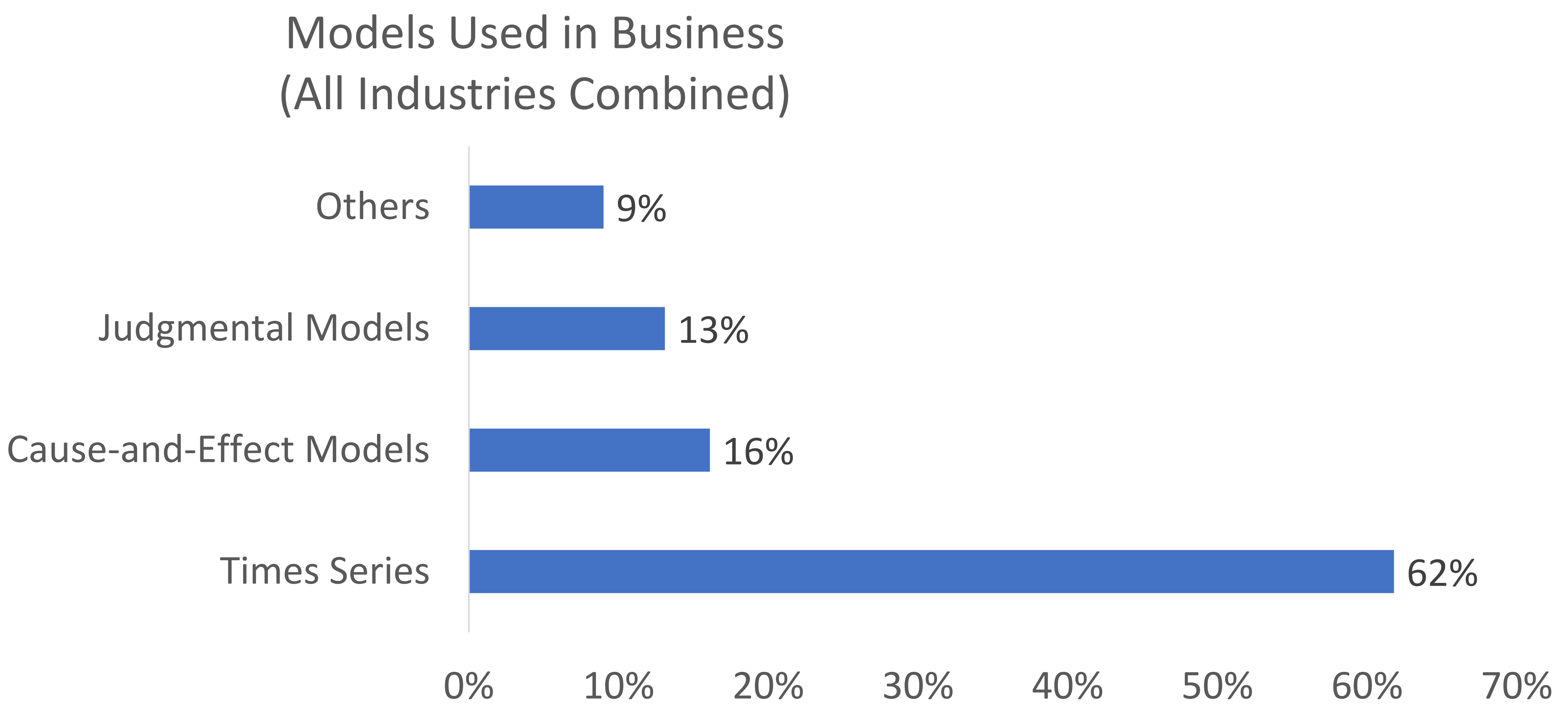
Kannapha Amaruchkul

CONFERENCE 2019 & The 4th BADS

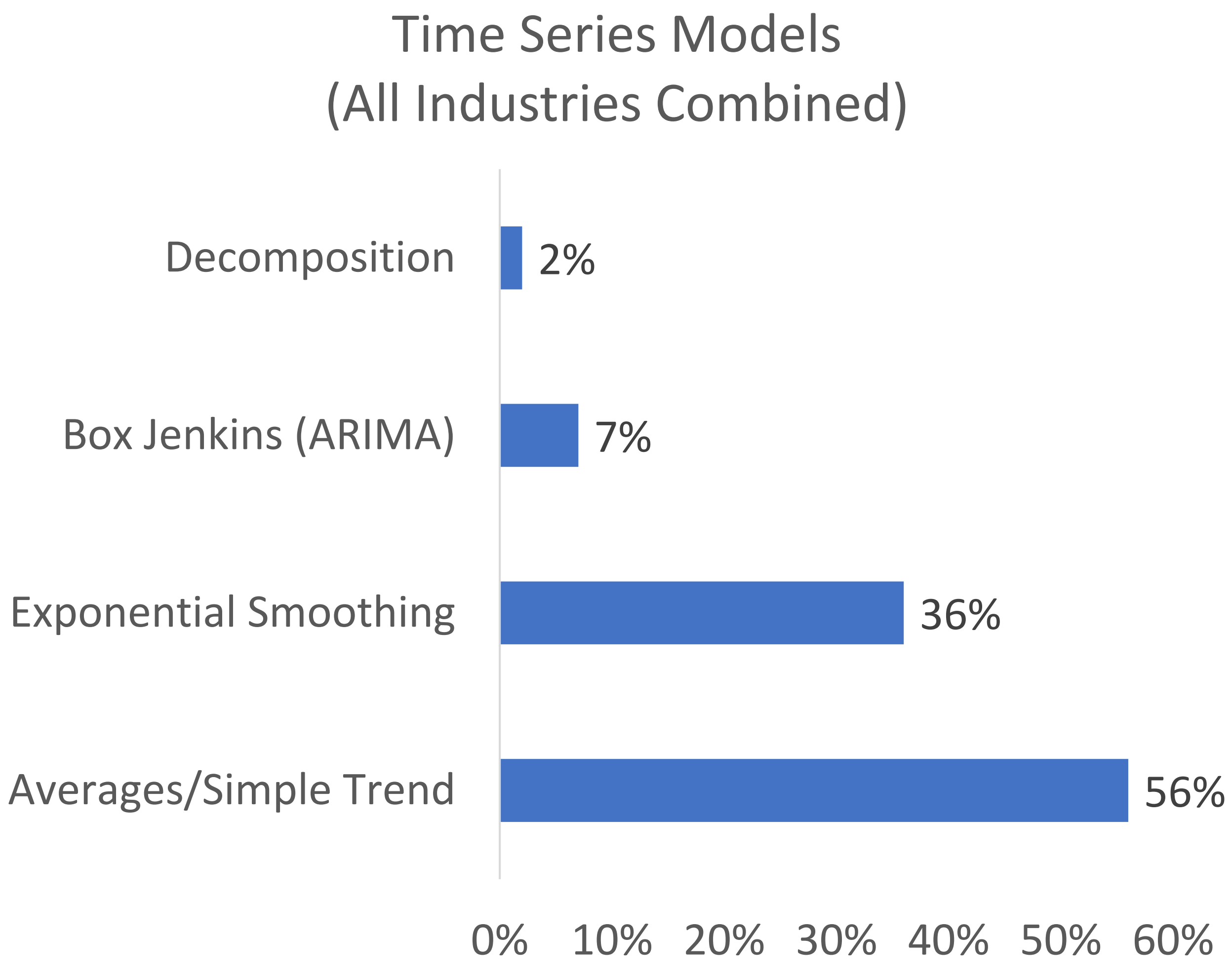
NIDA GSAS
Graduate School of Applied Statistics

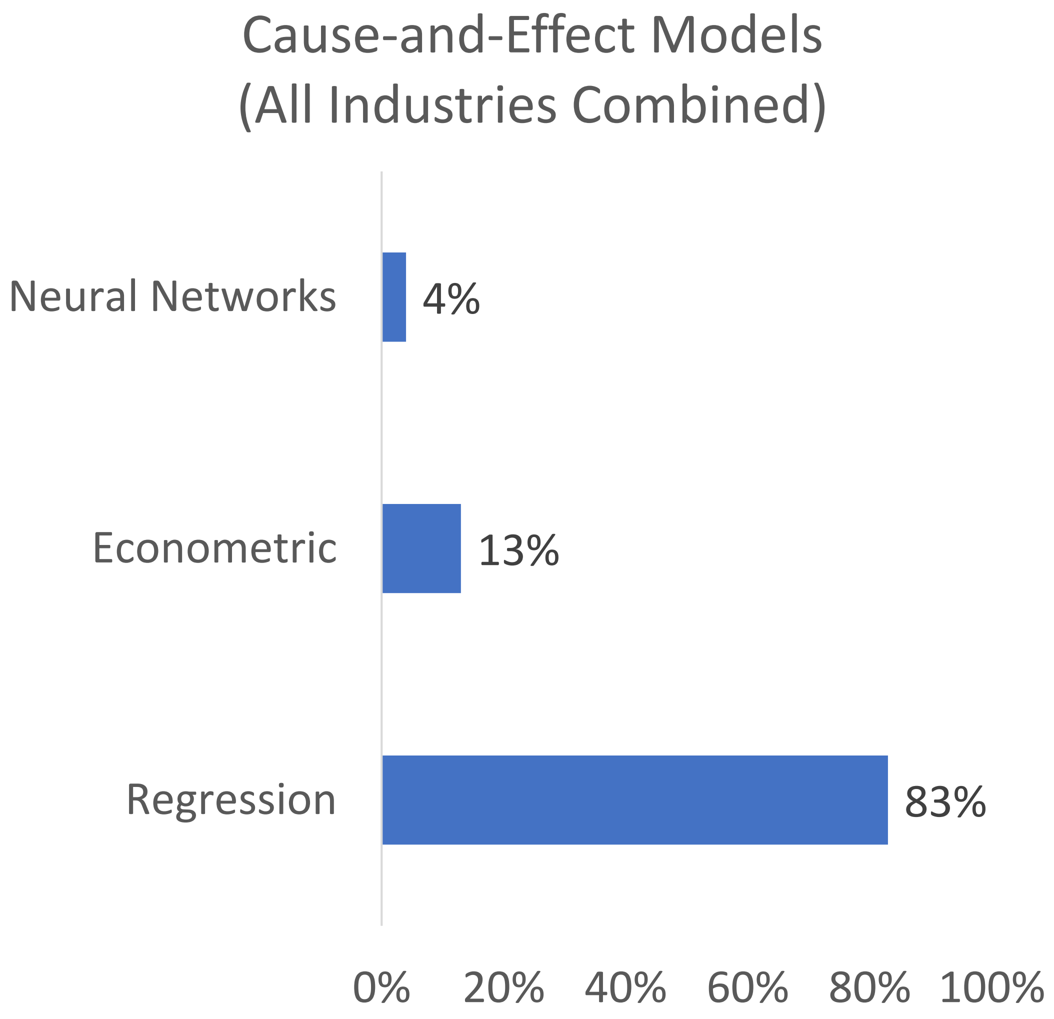
Outline

- Overview of time series models
- Forecasting process
- From demand forecasting (predictive) to inventory optimization (prescriptive)



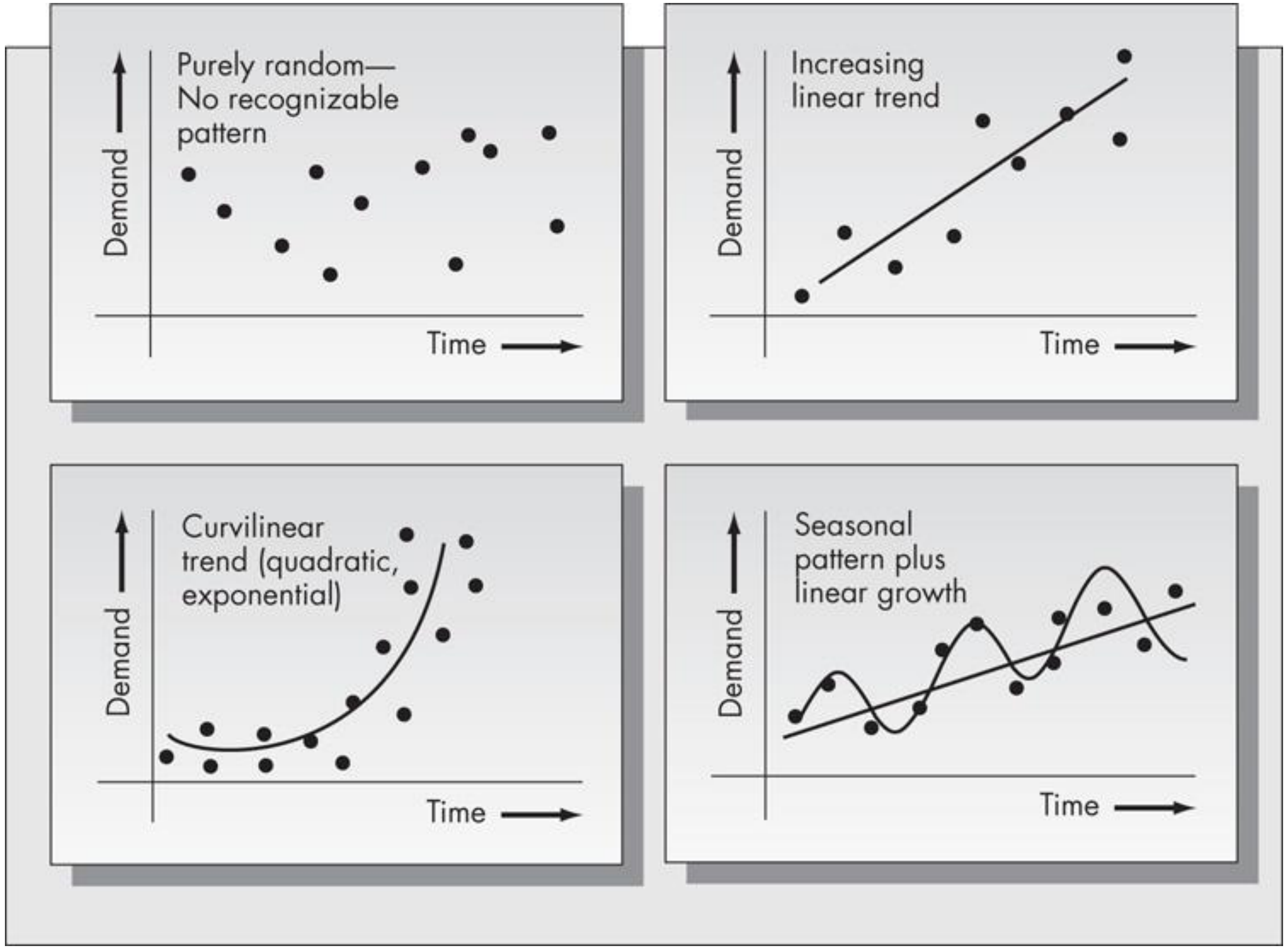
Source: Based on the IBF Survey of 2009





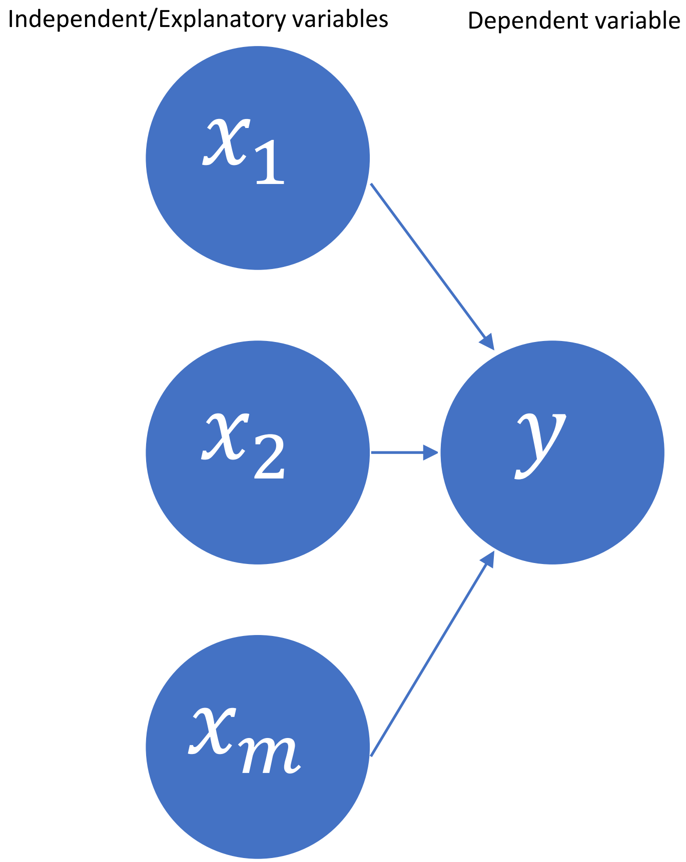
Source: Based on the IBF Survey of 2009

TS Models: Level *vs* Trend *vs* Seasonal

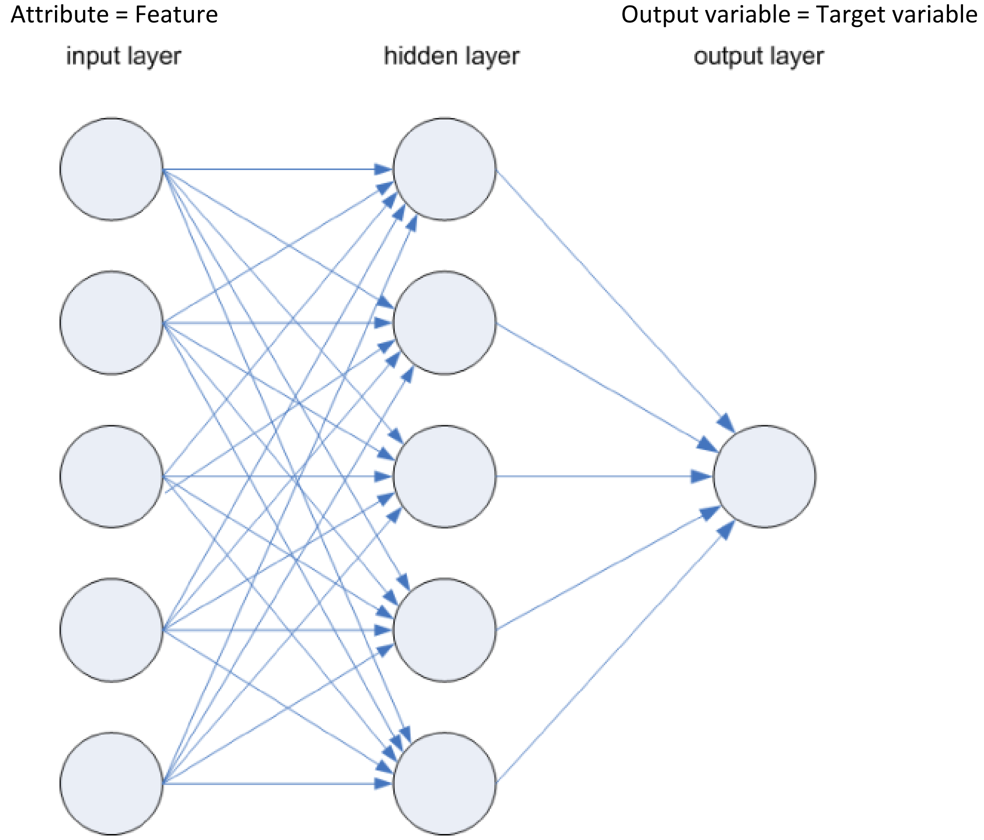


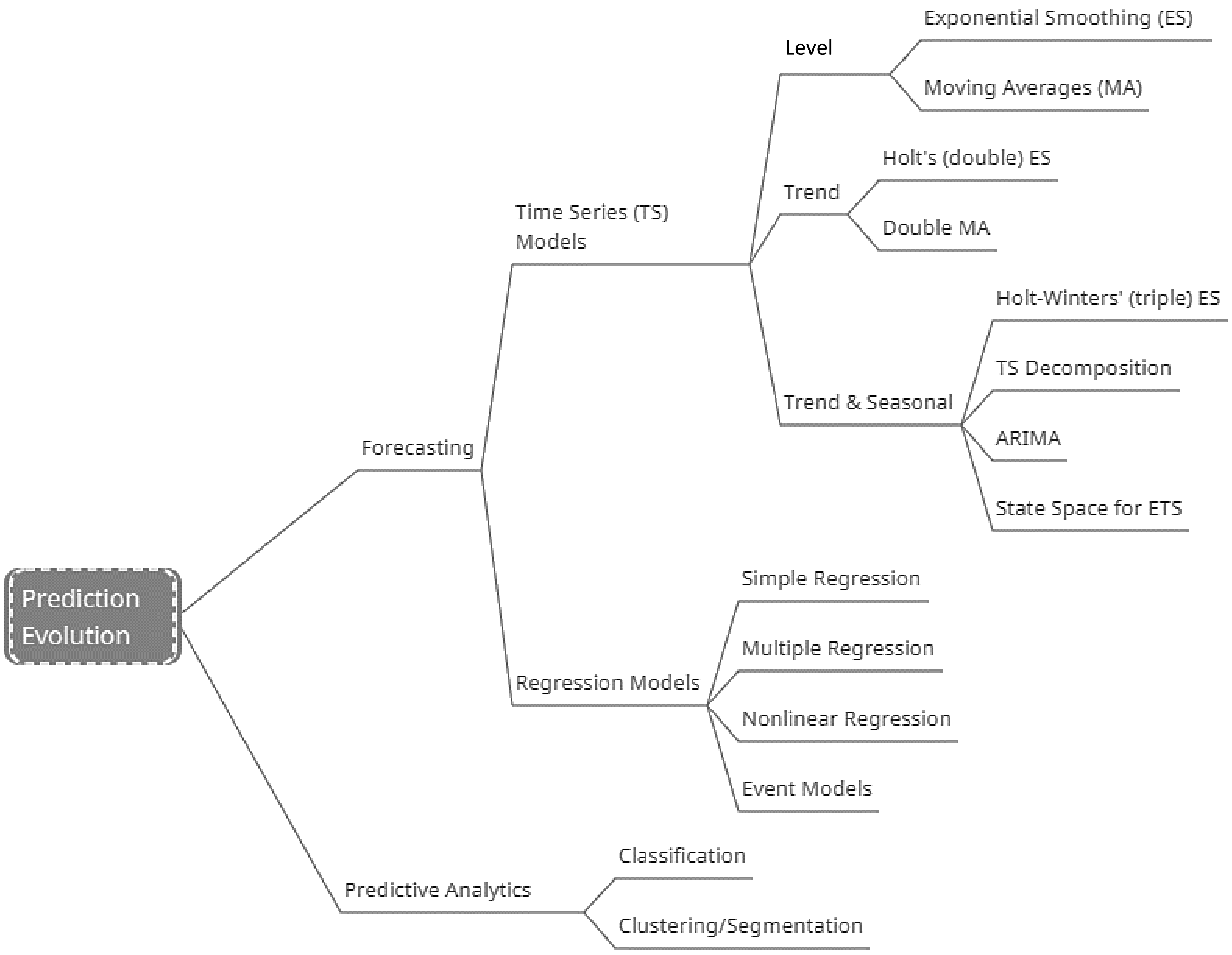
Cause-and-Effect Models

Multiple Regression



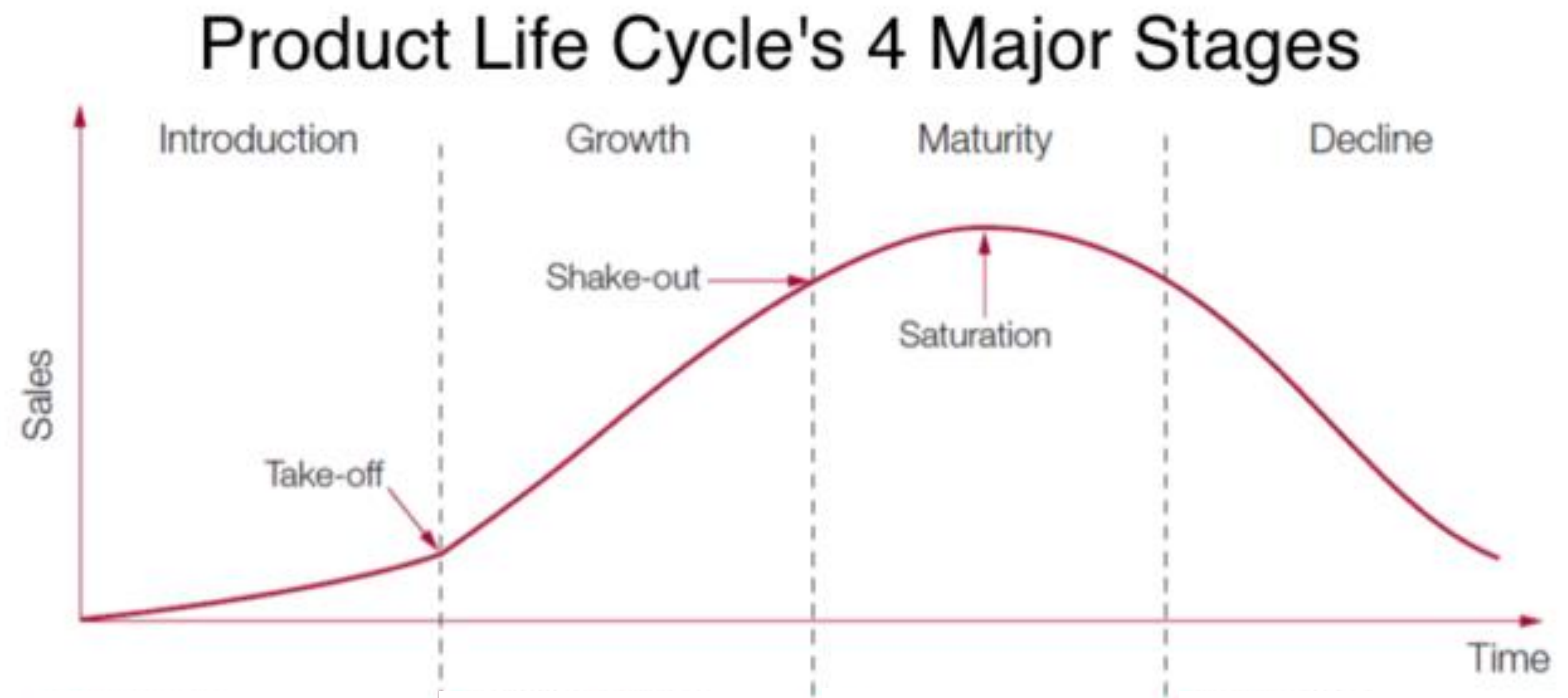
Neural Network




Prediction Evolution
Forecasting
Time Series (TS) Models
Level
Exponential Smoothing (ES)
Moving Averages (MA)
Trend
Holt's (double) ES
Double MA
Trend & Seasonal
Holt-Winters' (triple) ES
TS Decomposition
ARIMA
State Space for ETS
Regression Models
Simple Regression
Multiple Regression
Nonlinear Regression
Event Models
Predictive Analytics
Classification
Clustering/Segmentation

	Author	Year	Development
			ES widely used in business as ad hoc techniques for extrapolating
Smoothing Techniques	Brown	1963	ES received attention from statisticians *Smoothing, Forecasting and Prediction*. NJ: Prentice-Hall (1963)
	Holt	1957	Extended SES to include trend → Holt's method
	Winters	1960	Extended SES to include both trend and seasonality → Holt-Winters' model
State space model for ExponenTial Smoothing (ETS)	Pegels	1969	Provided simple classification of trend and seasonality patterns, depending on whether they are additive or multiplicative
	Gardner	1985	Extended Pegels' classification to include damped trend
	Synder	1985	Showed that SES could be considered as arising from innovation state space model (i.e., a model with a single source of error)
ARIMA	Yule	1927	Formulated autoregressive (AR) and moving average (MA) (Postulating that every time series can be regarded as realization of stochastic process)
	Box & Jenkins	1976	Developed the three-stage iterative cycle for identification, estimation, and verification (rightly known as Box-Jenkins approach) *Time Series Analysis: Forecasting and Control*, 2nd ed. SF: Holden-Day (1976)

Source: Gooijer and Hyndman (2006). 25 Years of Time Series Forecasting.



Introduction

- Qualitative methods (e.g., sales force composite, Delphi)
- Input-output analysis
- Historical analysis of comparable products

Growth

- Qualitative methods (e.g., consumer survey)
- Product growth model (e.g., Gompertz, logistic curves, Bass model)

Maturity

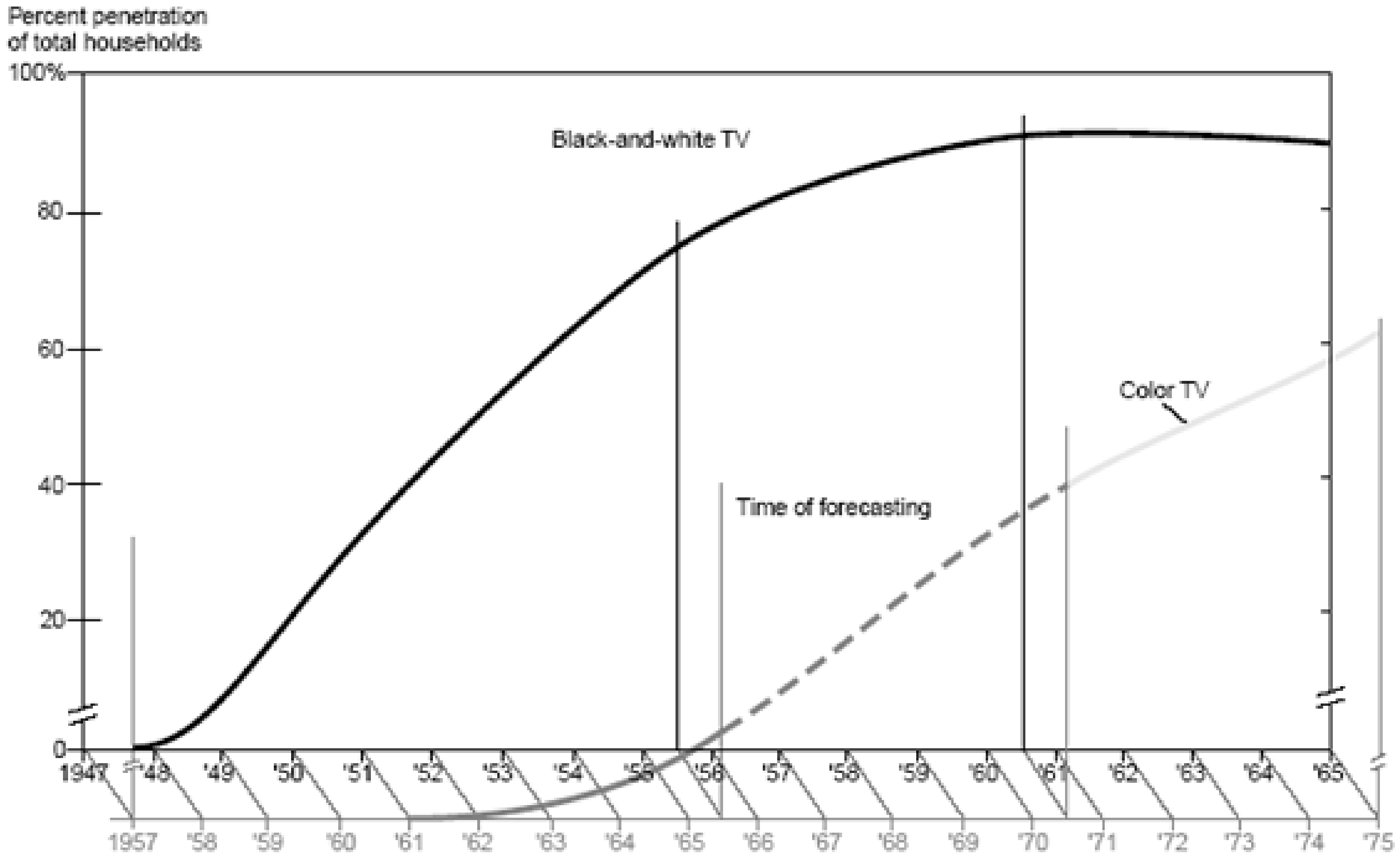
- Smoothing techniques
- Regression models

Decline

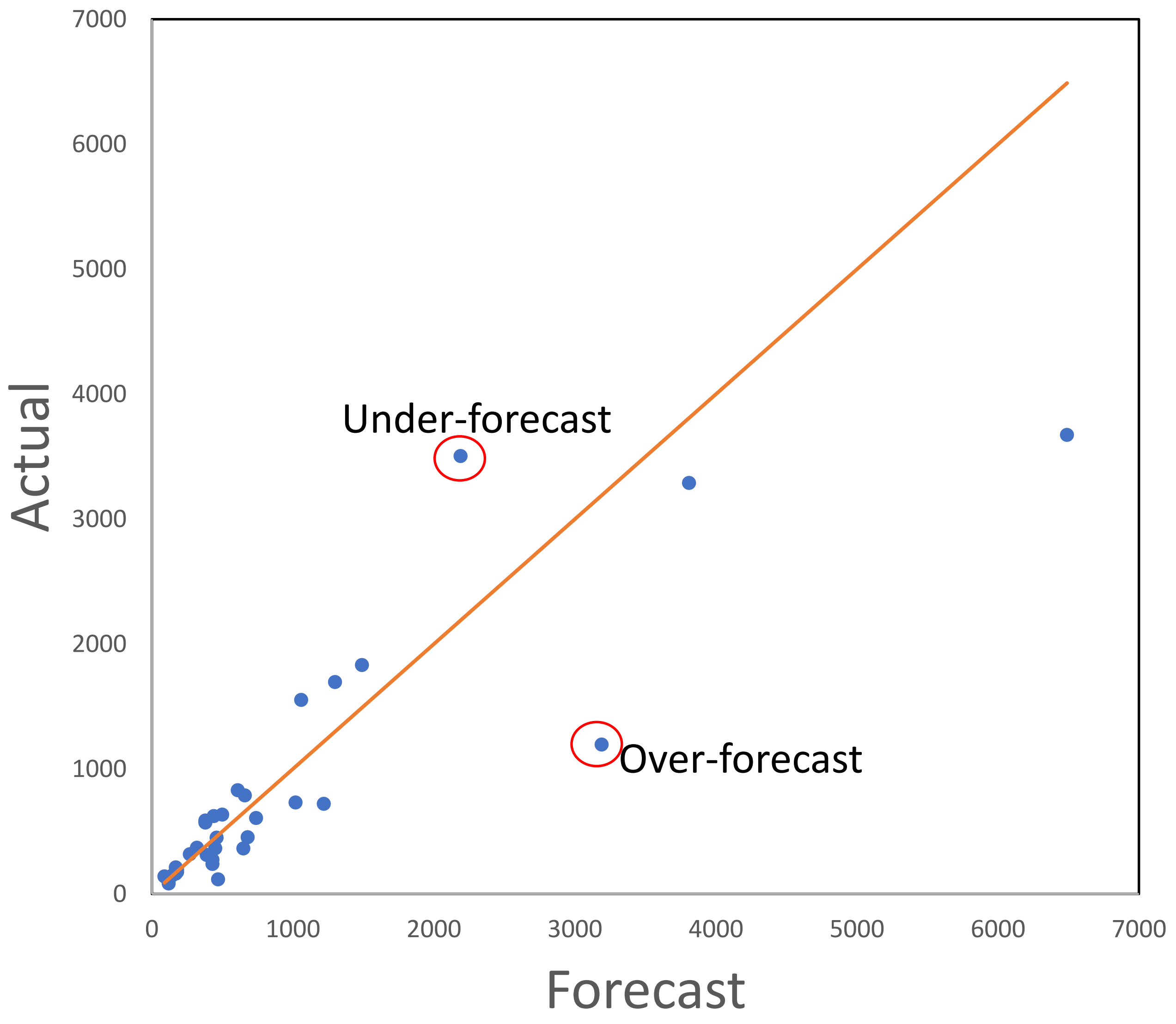
- TS analysis & projection
- Causal models

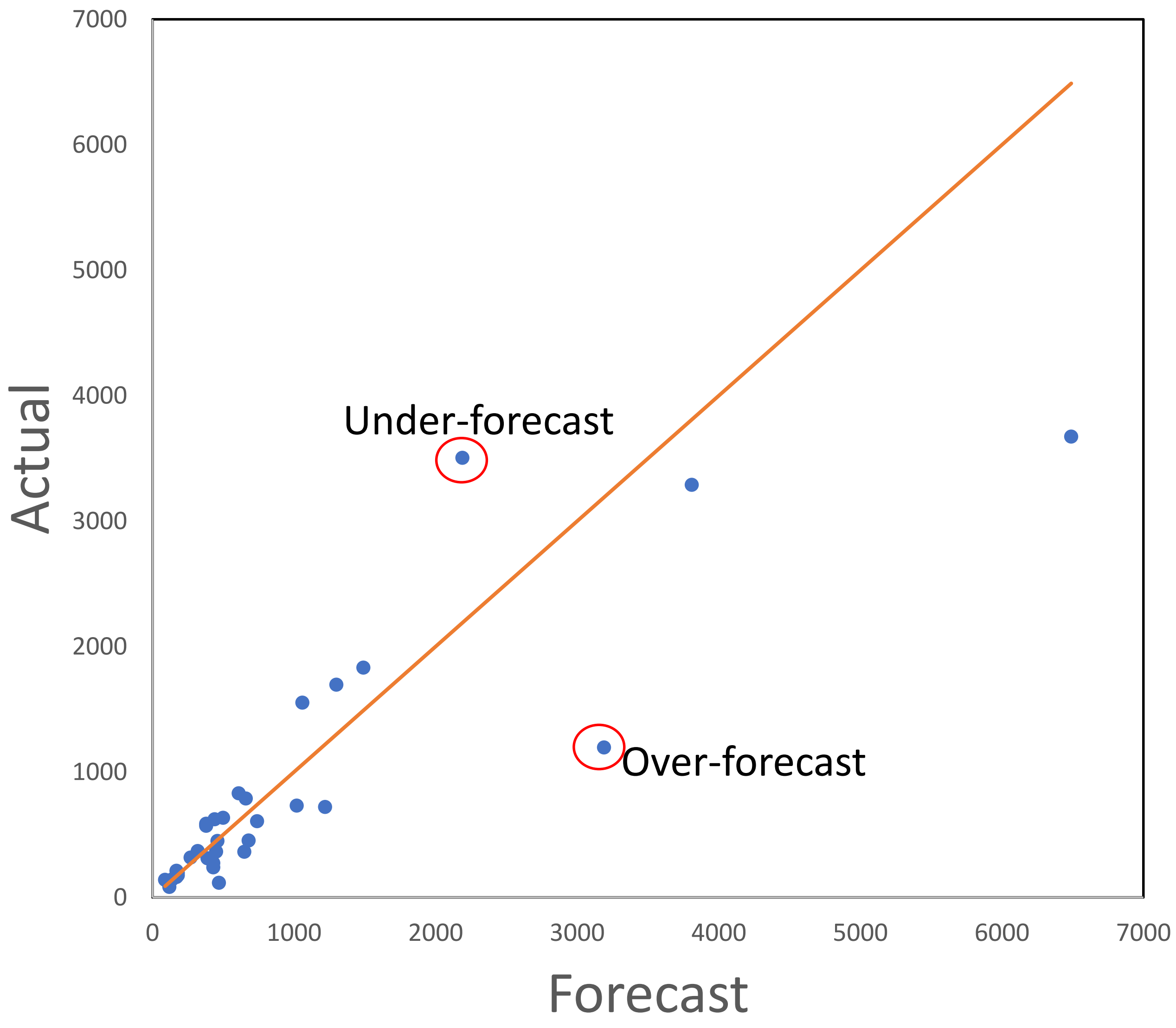
In 1965, we disaggregated the market for color television by income levels and geographical regions and compared these submarkets with the historical pattern of black-and-white TV market growth.

Exhibit V Long-term Household Penetration Curves for Color and Black-and-White TV

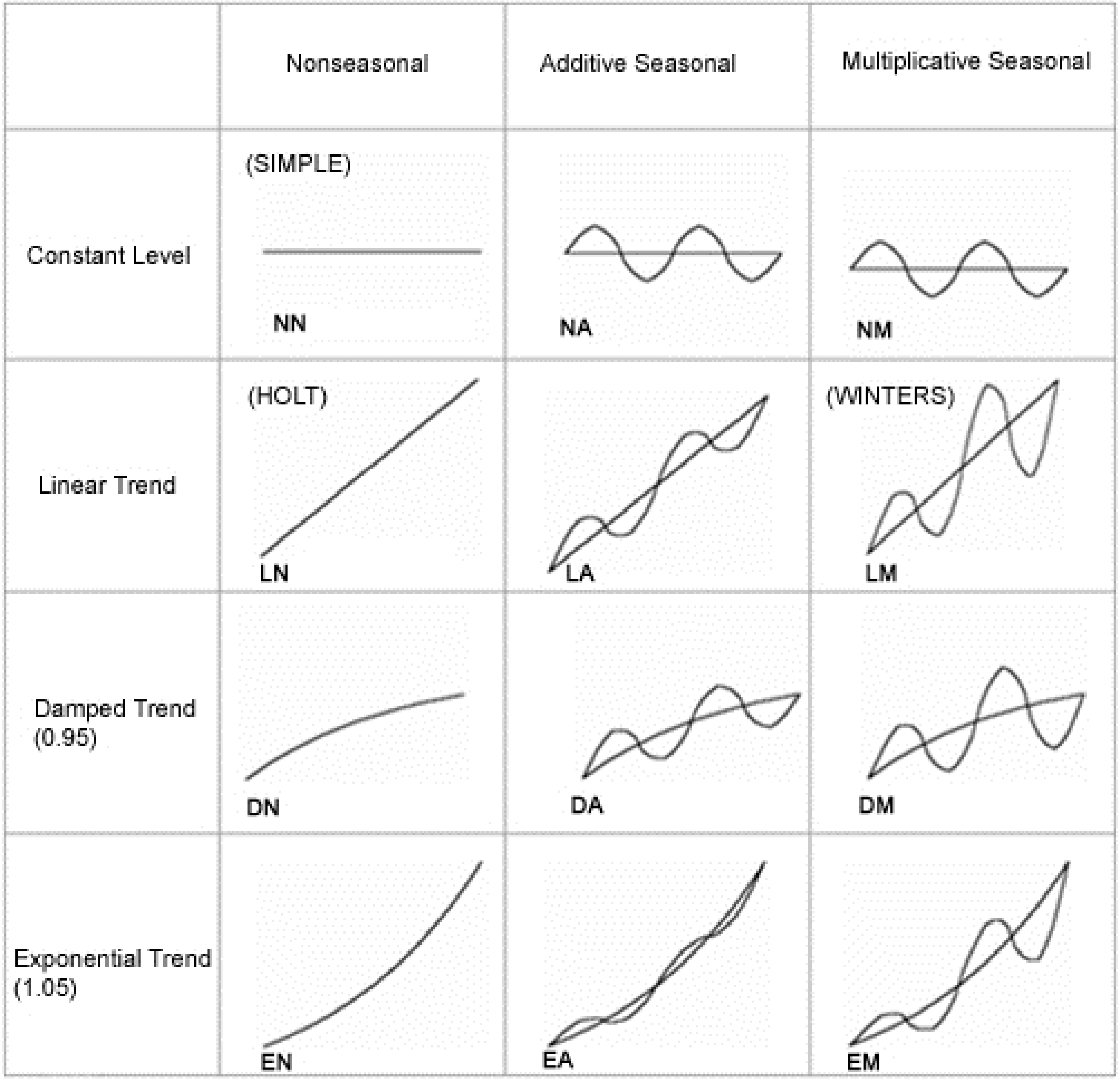


Source: J.C. Chambers, S.K. Mullick, and D.D. Smith (1971). How to choose the right forecasting techniques. Harvard Business Review. https://hbr.org/1971/07/how-to-choose-the-right-forecasting-technique


7000
6000
5000
4000
3000
2000
1000
0
Actual
Under-forecast
Over-forecast
0
1000
2000
3000
4000
5000
6000
7000
Forecast



"Forecasts usually tell us more of the ***forecaster*** *than of the future."* — Warren Buffett



Source: Gardner (1987)

Table 7.7: State space equations for each of the models in the ETS framework.

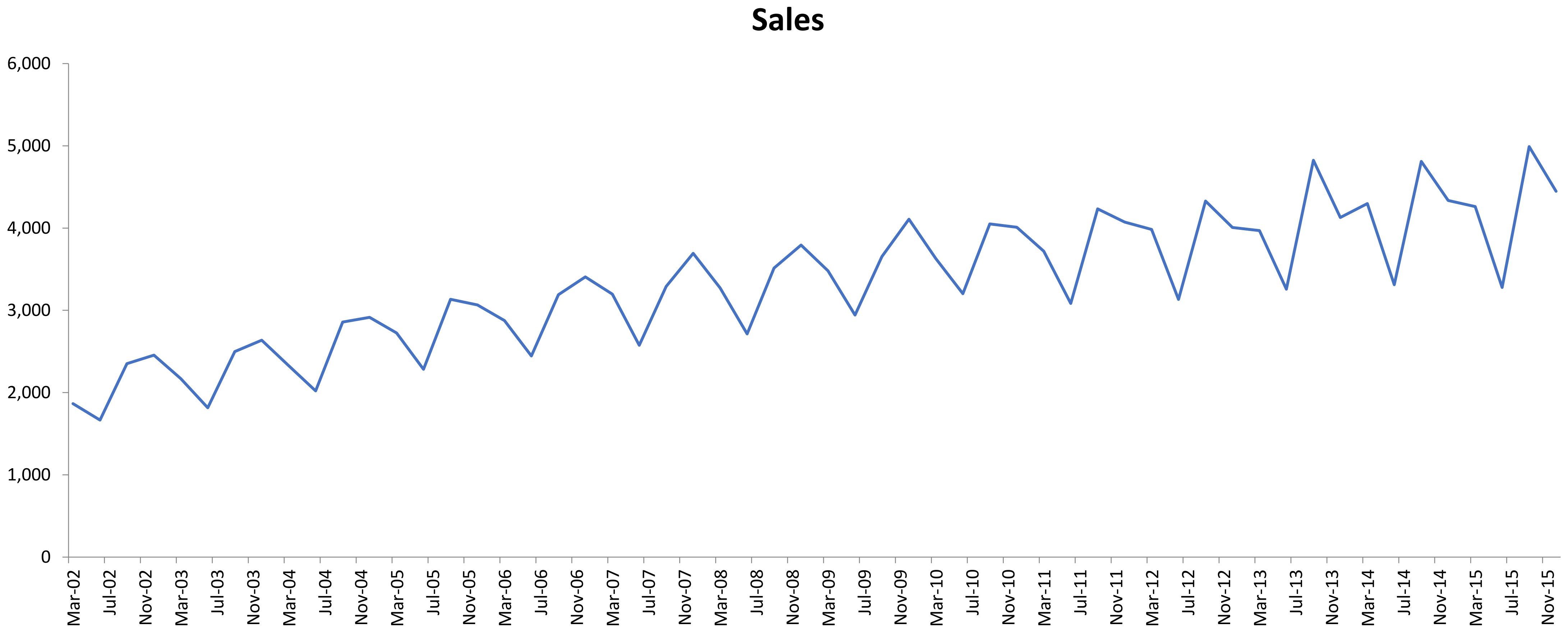
ADDITIVE ERROR MODELS

Trend	Seasonal N	Seasonal A	Seasonal M
N	$y_t = \ell_{t-1} + \varepsilon_t$ $\ell_t = \ell_{t-1} + \alpha\varepsilon_t$	$y_t = \ell_{t-1} + s_{t-m} + \varepsilon_t$ $\ell_t = \ell_{t-1} + \alpha\varepsilon_t$ $s_t = s_{t-m} + \gamma\varepsilon_t$	$y_t = \ell_{t-1}s_{t-m} + \varepsilon_t$ $\ell_t = \ell_{t-1} + \alpha\varepsilon_t/s_{t-m}$ $s_t = s_{t-m} + \gamma\varepsilon_t/\ell_{t-1}$
A	$y_t = \ell_{t-1} + b_{t-1} + \varepsilon_t$ $\ell_t = \ell_{t-1} + b_{t-1} + \alpha\varepsilon_t$ $b_t = b_{t-1} + \beta\varepsilon_t$	$y_t = \ell_{t-1} + b_{t-1} + s_{t-m} + \varepsilon_t$ $\ell_t = \ell_{t-1} + b_{t-1} + \alpha\varepsilon_t$ $b_t = b_{t-1} + \beta\varepsilon_t$ $s_t = s_{t-m} + \gamma\varepsilon_t$	$y_t = (\ell_{t-1} + b_{t-1})s_{t-m} + \varepsilon_t$ $\ell_t = \ell_{t-1} + b_{t-1} + \alpha\varepsilon_t/s_{t-m}$ $b_t = b_{t-1} + \beta\varepsilon_t/s_{t-m}$ $s_t = s_{t-m} + \gamma\varepsilon_t/(\ell_{t-1} + b_{t-1})$
A$_d$	$y_t = \ell_{t-1} + \phi b_{t-1} + \varepsilon_t$ $\ell_t = \ell_{t-1} + \phi b_{t-1} + \alpha\varepsilon_t$ $b_t = \phi b_{t-1} + \beta\varepsilon_t$	$y_t = \ell_{t-1} + \phi b_{t-1} + s_{t-m} + \varepsilon_t$ $\ell_t = \ell_{t-1} + \phi b_{t-1} + \alpha\varepsilon_t$ $b_t = \phi b_{t-1} + \beta\varepsilon_t$ $s_t = s_{t-m} + \gamma\varepsilon_t$	$y_t = (\ell_{t-1} + \phi b_{t-1})s_{t-m} + \varepsilon_t$ $\ell_t = \ell_{t-1} + \phi b_{t-1} + \alpha\varepsilon_t/s_{t-m}$ $b_t = \phi b_{t-1} + \beta\varepsilon_t/s_{t-m}$ $s_t = s_{t-m} + \gamma\varepsilon_t/(\ell_{t-1} + \phi b_{t-1})$

MULTIPLICATIVE ERROR MODELS

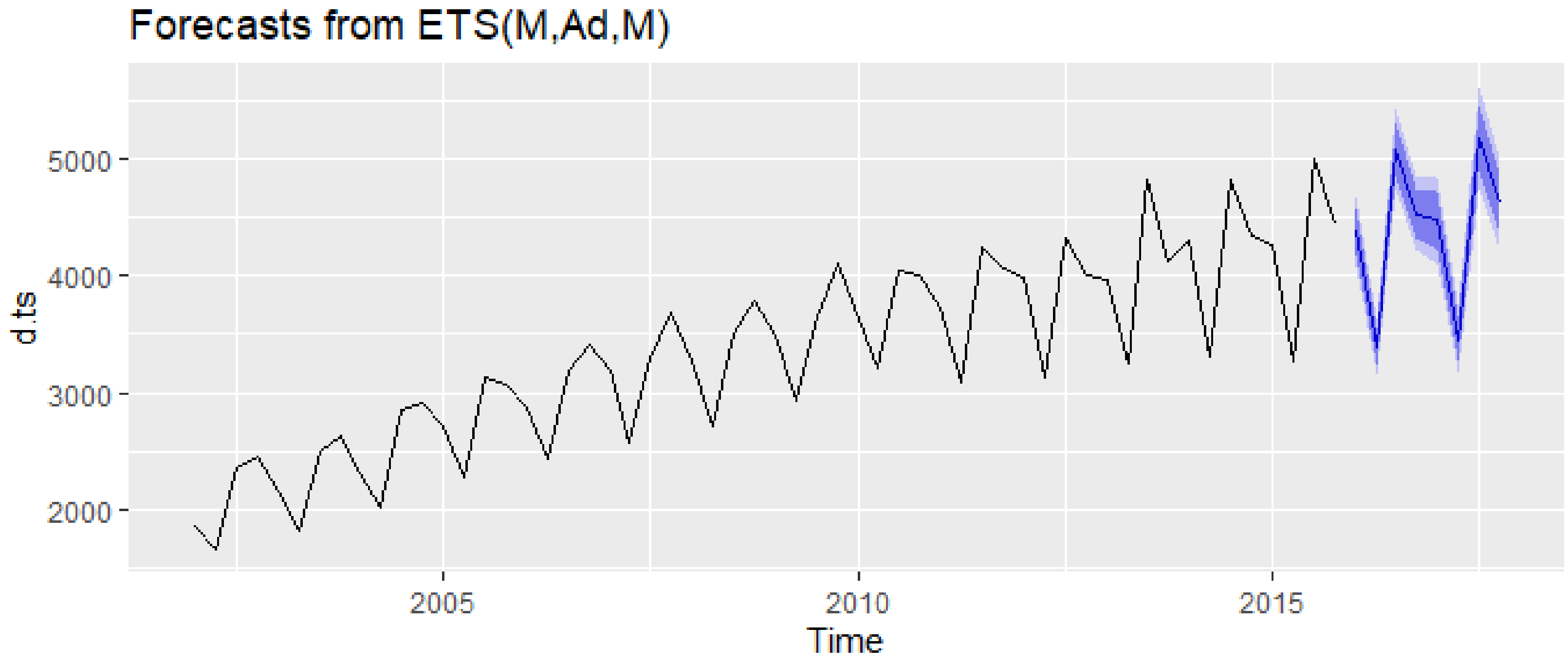
Trend	Seasonal N	Seasonal A	Seasonal M
N	$y_t = \ell_{t-1}(1 + \varepsilon_t)$ $\ell_t = \ell_{t-1}(1 + \alpha\varepsilon_t)$	$y_t = (\ell_{t-1} + s_{t-m})(1 + \varepsilon_t)$ $\ell_t = \ell_{t-1} + \alpha(\ell_{t-1} + s_{t-m})\varepsilon_t$ $s_t = s_{t-m} + \gamma(\ell_{t-1} + s_{t-m})\varepsilon_t$	$y_t = \ell_{t-1}s_{t-m}(1 + \varepsilon_t)$ $\ell_t = \ell_{t-1}(1 + \alpha\varepsilon_t)$ $s_t = s_{t-m}(1 + \gamma\varepsilon_t)$
A	$y_t = (\ell_{t-1} + b_{t-1})(1 + \varepsilon_t)$ $\ell_t = (\ell_{t-1} + b_{t-1})(1 + \alpha\varepsilon_t)$ $b_t = b_{t-1} + \beta(\ell_{t-1} + b_{t-1})\varepsilon_t$	$y_t = (\ell_{t-1} + b_{t-1} + s_{t-m})(1 + \varepsilon_t)$ $\ell_t = \ell_{t-1} + b_{t-1} + \alpha(\ell_{t-1} + b_{t-1} + s_{t-m})\varepsilon_t$ $b_t = b_{t-1} + \beta(\ell_{t-1} + b_{t-1} + s_{t-m})\varepsilon_t$ $s_t = s_{t-m} + \gamma(\ell_{t-1} + b_{t-1} + s_{t-m})\varepsilon_t$	$y_t = (\ell_{t-1} + b_{t-1})s_{t-m}(1 + \varepsilon_t)$ $\ell_t = (\ell_{t-1} + b_{t-1})(1 + \alpha\varepsilon_t)$ $b_t = b_{t-1} + \beta(\ell_{t-1} + b_{t-1})\varepsilon_t$ $s_t = s_{t-m}(1 + \gamma\varepsilon_t)$
A$_d$	$y_t = (\ell_{t-1} + \phi b_{t-1})(1 + \varepsilon_t)$ $\ell_t = (\ell_{t-1} + \phi b_{t-1})(1 + \alpha\varepsilon_t)$ $b_t = \phi b_{t-1} + \beta(\ell_{t-1} + \phi b_{t-1})\varepsilon_t$	$y_t = (\ell_{t-1} + \phi b_{t-1} + s_{t-m})(1 + \varepsilon_t)$ $\ell_t = \ell_{t-1} + \phi b_{t-1} + \alpha(\ell_{t-1} + \phi b_{t-1} + s_{t-m})\varepsilon_t$ $b_t = \phi b_{t-1} + \beta(\ell_{t-1} + \phi b_{t-1} + s_{t-m})\varepsilon_t$ $s_t = s_{t-m} + \gamma(\ell_{t-1} + \phi b_{t-1} + s_{t-m})\varepsilon_t$	$y_t = (\ell_{t-1} + \phi b_{t-1})s_{t-m}(1 + \varepsilon_t)$ $\ell_t = (\ell_{t-1} + \phi b_{t-1})(1 + \alpha\varepsilon_t)$ $b_t = \phi b_{t-1} + \beta(\ell_{t-1} + \phi b_{t-1})\varepsilon_t$ $s_t = s_{t-m}(1 + \gamma\varepsilon_t)$

ETS
Error=(M,A), Trend=(N,A, Ad), Season= (N,A,M)?



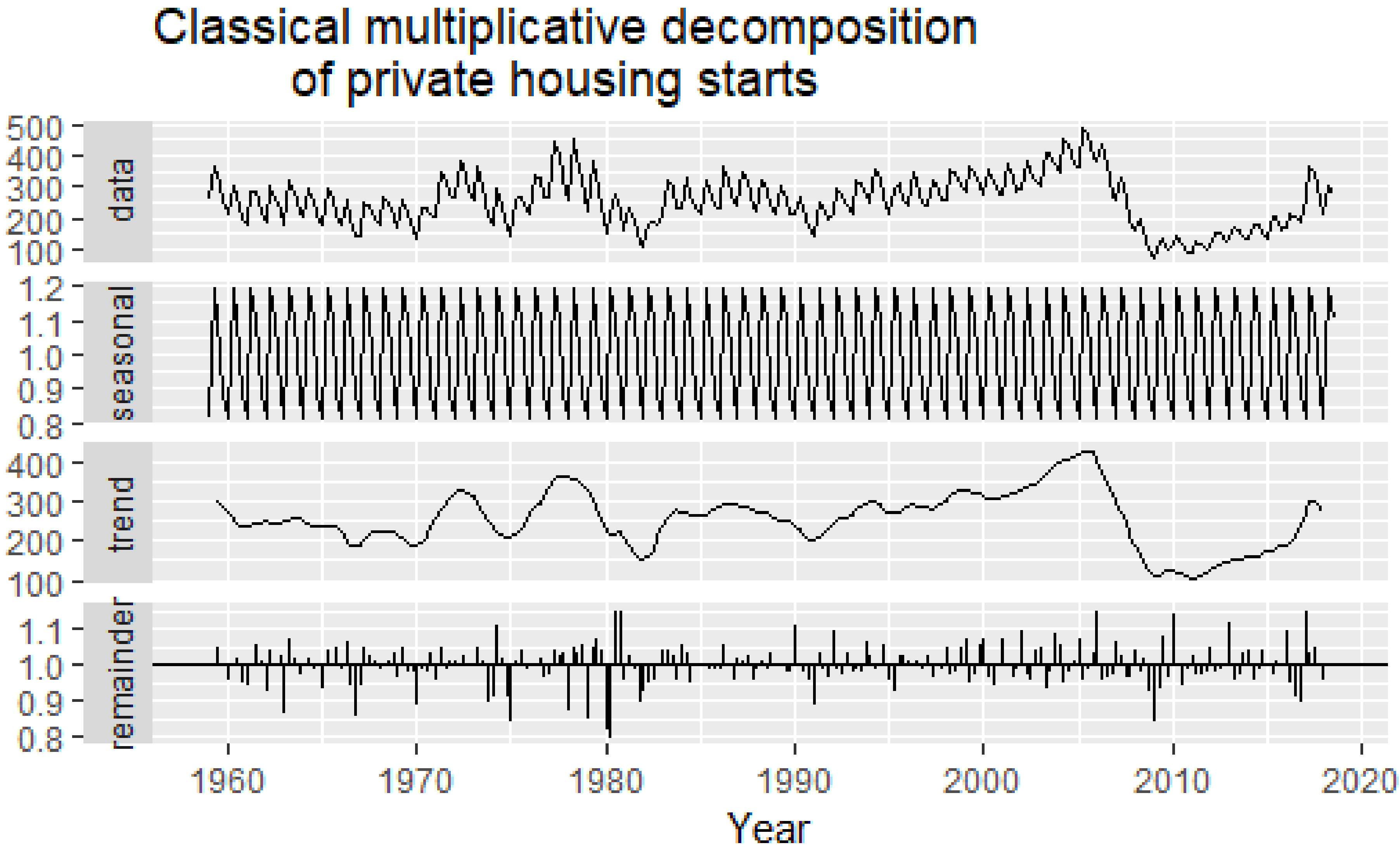
```
> library(forecast)
> d.ts <- ts(d$Sales, frequency = 4, start=c(2002,1), end=c(2015,4))
> fit1 <- forecast(d.ts)
> summary(fit1)
```

```
Forecast method: ETS(M,Ad,M)

Smoothing parameters:
    alpha = 1e-04
    beta  = 1e-04
    gamma = 0.6923
    phi   = 0.98

  AIC     AICc      BIC
764.9110 769.7999 785.1645


Error measures:
RMSE      MAE  MAPE
113.6239 85.76882 2.514105
```



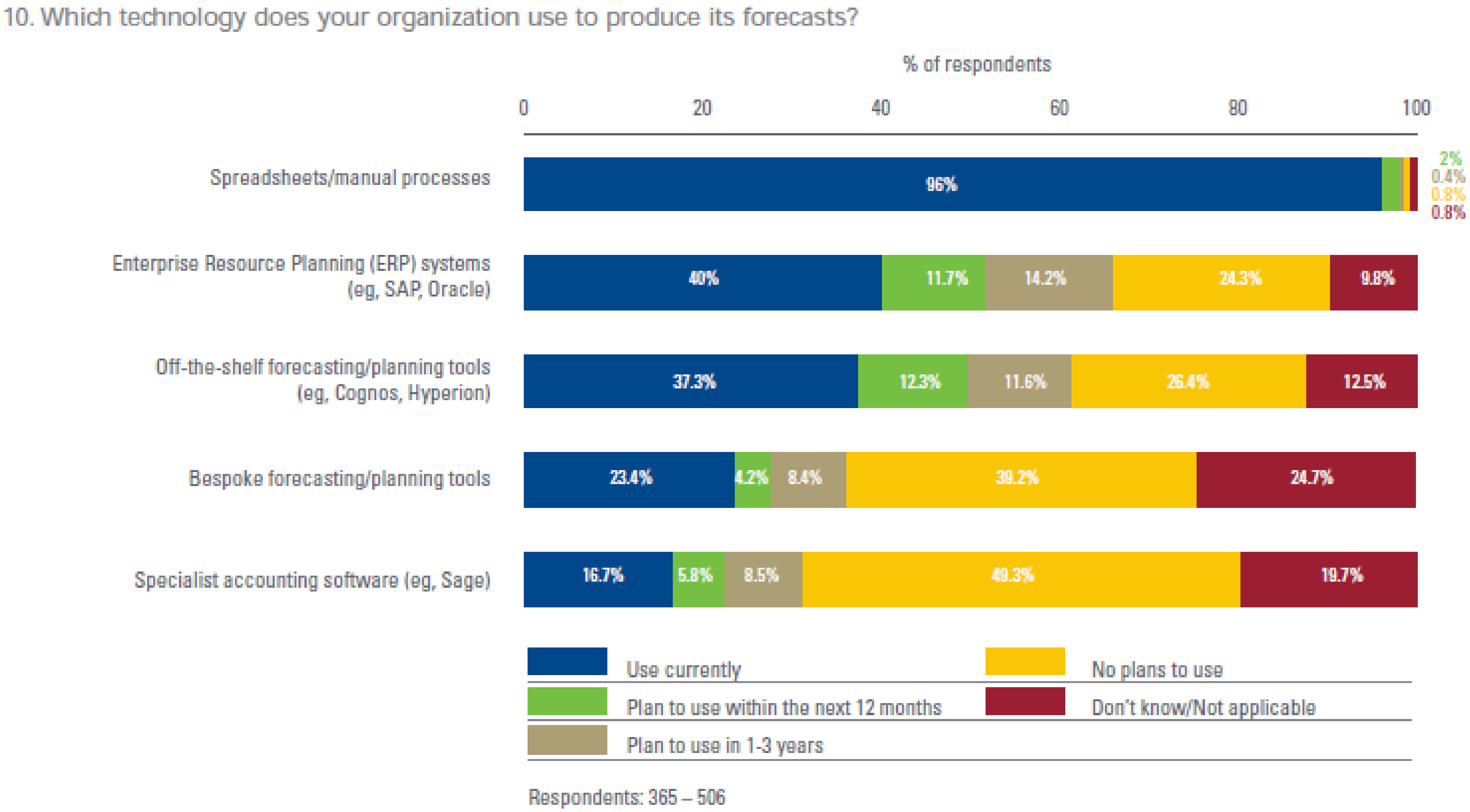
	Point Forecast	Lo 80	Hi 80	Lo 95	Hi 95
2016 Q1	4381.699	4185.239	4578.160	4081.239	4682.160
2016 Q2	3388.180	3236.265	3540.094	3155.847	3620.513
2016 Q3	5078.197	4850.508	5305.886	4729.976	5426.417
2016 Q4	4527.844	4324.831	4730.857	4217.362	4838.326

Quarter	Seasonal Index (SI)
1	0.82
2	1.19
3	1.11
4	0.89

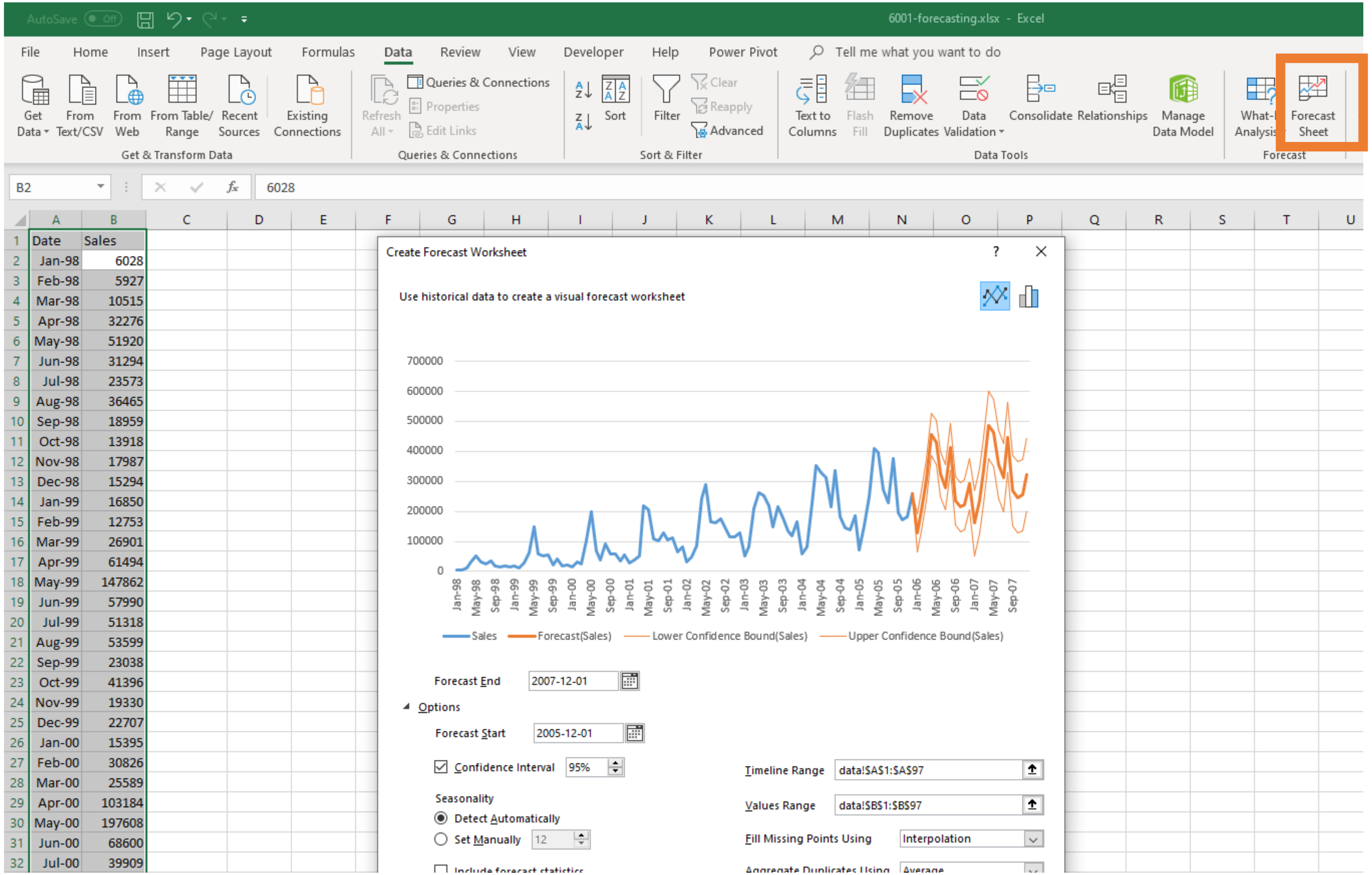


Other decomposition techniques: X11, Seasonal Extraction in ARIMA Time Series (SEATS), Seasonal and Trend decomposition using Loess (STL)

Overview of forecasting models



Source: KPMG (2007). Forecasting with confidence: Insights from leading finance functions



Statistic	Value
Alpha	0.5010
Beta	0.0010
Gamma	0.0010
MASE	0.9148
SMAPE	0.05
MAE	2.26
RMSE	2.82

`FORECAST.ETS` uses ETS(**E**rror=A, **T**rend=A, **S**eason=A), Additive Holt-Winters' method with additive errors

`FORECAST.ETS.CONFINT`

`FORECAST.ETS.STAT`

Forecast Process

Problem definition

Data collection

Preliminary (exploratory) analysis

Model selection

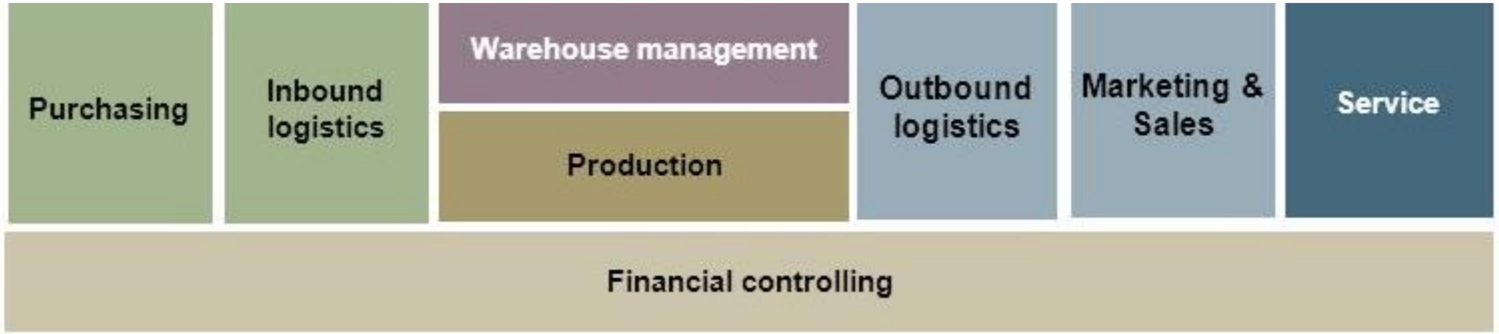
Model evaluation

Tracking results

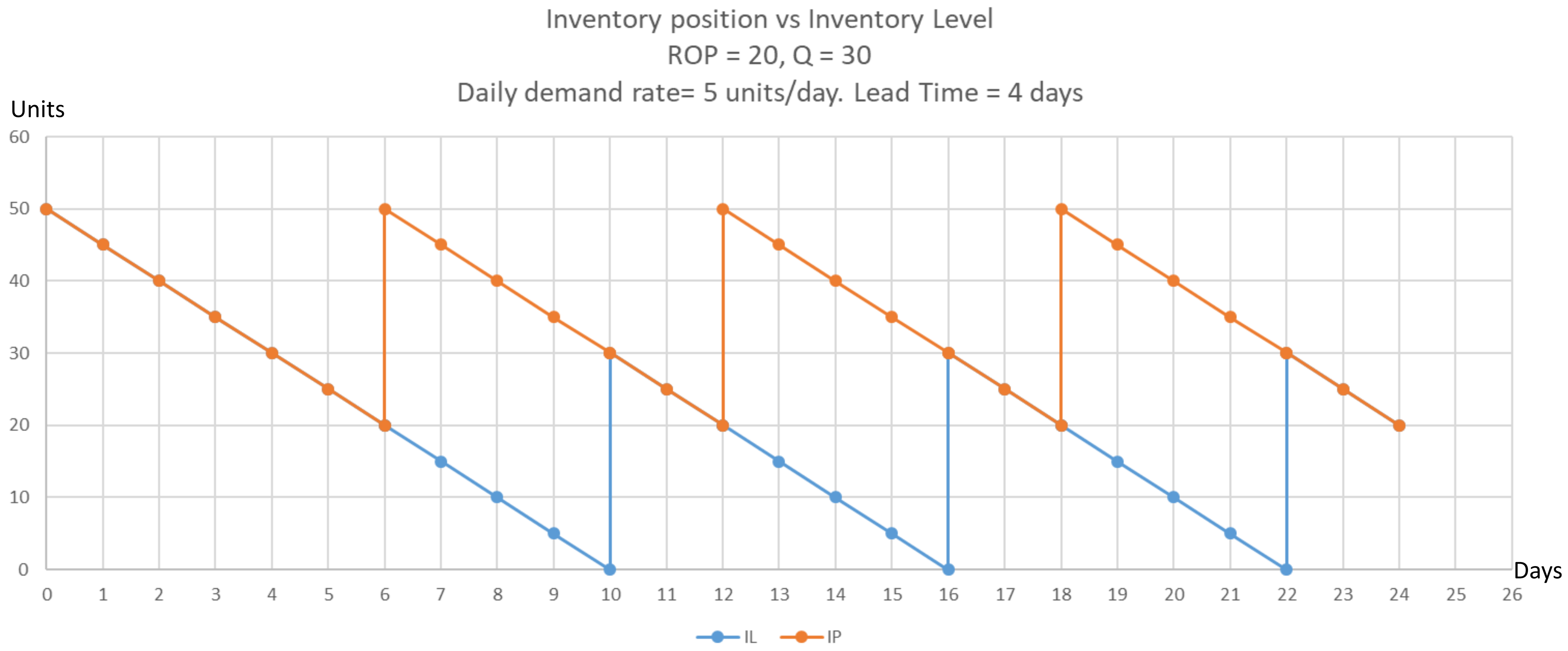
What is the purpose of forecast?
How is it to be used?

- Forecast requirements
  - Aggregate forecasts in dollar by month
  - Forecasts by plant, by month, in units, and at SKU level
  - Demand by category/brand, channel of distribution, region, market share of different categories/brands

- Forecast users



- Strategic decisions
- Operational decisions: Forecast horizon depends on lead time.



Censored/Truncated/Constrained Demand

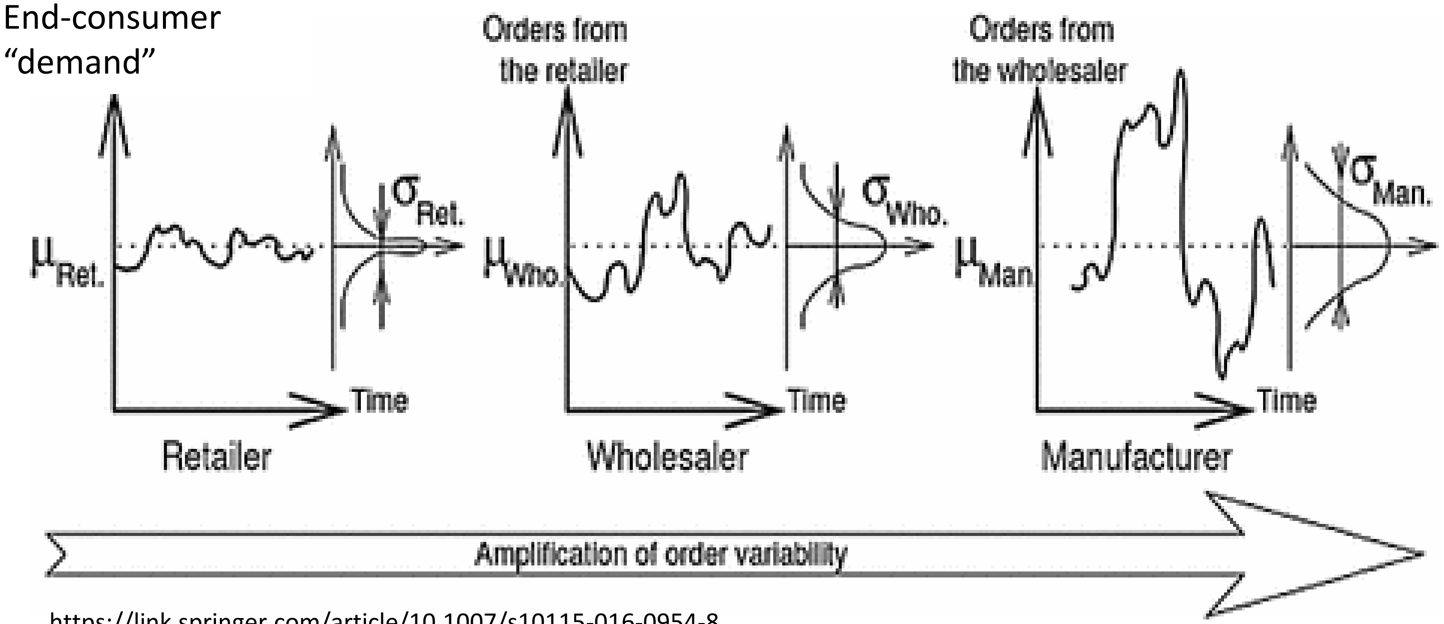
- Our goal is often to forecast *unconstrained* demand, but we only observe *constrained* demand.
- As an example, consider an airline with a 100-seat airplane, flying from AAA to BBB daily.
- If you computed the sample mean and sample variance of these numbers, they would ____________ (underestimate or overestimate) the true mean and variance of demand.

Days	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	Mean	Stdev
Passengers	70	50	100	100	100	80	30	60	100	100	90	50	100	100	40	78	26
Demand	70	50	?	?	?	80	30	60	?	?	90	50	?	?	40		

POS Reduces Bullwhip Effects

- Point of Sales (POS) *vs* Customer Order *vs* Shipment

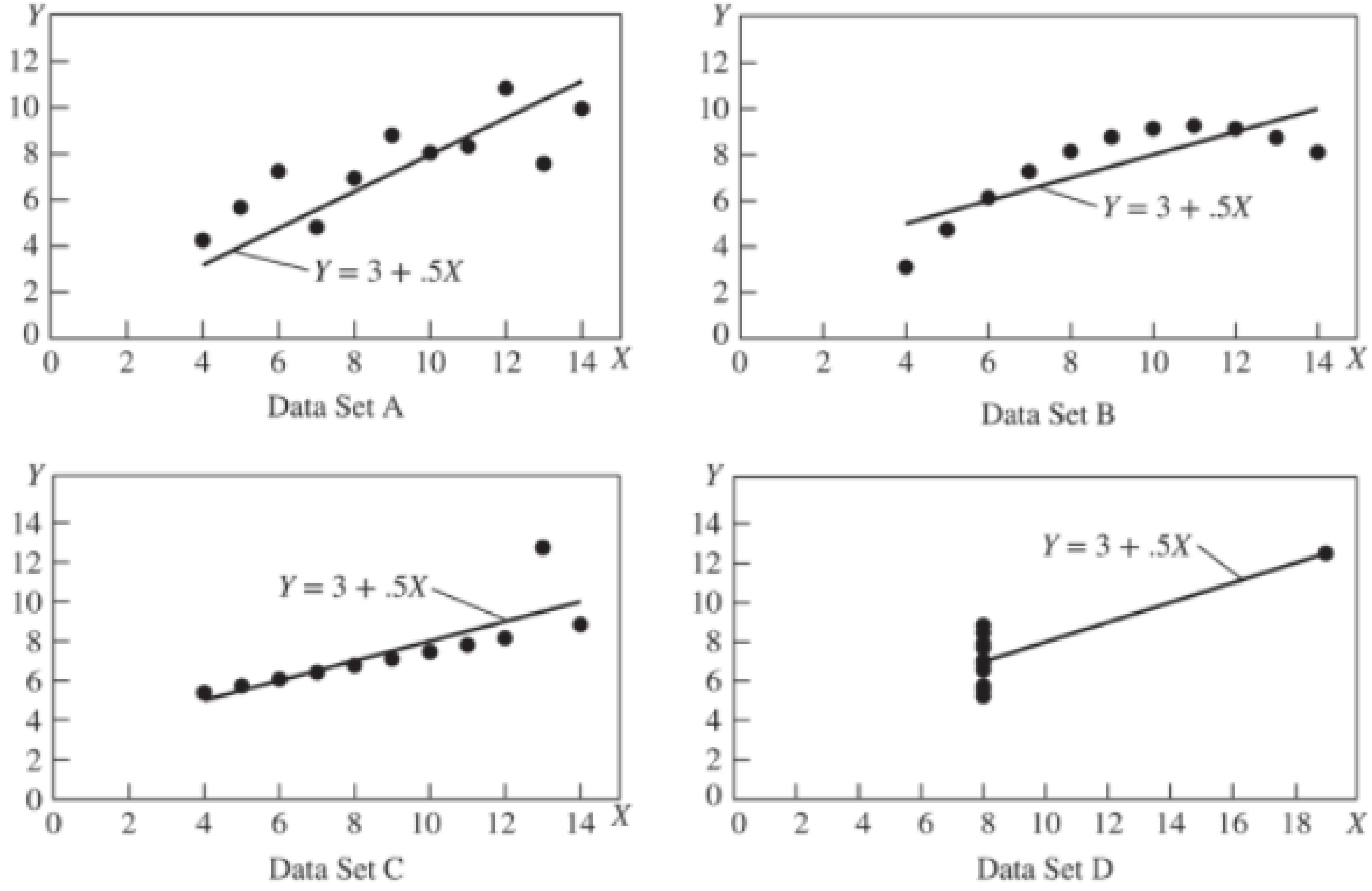
End-consumer
"demand"



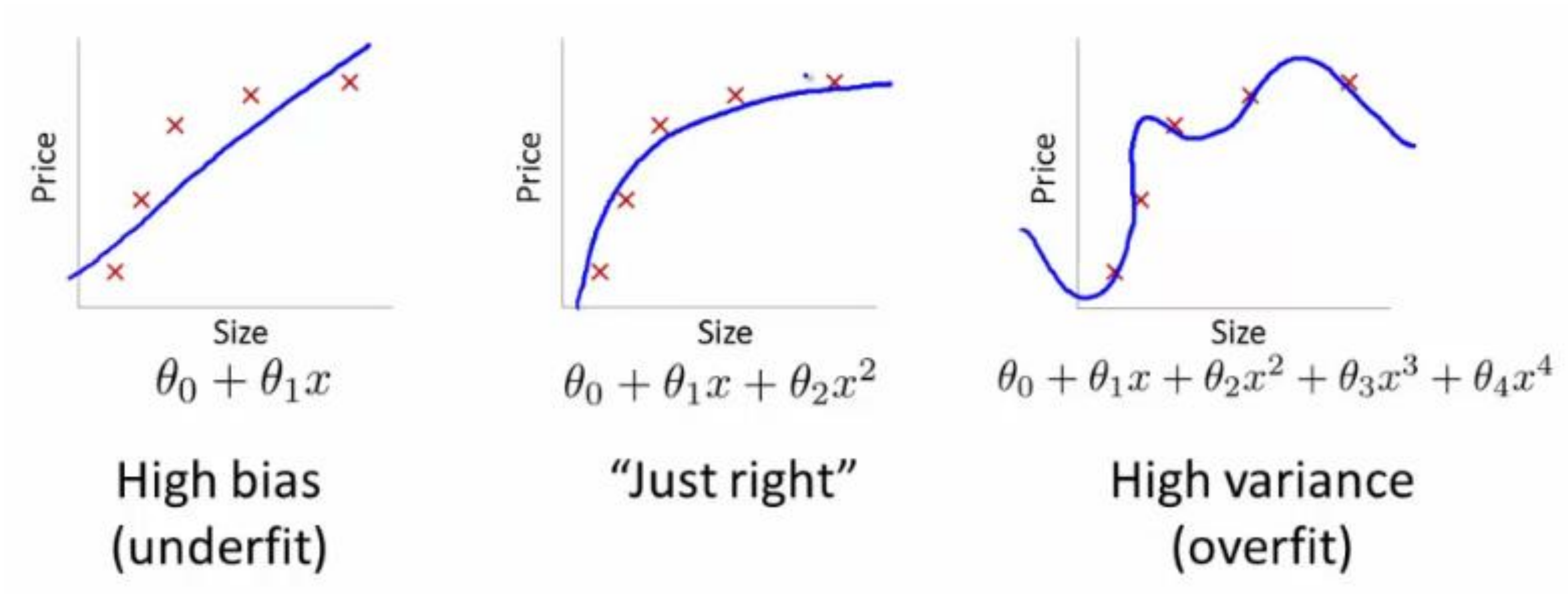
https://link.springer.com/article/10.1007/s10115-016-0954-8

Bullwhip Effect: Increase in order variability as we travel up in the chain

Anscombe's quartet



It is important to *look* at the data before plunging into data analysis and the selection of an appropriate set of forecasting techniques.


Price
Size
$\theta_0 + \theta_1 x$
High bias
(underfit)
Price
Size
$\theta_0 + \theta_1 x + \theta_2 x^2$
"Just right"
Price
Size
$\theta_0 + \theta_1 x + \theta_2 x^2 + \theta_3 x^3 + \theta_4 x^4$
High variance
(overfit)

Accuracy measures

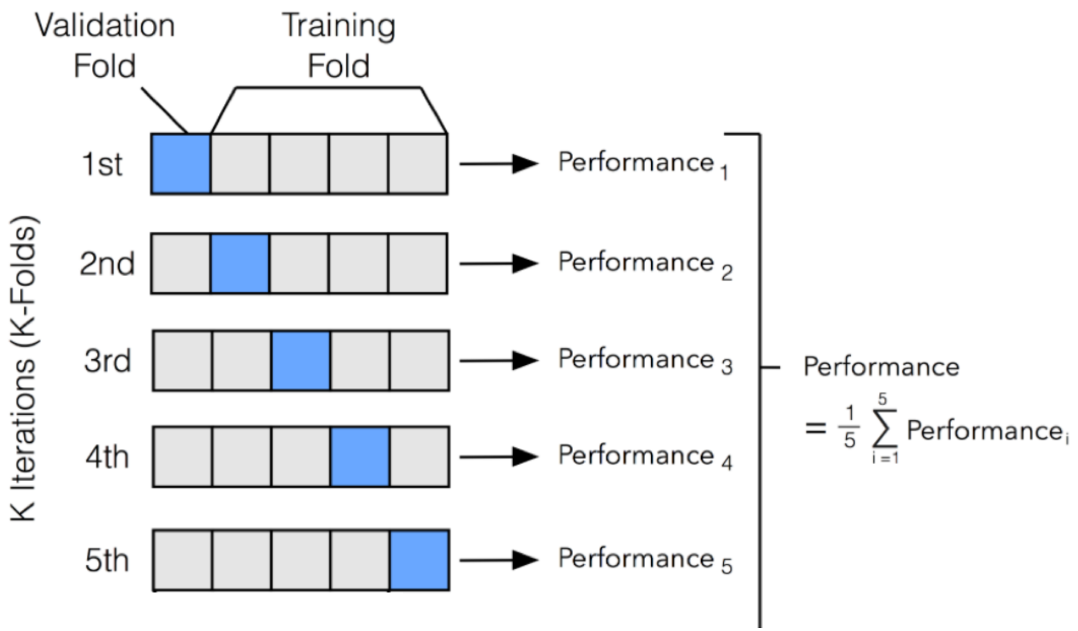
MSE	Mean Squared Error	$= \text{mean}(e_t^2)$
RMSE	Root Mean Squared Error	$= \sqrt{\text{MSE}}$
MAE	Mean Absolute Error	$= \text{mean}(\lvert e_t \rvert)$
MdAE	Median Absolute Error	$= \text{median}(\lvert e_t \rvert)$
MAPE	Mean Absolute Percentage Error	$= \text{mean}(\lvert p_t \rvert)$
MdAPE	Median Absolute Percentage Error	$= \text{median}(\lvert p_t \rvert)$
sMAPE	Symmetric Mean Absolute Percentage Error	$= \text{mean}(2\lvert Y_t - F_t \rvert/(Y_t + F_t))$
sMdAPE	Symmetric Median Absolute Percentage Error	$= \text{median}(2\lvert Y_t - F_t \rvert/(Y_t + F_t))$
MRAE	Mean Relative Absolute Error	$= \text{mean}(\lvert r_t \rvert)$
MdRAE	Median Relative Absolute Error	$= \text{median}(\lvert r_t \rvert)$
GMRAE	Geometric Mean Relative Absolute Error	$= \text{gmean}(\lvert r_t \rvert)$
RelMAE	Relative Mean Absolute Error	$= \text{MAE}/\text{MAE}_b\,.$
RelRMSE	Relative Root Mean Squared Error	$= \text{RMSE}/\text{RMSE}_b\,.$
LMR	Log Mean Squared Error Ratio	$= \log(\text{RelMSE})$
PB	Percentage Better	$= 100\,\text{mean}(I\{\lvert r_t \rvert < 1\})$
PB(MAE)	Percentage Better (MAE)	$= 100\,\text{mean}(I\{\text{MAE} < \text{MAE}_b\})$
PB(MSE)	Percentage Better (MSE)	$= 100\,\text{mean}(I\{\text{MSE} < \text{MSE}_b\})$

Source: Gooijer and Hyndman (2006). 25 Years of Time Series Forecasting.

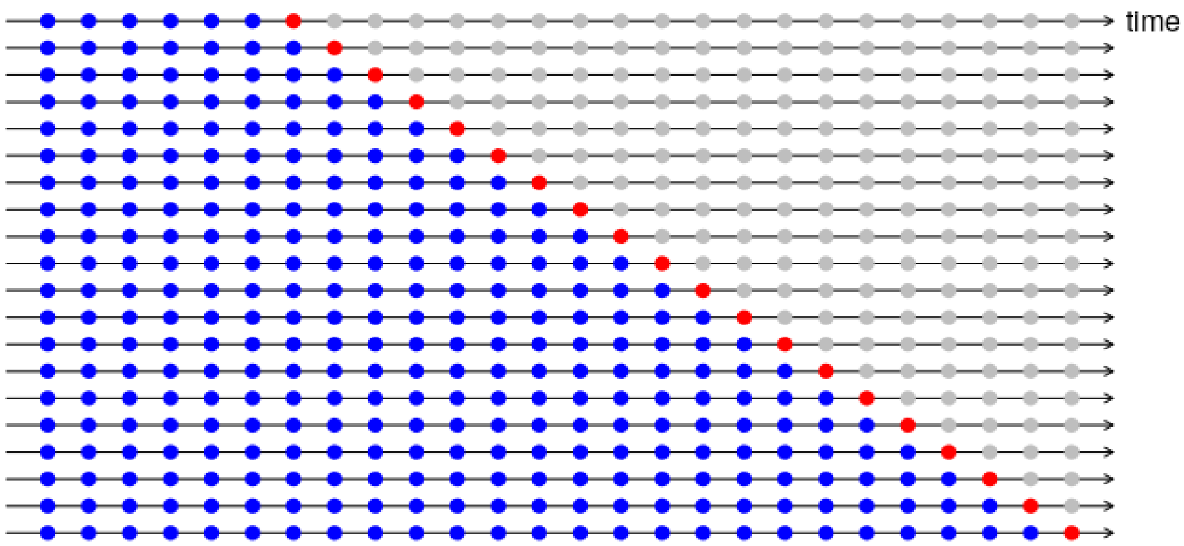
<table>
<tr><td colspan="3">Original Data</td></tr>
<tr><td colspan="2">Training Data</td><td>Testing Data (Holdout Sample)</td></tr>
<tr><td>Training Data</td><td>Validation Data</td><td>Testing Data</td></tr>
</table>

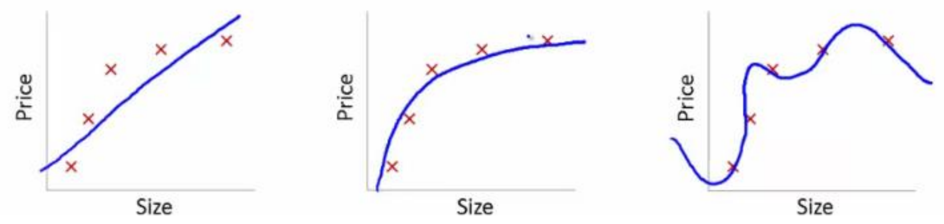
Original Data		
Training Data		Testing Data (Holdout Sample)
Training Data	Validation Data	Testing Data

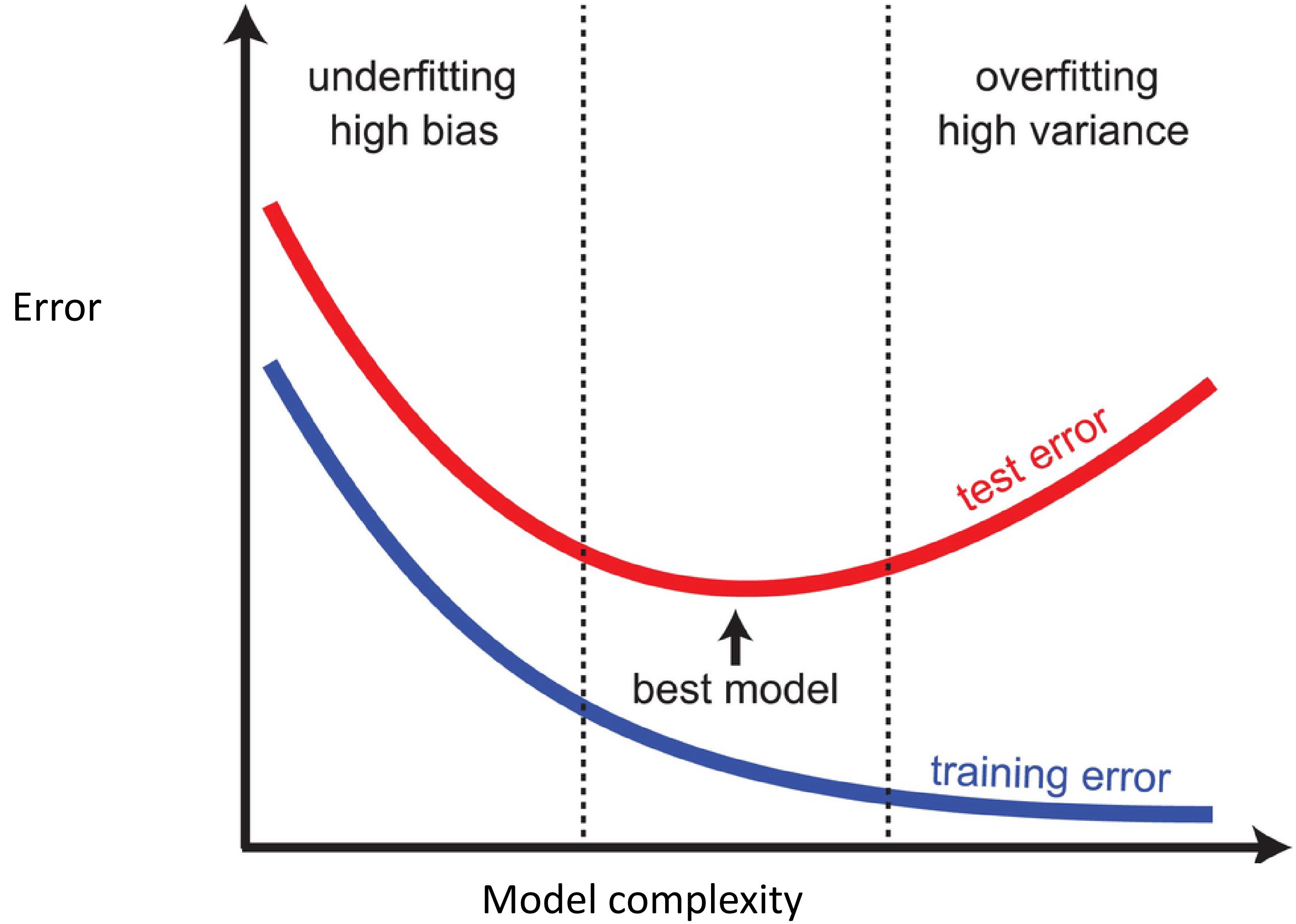
K-fold cross variation

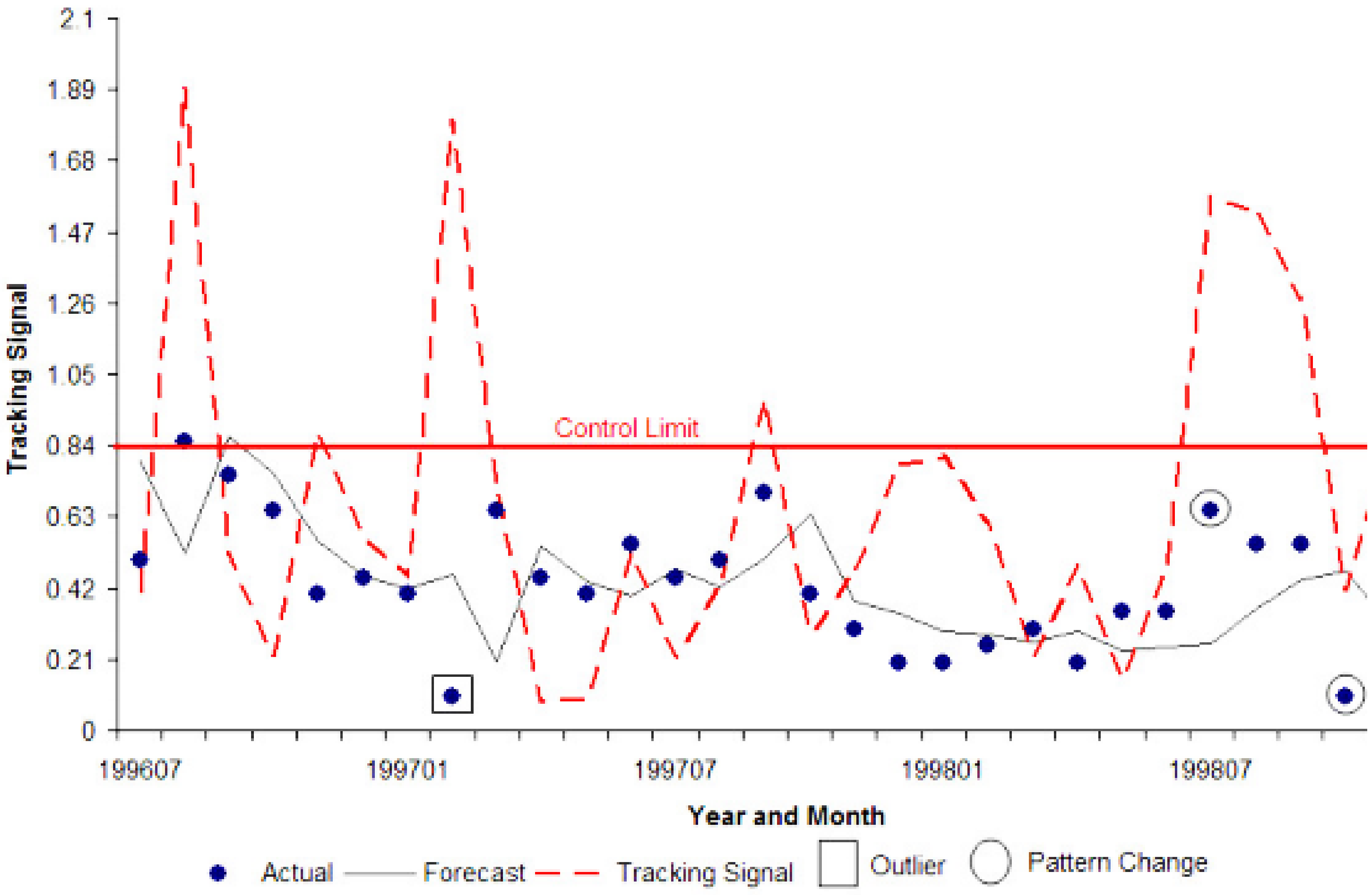


Time series cross variation




Price
Size
Price
Size
Price
Size


underfitting
high bias
overfitting
high variance
Error
test error
best model
training error
Model complexity



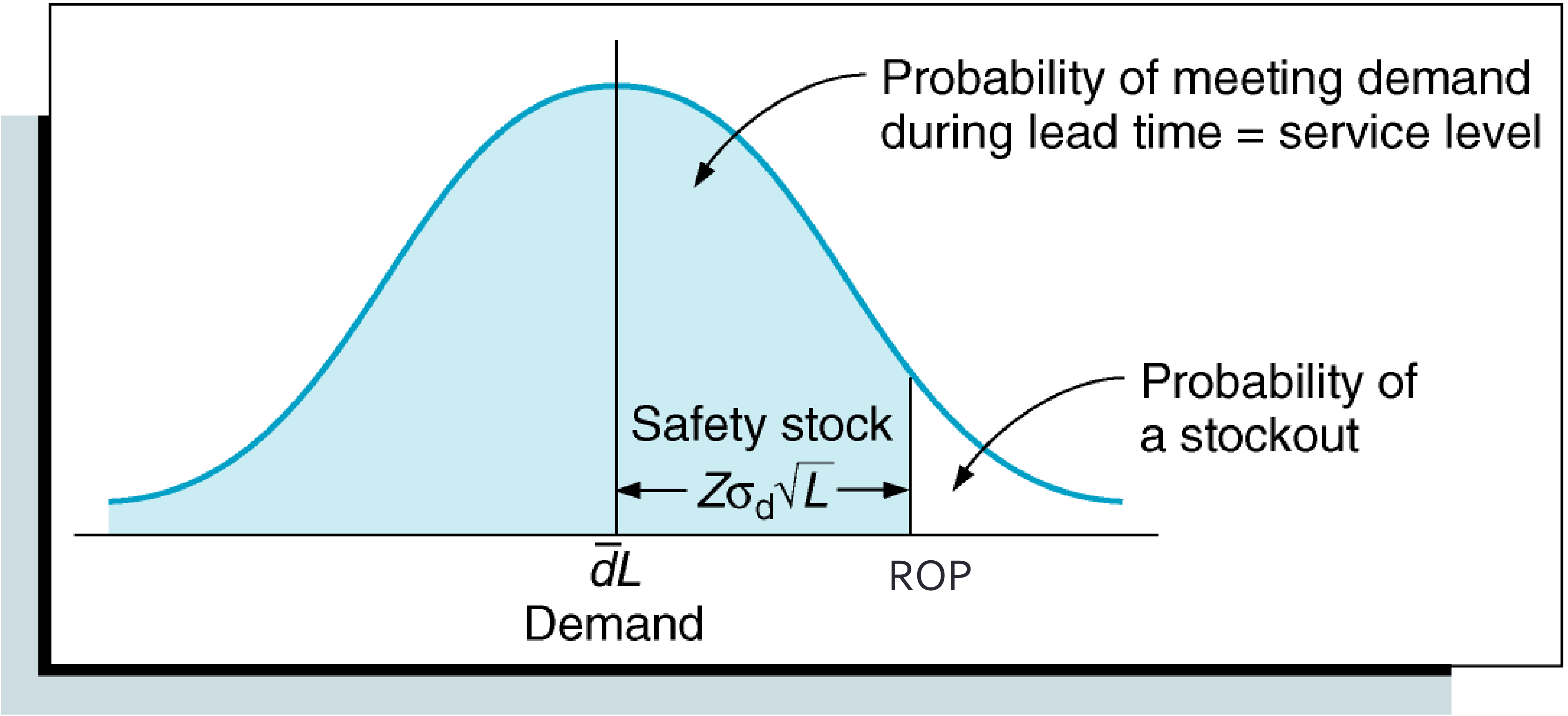
Ensure that process is in place to find exceptions and flag them (managing exceptions) so that corrective action can be taken

Forecasts are always wrong!

Famous predictions about computing

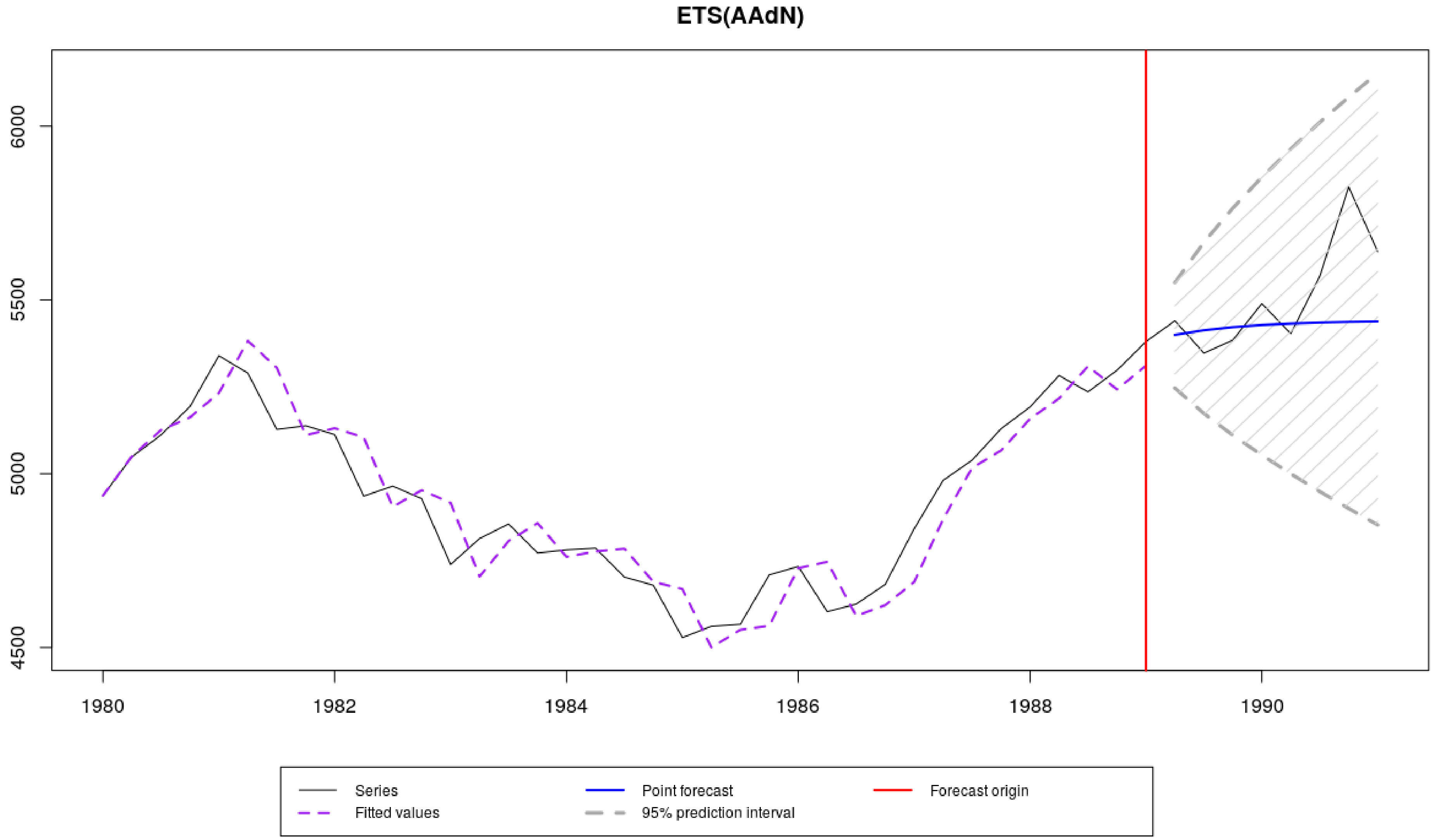
- *"I think there is a world market for maybe five computers."* (Chairman of IBM, 1943)
- *"Computers in the future may weight no more than 1.5 tons."* (Popular Mechanics, 1949)
- *" There is no reason anyone would want a computer in their home."* (President, DEC, 1977)

Safety Stock Calculation

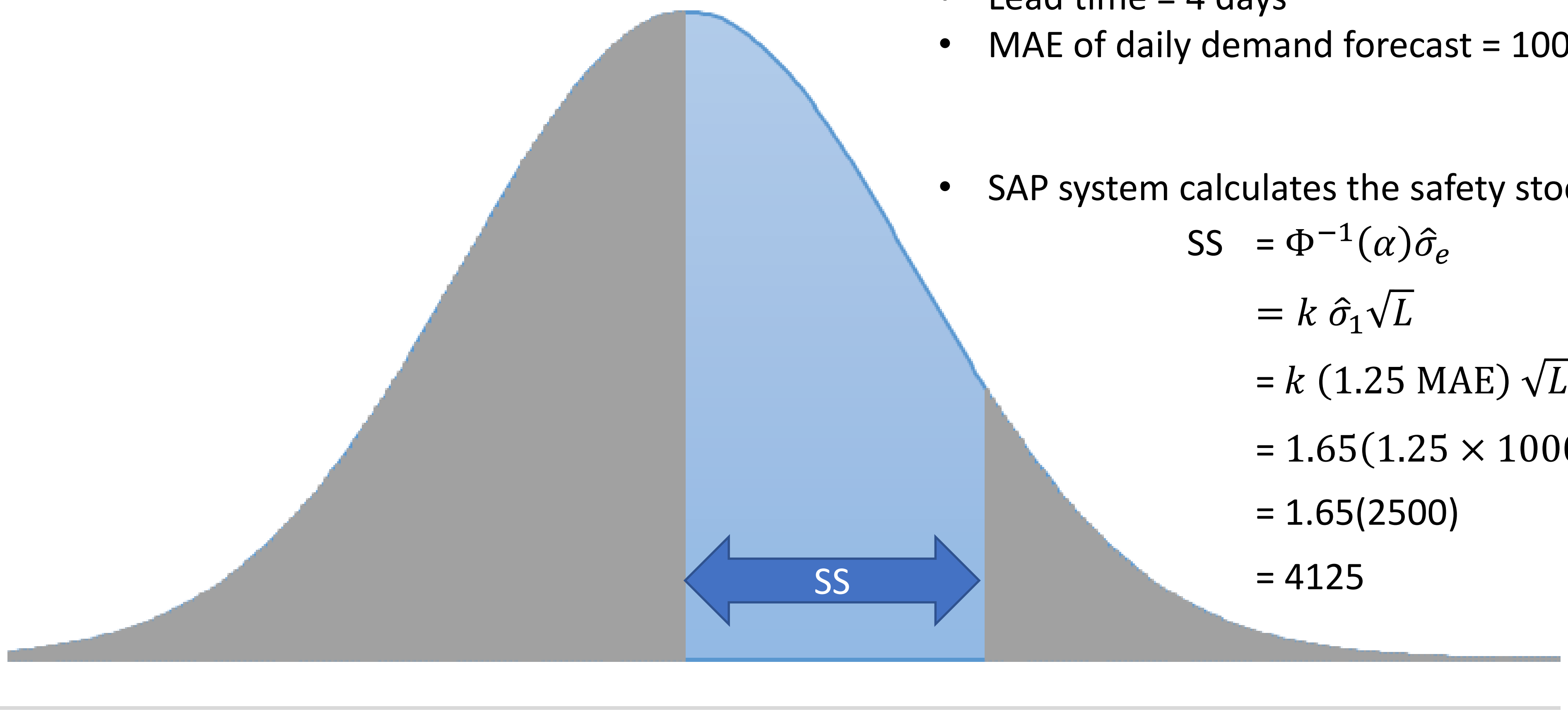


SL	90%	91%	92%	93%	94%	95%	96%	97%	98%	99%	99.9%
k	1.29	1.34	1.41	1.48	1.56	1.65	1.75	1.88	2.05	2.33	3.08

Interval forecasts provide insight to risks



Safety stock: Normal Approximation



Example

- Cycle service level = 95%
- Lead time = 4 days
- MAE of daily demand forecast = 1000 units

- SAP system calculates the safety stock as follows:

$$
\begin{aligned}
SS &= \Phi^{-1}(\alpha)\hat{\sigma}_e \\
&= k\,\hat{\sigma}_1\sqrt{L} \\
&= k\,(1.25\ \mathrm{MAE})\,\sqrt{L} \\
&= 1.65(1.25 \times 1000)\sqrt{4} \\
&= 1.65(2500) \\
&= 4125
\end{aligned}
$$

https://help.sap.com/doc/erp2005_ehp_02/6.02/en-US/37/67b65334e6b54ce10000000a174cb4/content.htm?no_cache=true

Aggregate forecasts are more accurate

Table 2. MAPEs for Monthly Sales Forecast Source: Jain & Malehorn (2006, Table 6.2)

Horizon	1 month			2 months			1 quarter			1 year		
Level	SKU	Category	Aggregate	SKU	Category	Aggregate	SKU	Category	Aggregate	SKU	Category	Aggregate
Automotive	25% n = 3	5% n = 1	36% n = 1	31% n = 3	33% n = 2	25% n = 2	42% n = 1			46% n = 1		10% n = 1
Computer/ Technology	19% n = 4	14% n = 4	12% n = 7	33% n = 2	11% n = 2	18% n = 4	30% n = 3	16% n = 4	25% n = 6	17% n = 2	30% n = 1	31% n = 4
Consumer Products	27% n = 35	20% n = 23	15% n = 21	29% n = 20	22% n = 14	15% n = 10	33% n = 11	23% n = 7	14% n = 6	48% n = 4	19% n = 4	8% n = 3
Food/ Beverages	26% n = 16	15% n = 10	18% n = 11	28% n = 10	22% n = 4	36% n = 5	26% n = 8	21% n = 3	40% n = 4	19% n = 4	14% n = 2	48% n = 3
Healthcare	25% n = 7	15% n = 6	9% n = 6	27% n = 5	19% n = 5	17% n = 5	41% n = 5	24% n = 5	25% n = 5	30% n = 2	20% n = 2	15% n = 2
Industrial Products	22% n = 4	15% n = 7	7% n = 8	16% n = 2	14% n = 5	8% n = 6	17% n = 3	15% n = 6	10% n = 7	40% n = 2	21% n = 5	15% n = 6
Pharma	26% n = 5	20% n = 4	23% n = 4	30% n = 3	35% n = 2	33% n = 2	31% n = 4	25% n = 4	25% n = 3	34% n = 4	35% n = 4	28% n = 3
Retail	24% n = 7	18% n = 4	7% n = 4	17% n = 5	17% n = 6	8% n = 4	24% n = 4	10% n = 3	9% n = 4	23% n = 4	6% n = 2	6% n = 3
Telco				30% n = 1	10% n = 1	30§ n = 1	40% n = 1	15% n = 1	35% n = 1			
Others	28% n = 13	21% n = 9	17% n = 16	23% n = 7	20% n = 5	11% n = 10	25% n = 6	15% n = 5	14% n = 9	15% n = 4	18% n = 4	12% n = 8
Overall	26% n = 94	18% n = 68	13% n = 80	27% n = 58	20% n = 46	15% n = 51	30% n = 46	19% n = 37	17% n = 45	29% n = 27	21% n = 24	16% n = 33

Warehouse pooling

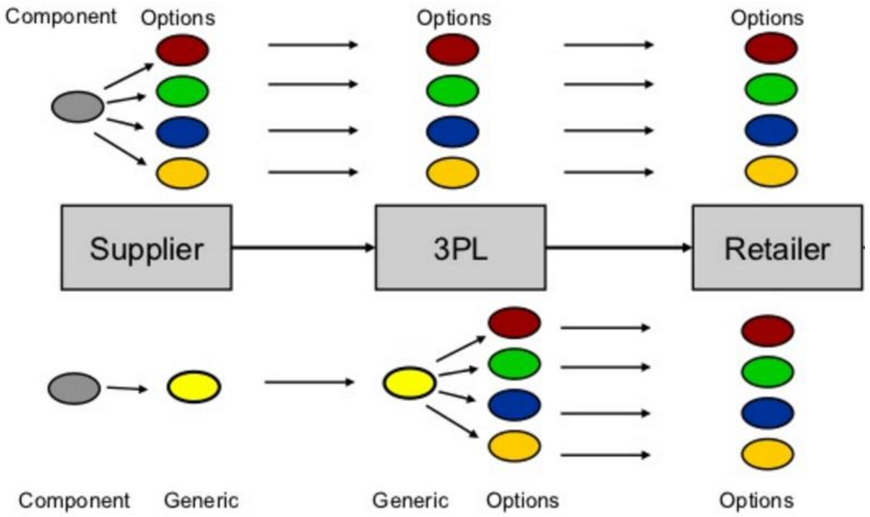
การกำหนดระดับสินค้าคงคลังที่เหมาะสมและการรวมคลังสินค้า: กรณีศึกษาคลังสินค้าในทวีปยุโรปของบริษัทผู้ผลิตสารเคลือบผิว

2/11/2018

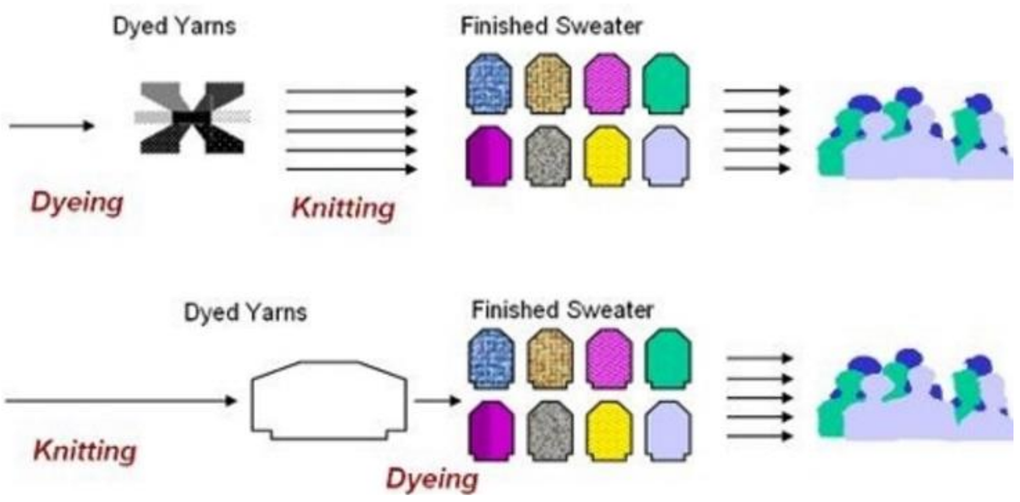
การกำหนดระดับสินค้าคงคลังที่เหมาะสมและการรวมคลังสินค้า: กรณีศึกษาคลังสินค้าในทวีปยุโรปของบริษัทผู้ผลิตสารเคลือบผิว อัญชลี แซ่เจียม กาญจ์นภา อมรัชกุล ศิวิกา ดุษฎีโหนด การประชุมวิชาการสถิติประยุกต์และเทคโนโลยีสารสนเทศระดับชา

http://logistics.nida.ac.th/optinvwh/

Delayed differentiation/Product Postponement



Source: Logistikgerechte-konzeption



Benetton's postponement

References

- J. E. Hanke and D. W. Wichern. *Business Forecasting.* Pearson Education, Inc., New Jersey, 2009.
- R. J. Hyndman, and G. Athanasopoulos. *Forecasting: Principles and Practices,* Second Edition, OText: Melbourne, Australia. OTexts.com/fpp2. Accesses on 10 May 2019.
- C. L. Jain. *Fundamentals of Demand Planning and Forecasting*. Graceway Publishing Company, Inc., New York, 2017.
- C. L. Jain and J. Malehorn. *Practical Guide to Business Forecasting*. Graceway Publishing Company, Inc., New York, 2005.
- B. Keating and J. H. Wilson. *Forecasting and Predictive Analytics*. McGraw Hill, Boson, 2019.
- K. Ord and R. Fildes and N. Kourentzes. *Principles of Business Forecasting*. Wessex Press, Inc., 2017.
- S. Makridakis and S. C. Wheelwright and R. J. Hyndman. *Forecasting: Methods and Applications*, Third Edition. John Wiley & Sons, Inc., New Jersey, 1998.
- D. C. Montgomery and C. L. Jennings and M. Kulahci. *Introduction to Time Series Analysis and Forecasting.* John Wiley & Sons, Inc., New Jersey, 2008.
- P. Tetlock and D. Gardner. *Superforecasting: The Art & Science of Prediction*. Penguin Random House, 2015.

From Demand Forecasting (Predictive) to Inventory Optimization (Prescriptive)

Kannapha Amaruchkul

"Big data is not about the data"

Gary King, Harvard University, making the point that while data is plentiful and easy to collect, the real values in in the analytics.

AS CONFERENCE 2019 & The 4th BADS

NIDA GSAS Graduate School of Applied Statistics