Technische Universität Berlin

Quantitative Decision Making: Time series Analysis

Faculty of Economics & Management, Institute of Technology and Management

06. August 2018

Discussion

98

… out of 100 Trial, on average, „blindfolded Monkey throwing darts at a newspaper's financial pages" beat market experts picking stocks" beat Market Experts in a comparison of annualized return from 1964 to 2010.

- There were no real monkeys involved – the idea for a scientific experiment leading to this conclusion is just based on a 1973 quote from a Princeton University Professor that blindfolded monkey could beat stock experts

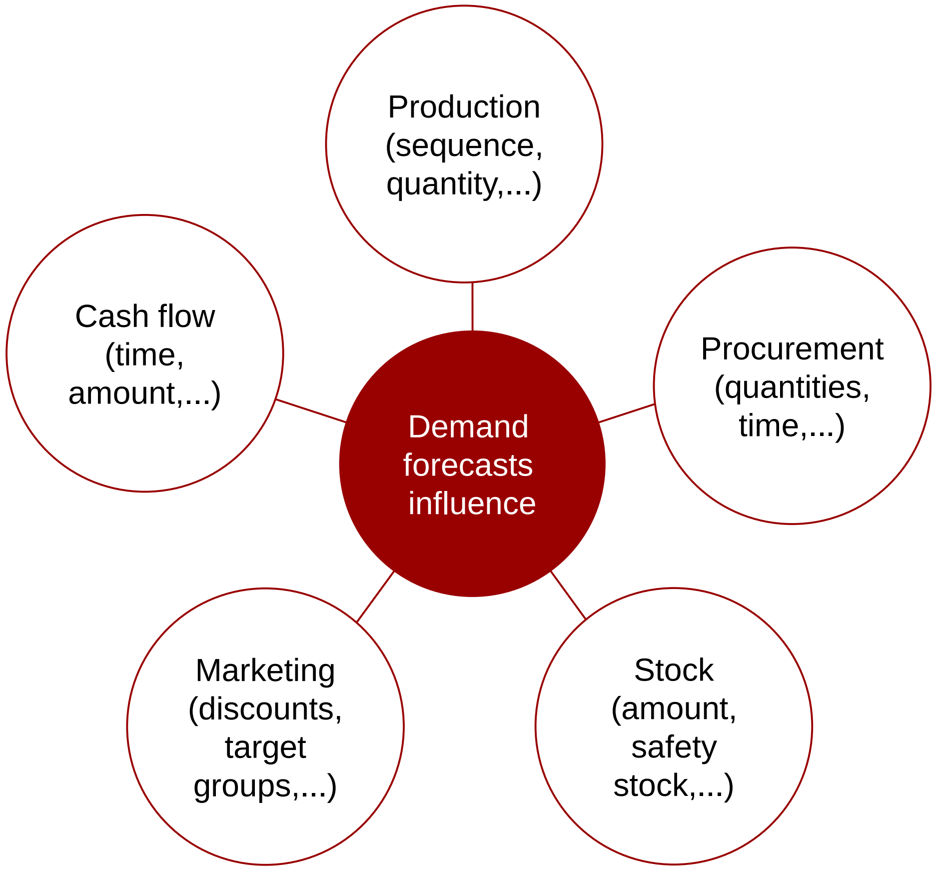
In the scientific experiment, for every year between 1964 and 2010, 100 random stock portfolios were selected out of a 1000 Stock pool and compared to the capitalization stock index

- The random portfolios usually did better, because more smaller companies were picked which had higher return on average
- However, smaller companies have a higher risk as well, resultingly it is not that simple as just investing in smaller companies

Agenda

1. Demand forecasting
2. Time series analysis I – naive and simple
3. Time series analysis II – pattern recognition

Demand Forecast



Strong deviations from the forecast to the actual value often has an impact on planning
→ resulting in costs

Source: Camm et. al (2015)

Quantitative methods for demand forecasting

- Quantitative methods for demand forecasting are useful if

1. Historical data exists on the variable to be predicted

2. The information is quantifiable

3. It can be assumed that past behaviour will **continue**

Source: Camm et. al (2015)

Characteristics of demand forecasts

- *„ Demand forecasts are always imprecise“*
  - …due to randomness. The aim is to separate the **systematic** part from the **random** part.
  - Therefore, the forecast error (error of the model) of the model should also be specified
  - The error is also used as a unit of uncertainty in the models working with the forecast
- Long-term forecasts are generally less accurate than short-term forecasts.
  - Measured by the relative ratio of standard deviations of the error to the mean value
- Aggregated forecasts are more accurate
  - Measured by the relative ratio of standard deviations of the error to the mean value.
  - In line with the pooling effect, large deviations at the disaggregated levels balance each other out at the aggregated level
- The further away a company is from the end customer (upstream goods flow), the higher the distortion of demand information.
  - See Bullwhip Effect

Source: Chopra und Meindl (2015)

Execution of the demand forecast

- Predicting demand is based on customer behavior in the past
- Customer demand depends on a variety of factors
  - Is the most important dependencies and relationships determinable → demand can be determined with a certain probability ("causal analysis" or "exploratory analysis")
  - The identification of "causal factors "* can already be improved by subjective, qualitative human input.
  - For example, past demand, delivery time, planned marketing measures, planned price reductions, economic situation, actions of competitors should be taken into account.
- … however, if these factors are not available or cannot be determined
  - Demand can be estimated with a time series analysis
  - Demand patterns from the past are continued

*häufig wird nicht der kausale Zusammenhang bestimmt, sondern eine Korrelation

Source: Chopra und Meindl (2015)

Methods of demand forecasting

*Actually wrong description

Forecasting

Qualitative

- Primarily subjective and based on human judgment.
- Useful in case of lack of historical data and availability of market experts. Used for very long-term forecasts.

Time series analysis

- Uses historical demand as an indicator for future demand
- Reasonable with no significant change in demand patterns.
- Simple model, but a "good starting point".

"Causal Analysis "*

- Methods assume a strong **correlation** between demand and environmental factors (weather, economic situation,...)
- Estimates of environmental factors are used accordingly to forecast demand.

Simulation

- Methods mimic customers' choices to predict demand.
- Simulation uses the methods of causal and time series analysis and combines them to evaluate changes.

Source: Chopra und Meindl (2015)

Summary: Demand Forecasting

- *"Prediction is always imprecise":* random events will always affect the actual event. However, the value of the prediction lies in **systematically** understanding the **non-random** part …
- … and to reduce the difference between prediction and actual event and to control uncertainty.
- Demand forecasts show exemplarily that predictions (depending on previous knowledge and information) can be solved with different techniques
- If data to examine possible correlations of the target variable with external factors are not possible, the demand can be estimated with time series analyses.

The following topic :

Demand is one of the most important variables to predict in logistics. The quantitative method that works with the least data is **time series analysis**.

Agenda

1. Demand forecasting
2. Time series analysis I – naive and simple
3. Time series analysis II – pattern recognition

Usability of time series analysis

- In addition to demand forecasts, time series analysis can be used in logistics in a variety of ways
- ...if stable patterns are available!
- Examples:
  - raw material prices
  - Delivery times / delivery delays
  - Reliability of suppliers
  - Performance comparison of delivery times
  - Monitoring of supply chain events
  - Discovery of Supply Chain Disruptions
  - Failure detection in production processes
  - …

Components of a time series analysis

- The actual realised value ("observed value") can be divided into two components:
  - $Observed\ value\ (y) = systematic\ component\ (S) + random\ component(R)$
  - $Observed\ value(y) = \underline{Prediction\ (\hat{y})} + Error\ (e)$

S — **Systematic component**

- Measures the expected value based on patterns such as level, trend, seasonality (fixed periods), or cycles (varying periods)

R — **Random component**

- The part of the prediction by which the systematic component deviates from the observation

ŷ — **Prediction**

- The estimation of the systematic component explained by the model
- (The circumflex / the "little hat" indicates that the value is an estimate)

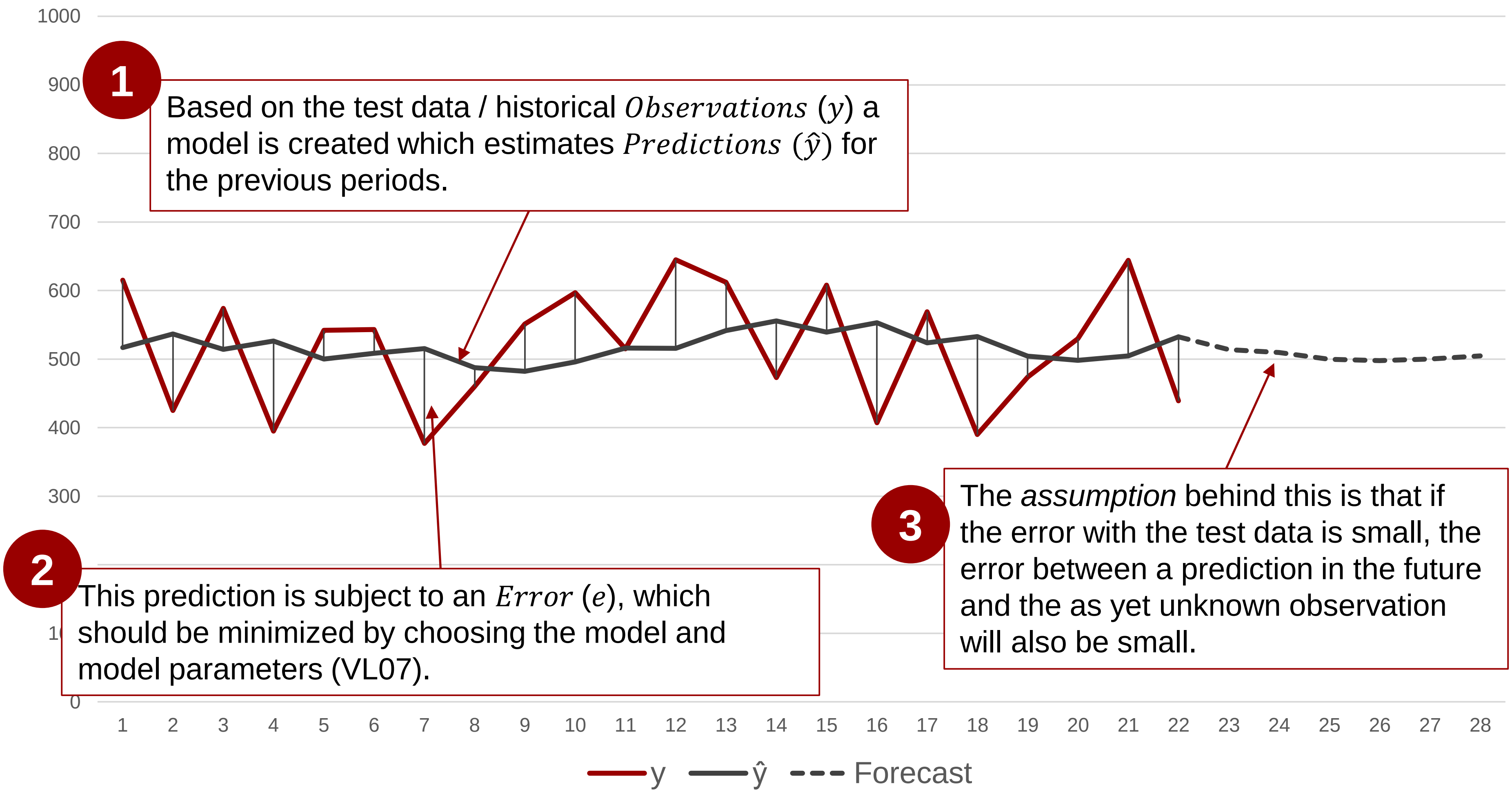
e — **Forecast Error**

- The difference between Prediction and Observed Value
- This should come as close as possible to the random component
- The size and variability must be evaluated, the direction of the error is irrelevant

$\neq$

Source: Chopra und Meindl (2015); Camm et al (2015)

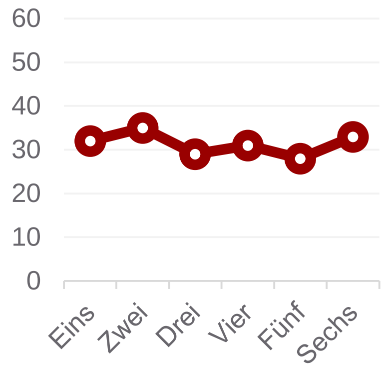
Motivation for modeling and error evaluation



Time series pattern

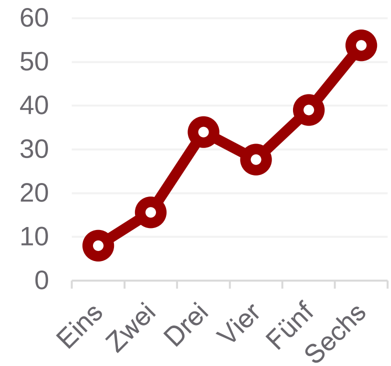
Level

Values fluctuate randomly around a constant level.



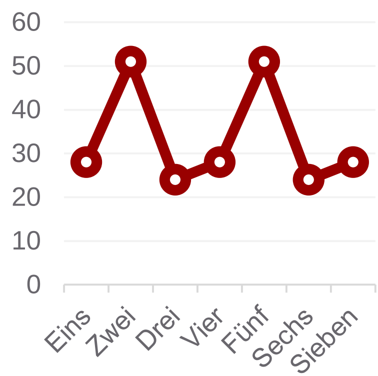
Trend

Gradual development to higher or lower values.



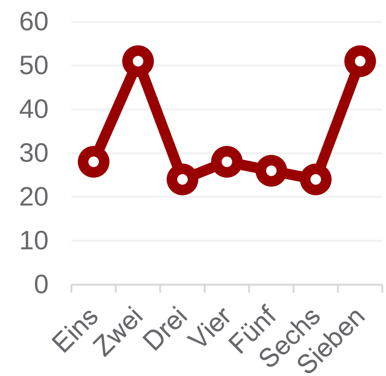
Seasonality

Recurring behavior of successive periods.



Cycle

Recurrent behavior in irregular periods.



Source: Chopra und Meindl (2015); Camm et al (2015)

Composition of the systematic component

- The aim of time series analysis is to predict the systematic component and to estimate the random component.

- Forms of calculation → basically similar procedure, but the accuracy of models varies depending on the situation

  - Multiplicative: systematic component = level x trend x seasonality

  - Additive: systematic component = level + trend + seasonality

  - Mixed: Systematic component = (level + trend) x seasonality

Source: Chopra und Meindl (2015); Camm et al (2015)

Types of methods

- Static methods
  - Static methods assume that the estimates of the patterns (level, trend, seasonality) do not change when new observations are available

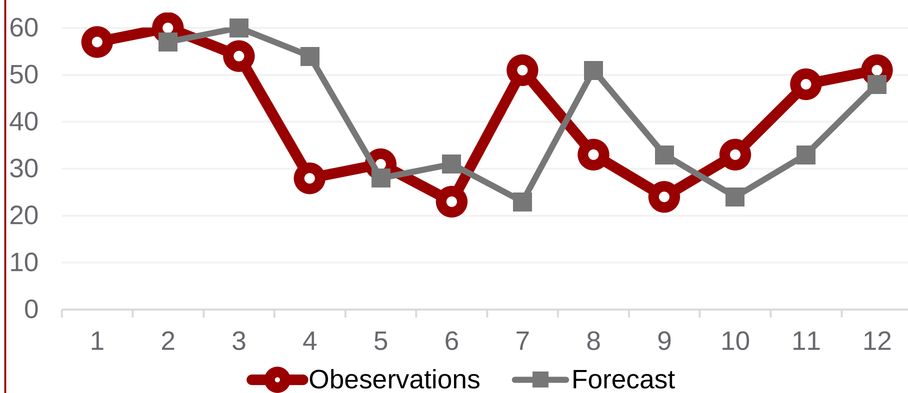
- Adaptive Methods
  - In adaptive methods, the estimates of the patterns (level, trend, seasonality) are adjusted after each new observation.
  - Shapes (exemplary):
    - Naive (Forecast is the observation of the last period)
    - Average over all previous periods → the average is adjusted for each new period
    - moving average
    - Exponential smoothing
    - Trend-corrected exponential smoothing (Holt's model)
    - Trend and seasonality corrected exponential smoothing (Holt-Winter model)

Source: Chopra und Meindl (2015); Camm et al (2015)

Simple models

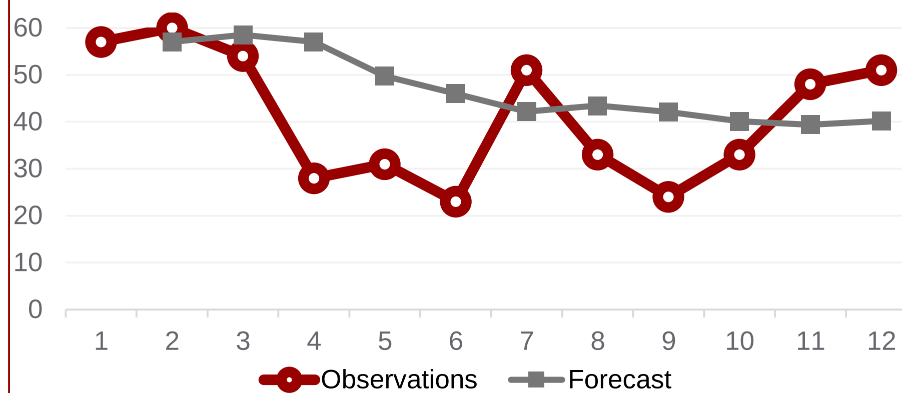
Naïve model

- The observation of the current period is the prediction of the next period
- $\hat{y}_{t,t+1} = y_t$



Durchschnitt über gesamte Historie

- The average over all periods considered is used as a prediction
- $\hat{y}_{t,t+1} = \frac{\sum_{i=0}^{n} y_i}{n}$



Nomenclature

$y_t - observed\ value\ in\ period\ t$

$\hat{y}_{t,t+x} - predicted\ value\ for\ period\ t + x\ on\ the\ basis\ of\ the\ values\ of\ period\ t$

$n\ - number\ of\ all\ periods$

Source: Chopra und Meindl (2015); Camm et al (2015)

Moving Average

Nomenclature

y_t – *observed value in period t*

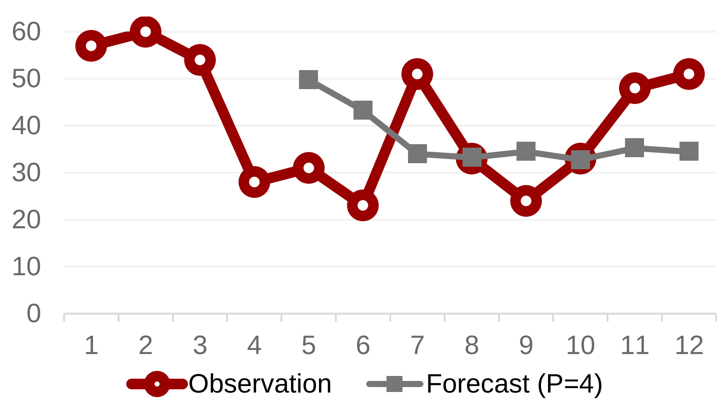
$\hat{y}_{t,t+x}$ – *predicted value for period t + x on the basis of the values of period t*

n – *Number of all periods*

P – *Number of all periods **considered***

Moving Average

- The mean of the previous fixed number of periods is used to make the prediction.
- Suitable for demand patterns: level
- $\hat{y}_{t,t+1} = \frac{\sum_{i=n-P}^{n} y_i}{P} = \frac{(y_{n-P}+y_{n-(P-1)}+\cdots+y_{n-(P-P)})}{P}$

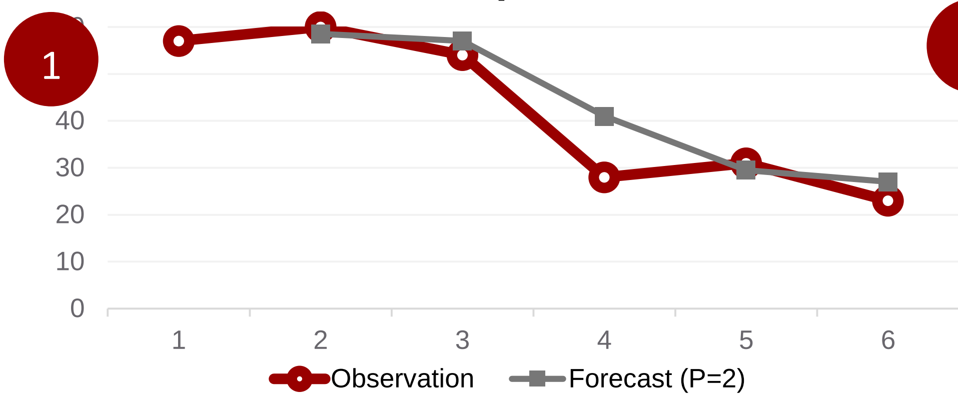


Period	Obersvation	Forecast (P=4)	Error
1	57	/	/
2	60	/	/
3	54	/	/
4	28	/	/
5	31	49,75	18,75
6	23	43,25	20,25
7	51	34	17
8	33	33,25	0,25
9	24	34,5	10,5
10	33	32,75	0,25
11	48	35,25	12,75
12	51	34,5	16,5

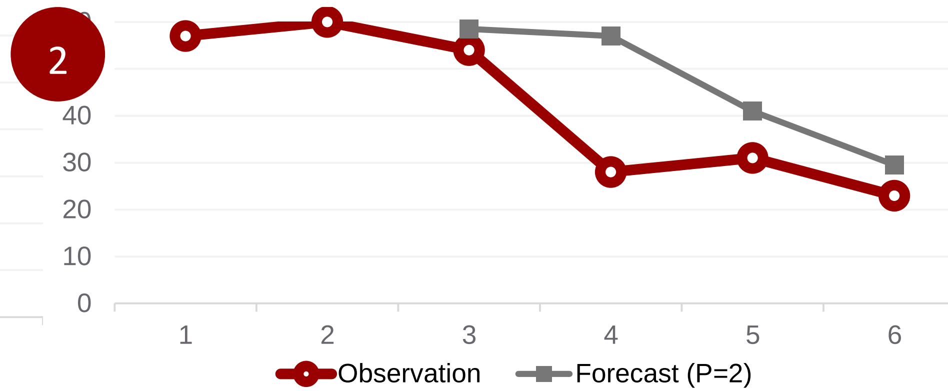
Source: Chopra und Meindl (2015); Camm et al (2015)

A word on representation

- According to the nomenclature, the forecast for period t+1 should be in the line of t. However, this results in a table that distorts the display, since the diagram does not show the forecast for the next period in the next period (1).
- If this is avoided and the forecast is in the line for the period (2) that is predicted, the correct formulation would be : $\hat{y}_{t-1,t} = \frac{\sum_{i=n-1-P}^{n-1} y_i}{P}$



Period	Observation	Forecast (P=2)
1	57	
2	60	58,5
3	54	57
4	28	41
5	31	29,5
6	23	27



Period	Observation	Forecast (P=2)
1	57	
2	60	
3	54	58,5
4	28	57
5	31	41
6	23	29,5

Calculation in Excel (See Problem Set 11-0)

Example - Solution will be presented

Naïve model

C4 =B3

	A	B	C
1			Naives Modell
2	Periode	y_t - Beob. Bedarf	ŷ_t-1,t
3	1	376	###
4	2	403	376
5	3	391	403
6	4	416	391
7	5	395	416
8	6	417	395
9	7	417	417
10	8	417	417
11	9	435	417
12	10	403	435
13	11	391	403
14	12	452	391
15	13	411	452
16	14	403	411
17	15	372	403
18	16	409	372
19	17	368	409
20	18	399	368
21	19	408	399
22	20	440	408
23	21	369	440
24	22	390	369
25	23	407	390
26	24	378	407
27	25	405	378

Total history

H8 =MITTELWERT($B$3:B7)

	A	B	H
1			Gesamte Historie
2	Periode	y_t - Beob. Bedarf	ŷ_t-1,t
3	1	376	###
4	2	403	376
5	3	391	389,5
6	4	416	390
7	5	395	396,5
8	6	417	396,2
9	7	417	399,6666667
10	8	417	402,1428571
11	9	435	404
12	10	403	407,4444444
13	11	391	407
14	12	452	405,5454545
15	13	411	409,4166667
16	14	403	409,5384615
17	15	372	409,0714286
18	16	409	406,6
19	17	368	406,75
20	18	399	404,4705882
21	19	408	404,1666667
22	20	440	404,3684211
23	21	369	406,15
24	22	390	404,3809524
25	23	407	403,7272727
26	24	378	403,8695652
27	25	405	402,7916667

Moving Average (P=2)

M8 =MITTELWERT(B6:B7)

	A	B	M
1			Moving Average (P=2)
2	Periode	y_t - Beob. Bedarf	ŷ_t-1,t
3	1	376	###
4	2	403	###
5	3	391	389,5
6	4	416	397
7	5	395	403,5
8	6	417	405,5
9	7	417	406
10	8	417	417
11	9	435	417
12	10	403	426
13	11	391	419
14	12	452	397
15	13	411	421,5
16	14	403	431,5
17	15	372	407
18	16	409	387,5
19	17	368	390,5
20	18	399	388,5
21	19	408	383,5
22	20	440	403,5
23	21	369	424
24	22	390	404,5
25	23	407	379,5
26	24	378	398,5
27	25	405	392,5

Forecast Error I/III

Four measures follow to evaluate the quality of the prediction models

- All are based on existing observations

Assumptions and procedures :

- the most accurate method for known data produces the most likely prediction for unknown data
- All measurements are based on the prediction error e:

$$e_t = y_t - \hat{y}_t$$

Note $\boldsymbol{t}$: The period on which the forecast was based is not relevant here. It is relevant that forecasts and observations of the same period are compared. (See "A word for representation")

Note:

- Every measure has advantages and disadvantages. It therefore makes sense to compare several measures
- In some prediction models there is no prediction for an initial number of periods → omit these in the error analysis.

Source: Chopra und Meindl (2015); Camm et al (2015)

Forecast Error II/III

Nomenclature

Values:

y_t – *observed value in period t*

$\hat{y}_{t,t+x}$ – *predicted value for period $t + x$ on the basis of the values of period t*

e – *Error*

Indices:

n – *Number of all periods*

P – *Number of all periods **considered***

k – *Number of periods, for which no prediction is possible*

t – *Control variable of the sum (considered period)*

Mean Forecast Error (MFE)

- Average error.
- Positive and negative errors can balance each other out with this measure!

$$MFE = \frac{\sum_{t=k+1}^{n} e_t}{n-k}$$

Mean absolute Error (MAE)

- Average absolute error
- Prevents compensation of positive and negative errors, but depending on scaling of the data and therefore difficult to compare
- All errors equally weighted

$$MAE = \frac{\sum_{t=k+1}^{n} |e_t|}{n-k}$$

Source: Chopra und Meindl (2015); Camm et al (2015)

Forecast Error II/III

Nomenclature

Values:

y_t – *observed value in period* t

$\hat{y}_{t,t+x}$ – *predicted value for period* $t+x$ *on the basis of the values of period* t

e – *Error*

Indices:

n – *Number of all periods*

P – *Number of all periods considered*

k – *Number of periods, for which no prediction is possible*

t – *Control variable of the sum (considered period)*

Mean squared Error (MSE)

- Average square error
- Prevents compensation of positive and negative errors, but depending on scaling of the data and therefore difficult to compare
- Larger errors are weighted more heavily due to squaring

$$MSE = \frac{\sum_{t=k+1}^{n} e_t^2}{n-k}$$

The square error has a unit that cannot be interpreted, therefore the root mean squared error (RSME) is also used:

$$\text{RSME} = \sqrt{\frac{\sum_{t=k+1}^{n} e_t^2}{n-k}}$$

Mean Absolute Percentage Error (MAPE)

- Average percentage error
- Enables comparability with different scaling
- All errors equally weighted

$$MAPE = \frac{\sum_{t=k+1}^{n} \left| \left(\frac{e_t}{y_t}\right) * 100 \right|}{n-k}$$

For observations with value 0, the percentage error cannot be calculated. Therefore the symmetric mean absolute percentage error (sMAPE) is also used:

$$\text{sMAPE} = \frac{1}{n-k} * \sum_{t=k+1}^{n} \frac{|e_t|}{\frac{y+\hat{y}}{2}}$$

Source: Chopra und Meindl (2015); Camm et al (2015)

Calculation in Excel (See Problem Set 6-0)

	A	B	C	D	E	F	G
1			Naives Modell				
2	Periode	y_t - Beob. Bedarf	y_t-1,t	e_t	|e_t|	(e_t)^2	|(e_t/y_t)*100|
3	1	376	###	###	###	###	###
4	2	403	376	=C4-$B4	=ABS(D4)	=D4^2	=ABS((D4/$B4)*100)
5	3	391	403	12	12	144	3,069053708
6	4	416	391	-25	25	625	6,009615385
7	5	395	416	21	21	441	5,316455696
8	6	417	395	-22	22	484	5,275779376
9	7	417	417	0	0	0	0
10	8	417	417	0	0	0	0
11	9	435	417	-18	18	324	4,137931034
12	10	403	435	32	32	1024	7,94044665
13	11	391	403	12	12	144	3,069053708
14	12	452	391	-61	61	3721	13,49557522
15	13	411	452	41	41	1681	9,9756691
16	14	403	411	8	8	64	1,985111663
17	15	372	403	31	31	961	8,333333333
18	16	409	372	-37	37	1369	9,046454768
19	17	368	409	41	41	1681	11,14130435
20	18	399	368	-31	31	961	7,769423559
21	19	408	399	-9	9	81	2,205882353
22	20	440	408	-32	32	1024	7,272727273
23	21	369	440	71	71	5041	19,24119241
24	22	390	369	-21	21	441	5,384615385
25	23	407	390	-17	17	289	4,176904177
26	24	378	407	29	29	841	7,671957672
27	25	405	378	-27	27	729	6,666666667
28							
29			*MFE*	=MITTELWERT(D4:D27)			
30			*MAE*		=MITTELWERT(E4:E27)		
31			*MSE*			=MITTELWERT(F4:F27)	
32			*MAPE*				=MITTELWERT(G4:G27)

Problem 11-1

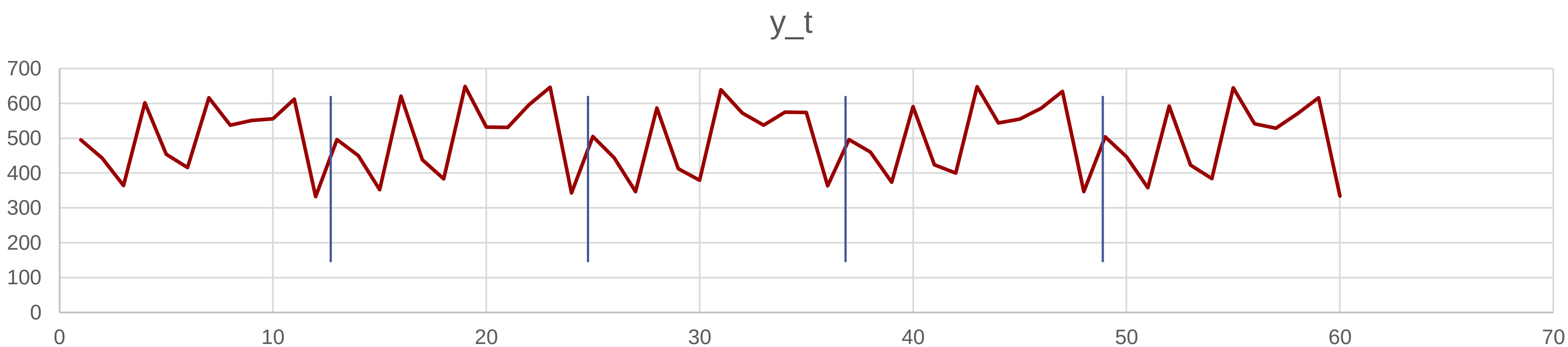
- BEAR Furniture manufactures the "Balu" chest of drawers for several branches of a leading furniture store. So far, the company has produced the average monthly demand (500 units). However, if demand deviates, this generates high costs:
  - (1) Chests of drawers that are produced too much generate storage costs per piece of 9
  - (2) insufficiently produced chests of drawers generate contractual penalties (discount of 15 per piece) for the delayed delivery time.
- Therefore BEAR Furniture wants to develop a prediction model for the demand. You have the sales data for the last 60 months.

a) Set up a naive model, overall historical model, a moving average (P=3) and moving average (P=6) and compare these with regard to the prediction accuracy!

b) Which model would minimize costs?

Solution 11-1

- The MA3 model achieves the best overall values for MFE, MAE and MAPE. The best MSE in overall history. (values see Excel file)
- The demand shows a clear seasonality, while the available models are suitable for a relatively stable level. None of the models is suitable for this problem.



- On average, costs of 85 euros per period are avoided.

Summary: Time series analysis

- An observed value should be divided into a systematic component and a random component. The aim of the model is to minimize the error to the level of the random component.
  - The error will never be zero and the prediction will never be accurate.
- Simple methods suitable for the time series pattern level are the Naïve model, the average of the entire history and the moving average.

The following topic:

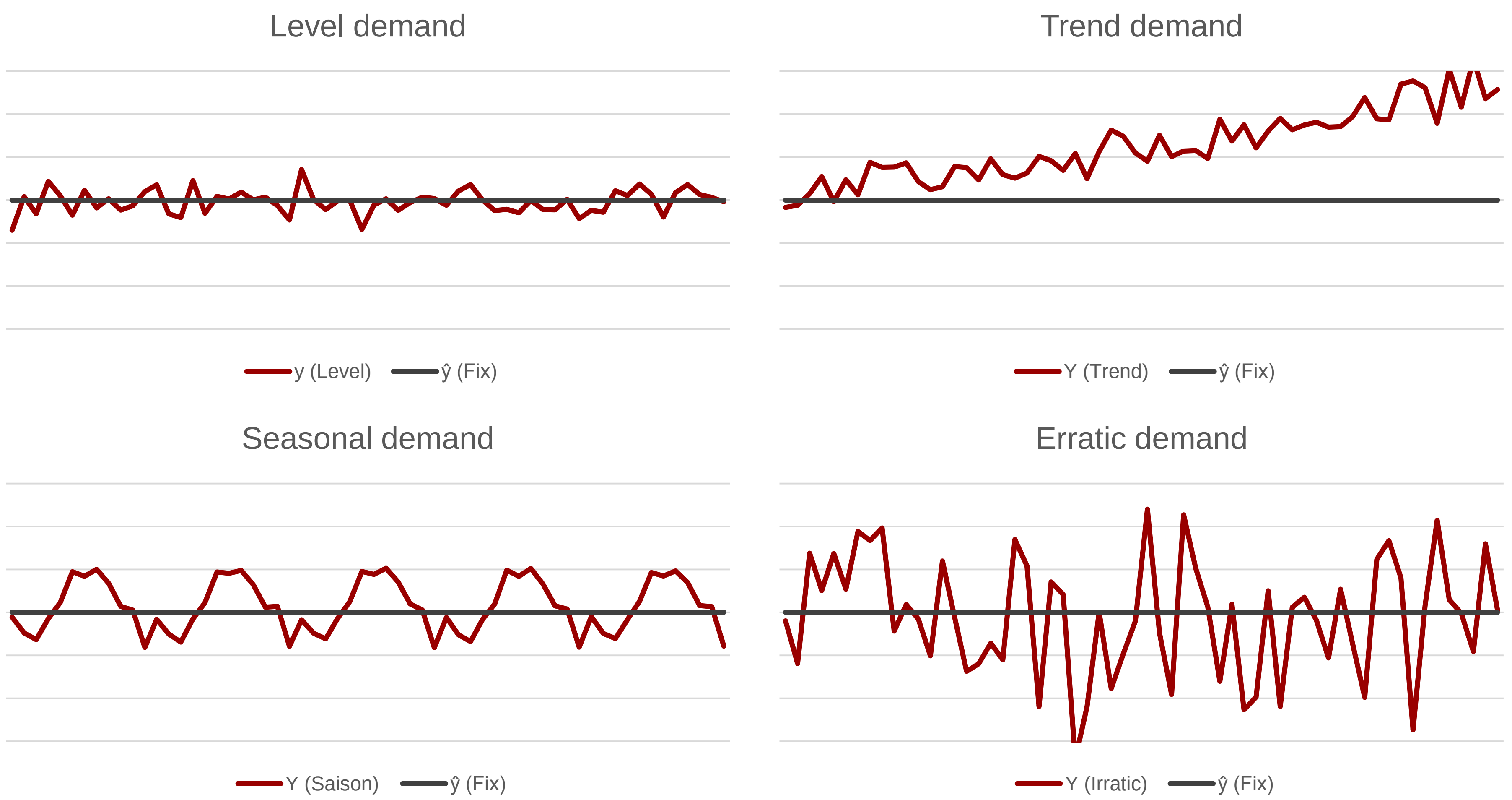
Based on the knowledge of simple models, models with **assumed time series patterns** are considered

Agenda

1. Demand forecasting
2. Time series analysis I – naive and simple
3. Time series analysis II – pattern recognition

The effects of different models on different demand patterns I

Baseline: static prediction



Simple exponential smoothing

Nomenclature

Values:

y_t – observed value in period t

$\hat{y}_{t,t+x}$ – predicted value for period $t + x$ on the basis of the values of period t

e – Error

α – Smoothing constant $(0 \leq \alpha \leq 1)$

Indices:

n – Number of all periods

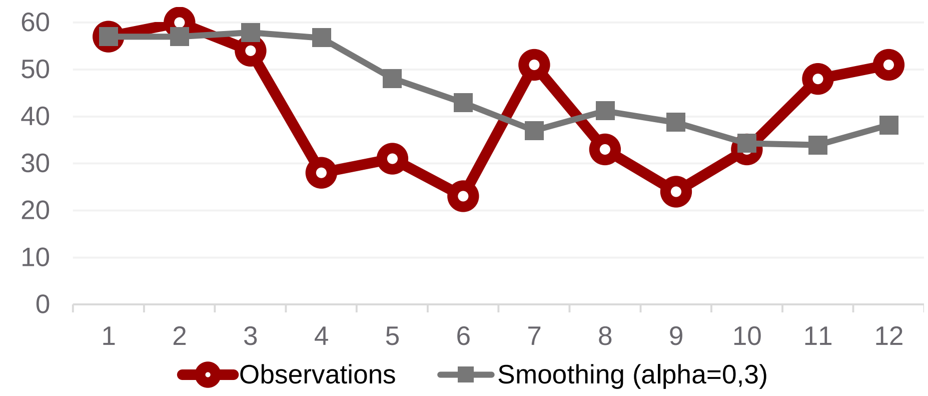
k – Number of periods, for which no prediction is possible

t – Control variable of the sum (considered period)

Simple exponential smoothing

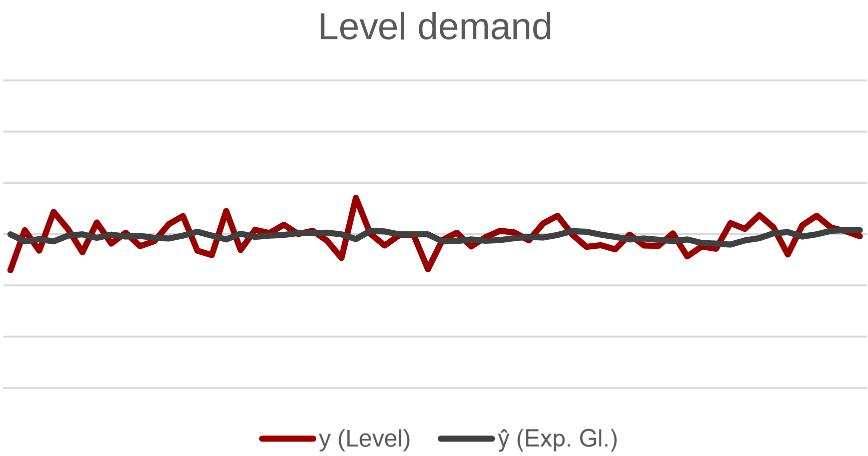
- Exponential smoothing based on weighted average of the past
- Suitable for demand patterns: level
- Prediction ($\hat{y}_{t,t+1}$ for period $t + 1$ **weights** the observation (y_t) in period t with smoothing constant (α and the prediction of period t with counterweight $(1 - \alpha)$.
  - $\hat{y}_{t,t+1} = \alpha * y_t + (1 - \alpha) * \hat{y}_{t-1,t}$

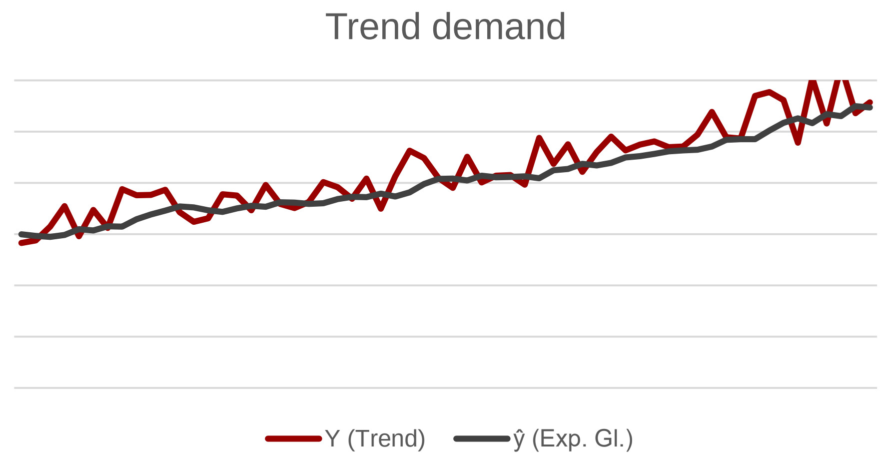
Accordingly, the observation from the period $t - 1$ is weighted with $\alpha * (1 - \alpha)$ etc.

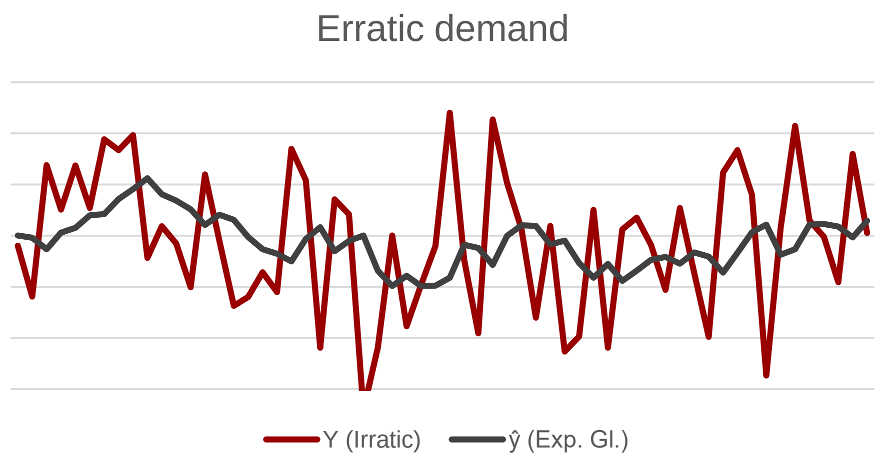


The effects of different models on different demand patterns II

Exponential smoothing: Decreasingly weighted past

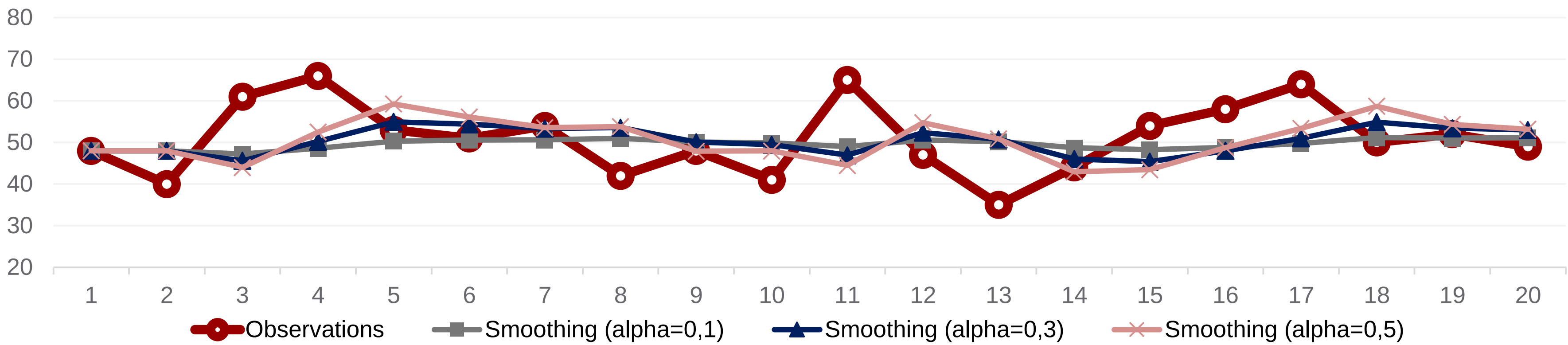






Smoothing constant (exponential smoothing)

- The smoothing constant should be determined by minimizing the prediction error (MSE).
  - In a (non-linear) optimization the error would be used as target variable, α as decision variable and $0 \leq \alpha \leq 1$ as constraint (see VL10 - Prescriptive Analytics 1).
- Depending on the source, α is recommended by 0.2 to 0.3
  - The Data Analysis Add-In of Excel calls $(1-\alpha)$ a smoothing parameter!



MSE = 79,58 MSE = 88,37 MSE = 96,51

Optimal MSE (alpha = 0,042) = 78,11

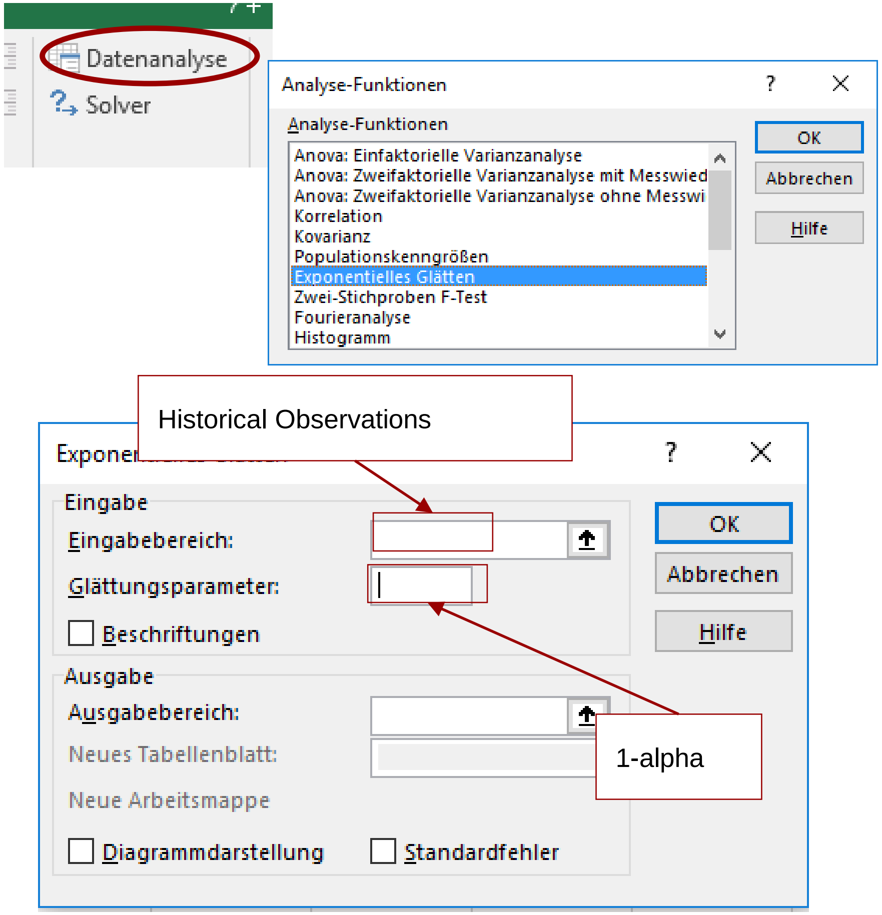
Exponential Smoothing - Example

Manual calculation

	A	B	C
1		*alpha*	0,2
2			
3	**Periode**	**y**	**ŷ (EG)**
4	1	48	48
5	2	40	48
6	3	61	46,4
7	4	66	49,32
8	5	53	52,656
9	6	51	52,7248
10	7	54	52,37[illegible]84
11	8	42	52,703872
12	9	48	50,5630976
13	10	41	50,050[illegible]
14	11	65	48,240[illegible]
15	12	47	51,59[illegible]
16	13	35	50,6738448
17	14	44	47,5390758
18	15	54	46,8312607
19	16	58	48,2650085
20	17	64	50,2120068
21	18	50	52,9696055
22	19	52	52,3756844
23	20	49	52,3005475

=$C$1*B8+(1-$C$1)*C8

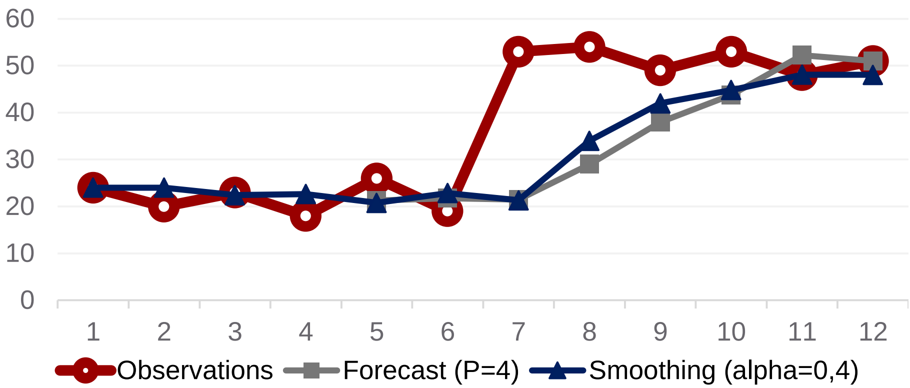
Calculation with Excel data analysis



Problems of level models with trends and jumps in demand

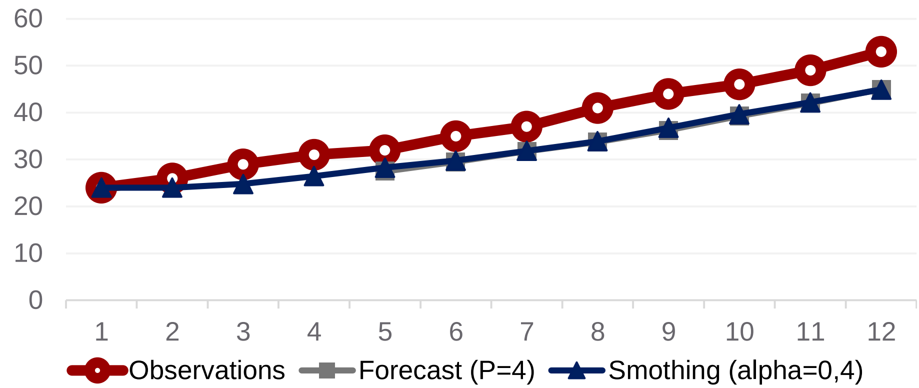
Case 1: Changing the level

→ The adjustment is very slow.



Case 2: Simple trend

→ The trend is always underestimated by the forecast



Source: Chopra und Meindl (2015); Camm et al (2015)

Trend-corrected exponential smoothing (Holt-method)

Nomenclature

Values:

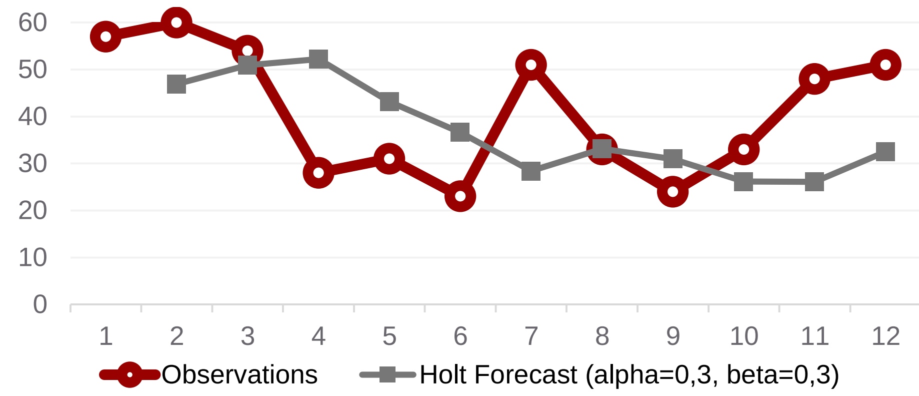
y_t – observed value in period t
$\hat{y}_{t,t+x}$ – predicted value for period $t+x$ on the basis of the values of period t
e – Error
α – Smoothing constant level $(0 \le \alpha \le 1)$
β – Smoothing constant level $(0 \le \beta \le 1)$
L – Level prediction
T – Trend prediction

Indices:

n – Number of all periods
k – Number of periods, for which no prediction is possible
t – Control variable of the sum (considered period)

Holt-method

- Based on exponential smoothing → uses smoothing constant (α)
- Suitable for demand patterns: Level (L) + Trend (T)
- The smoothing constant for level (α) is supplemented by a second for trend (β).
  - $L_t = \alpha * y_t + (1-\alpha) * (L_{t-1} + T_{t-1})$
  - $T_t = \beta * (L_t - L_{t-1}) + (1-\beta) * T_{t-1}$
- Samples are calculated individually and combined for prediction.
  - $\hat{y}_{t,t+1} = L_t + T_t$
- Forecast for several periods x
  - $\hat{y}_{t,t+x} = L_t + x * T_t$

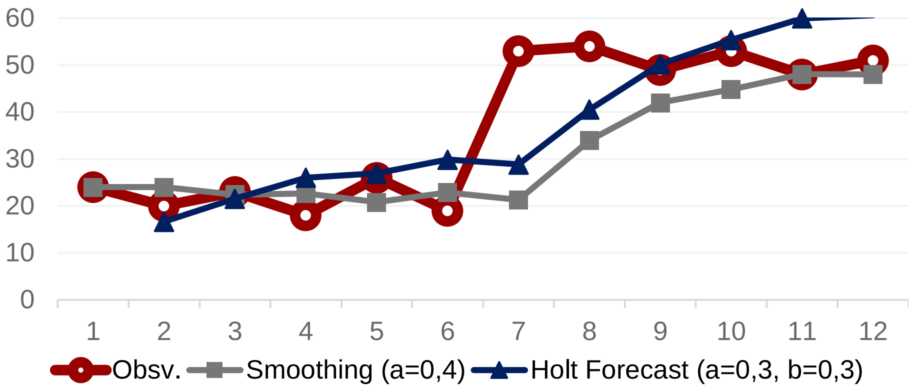


Source: Chopra und Meindl (2015); Camm et al (2015)

Comparison of level with level + trend model

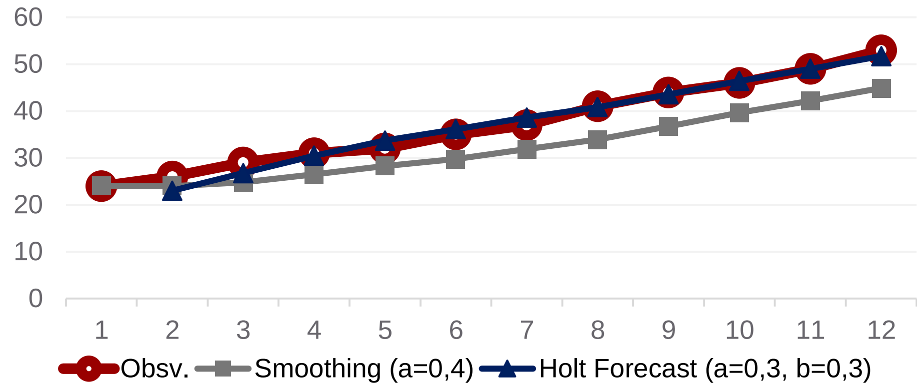
Case 1: Changing the level

→ the holt method adapts faster
→ overestimates the demand for a certain period of time



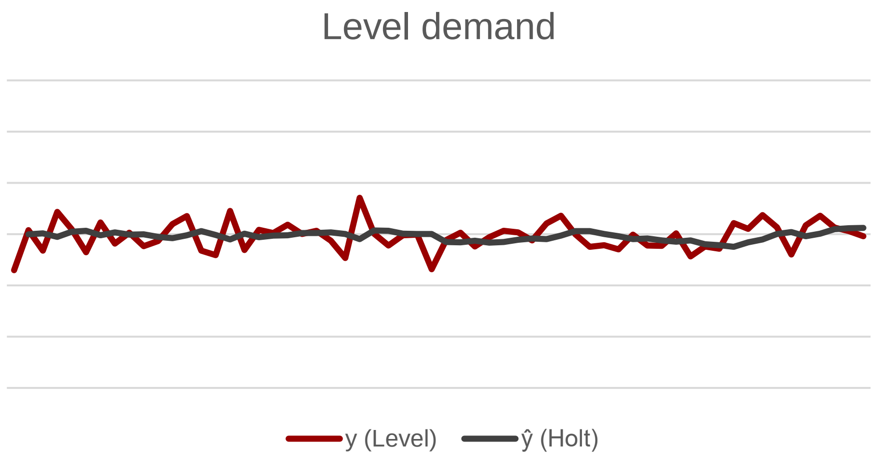
Case 2: Simple trend

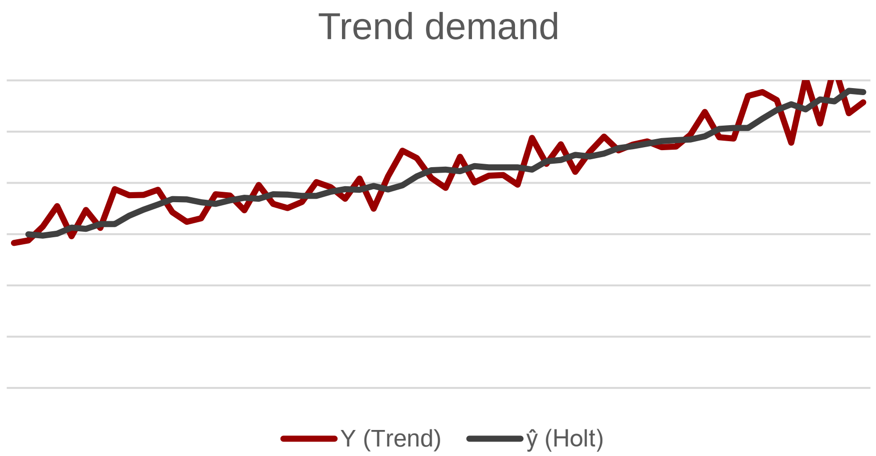
→ Accuracy of the Holt model is clearly superior

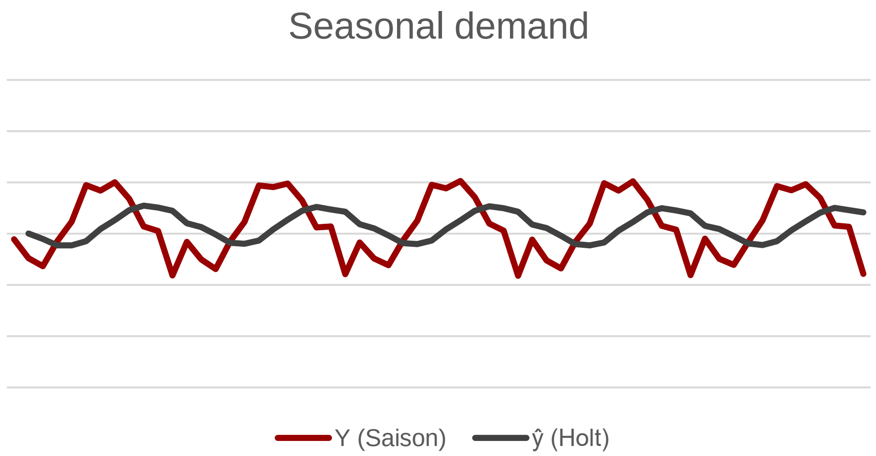


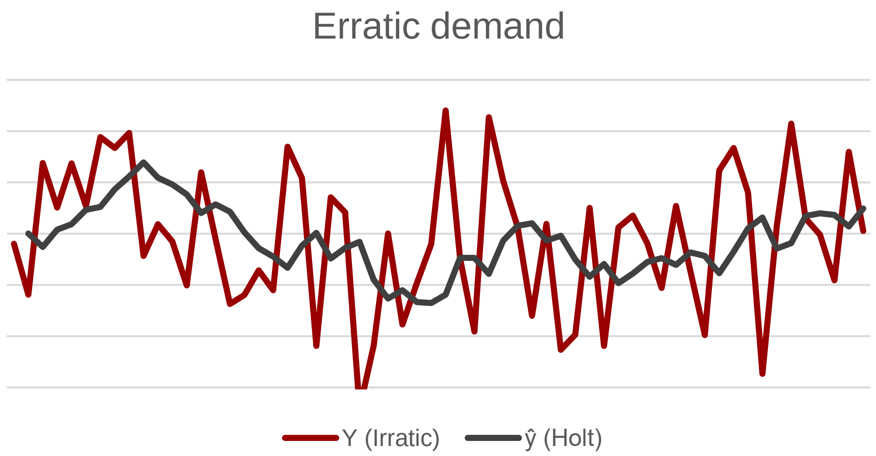
Source: Chopra und Meindl (2015); Camm et al (2015)

Holt Model: Consideration of trends









Smoothing constant and initiation values (Holt)

- Smoothing constants $\{\alpha,\beta\}$ should be determined by minimizing the prediction error (MSE)
  - In a (non-linear) optimization the error would be used as target variable, α and β as decision variables and $0 \leq \alpha \leq 1$ and $0 \leq \beta \leq 1$ as constraint (see VL11 - Prescriptive Analytics 1).

- The initial values for level and trend patterns can be determined by a simple linear regression.
  - In a linear regression the observations (y) would be the dependent variable, the periods (t) the independent variables. The intersection represents the initial value for L and T the slope the initial value for T

Source: Chopra und Meindl (2015); Camm et al (2015)

Holt Model - Example

	A	B	C	D	E
1		*alpha*	0,2		
2		*beta*	0,2		
3	**Periode**	**y**	**ŷ (Ht)**	**L_t**	**T_t**
4	1	48		51	0,02
5	2	40	51,02	48,816	-0,4208
6	3	61	48,3952	50,91616	0,083392
7	4	66	50,999552	53,9996416	0,68340992
8	5	53	54,6830515	54,34644122	0,616087859
9	6	51	54,9625291	54,17002326	0,457586696
10	7	54	54,62761	54,50208797	0,432482298
11	8	42	54,9345703	52,34765621	-0,084900513
12	9	48	52,2627557	51,41020456	-0,25541074
13	10	41	51,1547938	49,12383505	-0,661602493
14	11	65	48,4622326	51,76978605	-9,17956E-05
15	12	47	51,7696943	50,8157554	-0,190879566
16	13	35	50,6248758	47,49990067	-0,815874599
17	14	44	46,6840261	46,14722086	-0,923235642
18	15	54	45,2239852	46,97918817	-0,572195051
19	16	58	46,4069931	48,7255945	-0,108474775
20	17	64	48,6171197	51,69369578	0,506840436
21	18	50	52,2005362	51,76042897	0,418818987
22	19	52	52,179248	52,14339837	0,411649069
23	20	49	52,5550474	51,84403795	0,269447172

=D4+E4

=$C$1*B11+(1-$C$1)*C11

=$C$2*(D17-D16)+(1-$C$2)*E16

Problem 11-2

- BEAR Furniture manufactures the "Balu" chest of drawers for several branches of a leading furniture store. So far, the company has produced the average monthly demand (500 units). If demand deviates, however, this generates high costs, because too much produced chests of drawers cause costs of 9 per piece for storage and too little produced chests of drawers cause contractual penalties (discount of 15 per piece) for the delayed delivery time. Therefore BEAR Furniture wants to develop a prediction model for the demand. You have the sales data for the last 60 months.
- Set up a model after exponential smoothing (alpha=0.2) and a Holt-model (alpha=0.3, beta = 0.3) and compare these with regard to the prediction accuracy!

a) Which model would minimize costs?

Solution 11-2

a) Both models do not achieve any improvement compared to the moving average (P=2)

- Exponential smoothing: MSE = 11793.78522; MAPE = 19.93476512
- Holt Model: MSE = 115234.49112; MAPE = 20.61631247

b) Both models do not improve costs

- Fixed production: Costs = 1044.4 per period
- Exponential smoothing: Costs = 1101.47 per period
- Holt Model: Cost = 1138.26 per period

- Both models are unsuitable for dealing with strong fluctuations / seasonality

Trend and seasonally corrected exponential smoothing (Holt-Winter method)

Nomenclature

Values:

y_t – *observed value in period t*
$\hat{y}_{t,t+x}$ – *predicted value for period* $t + x$ *on the basis of the values of period t*
e – *Error*
α – *Smoothing constant level* $(0 \leq \alpha \leq 1)$
β – *Smoothing constant level* $(0 \leq \beta \leq 1)$
γ – *Smoothing constant season* $(0 \leq \gamma \leq 1)$
L – Level prediction
T – Trend prediction
S – Seasonal adjustment

Indices:

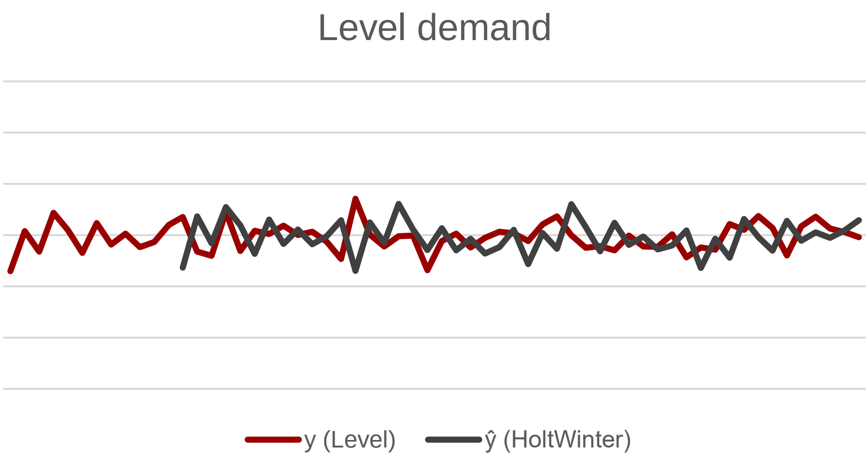
t – *Control variable of the sum (considered period)*
p – *Number of periods of a season*

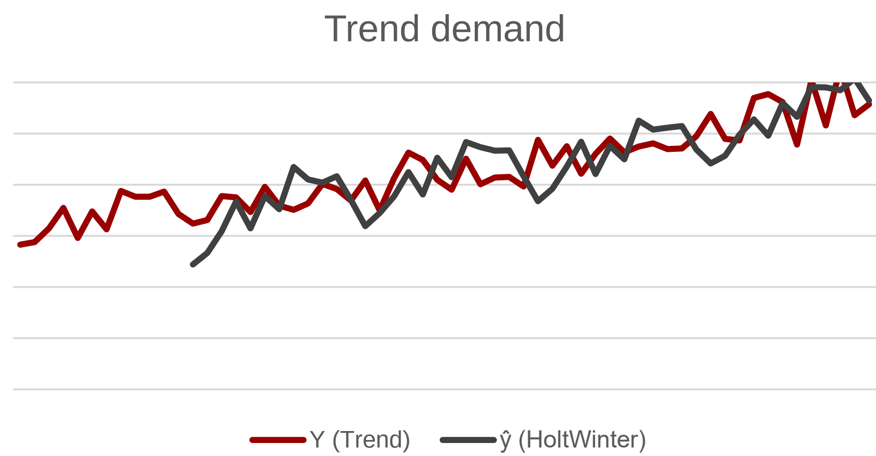
Holt-Winter method

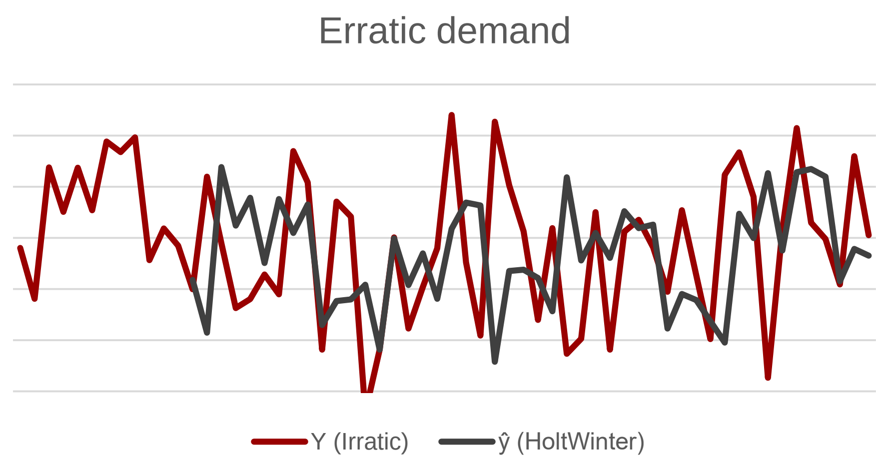
- Based on the holt method → uses smoothing constants $\{\alpha,\beta\}$
- Suitable for demand patterns: (Level (L) + Trend (T)) + Seasonality (S)
- Smoothing constants for level (α) and trend (β) are supplemented by third for season (γ).
  - $L_t = \alpha * (y_t - S_{t-p}) + (1 - \alpha) * (L_{t-1} + T_{t-1})$
  - $T_t = \beta * (L_t - L_{t-1}) + (1 - \beta) * T_{t-1}$
  - $S_t = \gamma * (y_t - L_t) + (1 - \gamma) * S_{t-p}$
- The samples are calculated individually and put together for the prediction.
  - $\hat{y}_{t,t+1} = L_t + T_t + S_{t+1-p}$
- Forecast for several periods x
  - $\hat{y}_{t,t+x} = L_t + x * T_t + S_{t+x-p}$

The effects of different models on different demand patterns IV

Holt-Winter: Consideration of trends and seasonality







Trend-corrected exponential smoothing (Holt-Winter method)

- The smoothing constant should be determined by minimizing the prediction error (MSE).
  - In a (non-linear) optimization, the error would be used as a target variable, α, β and γ as decision variables and 0≤α≤1, 0≤β≤1 and 0≤γ≤1 as secondary condition (see VL10 - Prescriptive Analytics 1).

- The initial values for level and trend patterns can be determined by a simple linear regression.
  - In a linear regression the observations (y) would be the dependent variable, the periods (t) the independent variables. The intersection represents the initial value for L and T the slope the initial value for T

- The initial values for seasonality patterns can be determined by the deviation from the mean value of the first period set. This means that no forecast error can be determined for the first p values.
  - $S_t = y_t - \frac{\sum_{i=1}^{p} y_i}{p} \quad t \in \{1, 2, 3, \ldots, p\}$

Problem 11-3

Example - Solution is presented

- BEAR Furniture manufactures the "Balu" chest of drawers for several branches of a leading furniture store. So far, the company has produced the average monthly demand (500 units). If demand deviates, however, this generates high costs, because too much produced chests of drawers cause costs of 9 per piece for storage and too little produced chests of drawers cause contractual penalties (discount of 15 per piece) for the delayed delivery time. Therefore BEAR Furniture wants to develop a prediction model for the demand. You have the sales data for the last 60 months.

a) Set up model after holt-winter (alpha=0.2; beta=0.2; gamma=0.2) and determine the prediction accuracy!

b) Which model would minimize costs?

Solution 11-3 I/III

Example - Solution is presented

	A	B	C	D	E	F
1			alpha	0,2		
2			beta	0,3		
3			gamma	0,2		
4			**Holt-Winter Modell**			
5	***Periode***	***y_t***	***ŷ_t (HW)***	***L_t***	***T_t***	***S_t***
6	1	495	#NV	#NV	#NV	-3,16666667
7	2	443	#NV	#NV	#NV	-55,1666667
8	3	364	#NV	#NV	#NV	-134,166667
9	4	602	#NV	#NV	#NV	103,833333
10	5	454	#NV	#NV	#NV	-44,1666667
11	6	416	#NV	#NV	#NV	-82,1666667
12	7	616	#NV	#NV	#NV	117,833333
13	8	537	#NV	#NV	#NV	38,8333333
14	9	551	#NV	#NV	#NV	52,8333333
15	10	556	#NV	#NV	#NV	57,8333333
16	11	612	#NV	#NV	#NV	113,833333
17	12	332	#NV	493	0,2	-166,166667
18	13	496	490,033333	494,393333	0,558	-2,212
19	14	450	439,784667	496,9944	1,17092	-53,5322133
20	15	352	363,998653	495,765589	0,4510008	-136,086451
21	16	621	600,049923	500,406605	1,70800539	107,185346
22	17	438	457,947944	498,125022	0,51112874	-47,3583377
23	18	383	416,469484	491,942254	-1,4970403	-87,5217841
24	19	649	608,278547	498,589504	0,94624688	124,348766
25	20	532	538,369084	498,261934	0,56410181	37,8142798
26	21	531	551,659369	494,694162	-0,67546035	49,5278342
27	22	596	551,852035	502,848295	1,97341754	64,8970077
28	23	646	618,655046	510,290703	3,6141148	118,208526
29	24	343	347,738151	512,957188	3,32982572	-166,924771
30	25	505	514,075013	514,472011	2,78532492	-3,66400215

Initial values from deviation from the mean value of the first 12 period

Initial values from linear regression

$$\hat{y}_{t,t+1} = L_t + T_t + S_{t+1-p}$$

=D17+E17+F6

Solution 11-3 II/III

Example - Solution is presented

	A	B	C	D	E	F
1			alpha	0,2		
2			beta	0,3		
3			gamma	0,2		
4			**Holt-Winter Modell**			
5	***Periode***	***y_t***	***ŷ_t (HW)***	***L_t***	***T_t***	***S_t***
6	1	495	#NV	#NV	#NV	-3,16666667
7	2	443	#NV	#NV	#NV	-55,1666667
8	3	364	#NV	#NV	#NV	-134,166667
9	4	602	#NV	#NV	#NV	103,833333
10	5	454	#NV	#NV	#NV	-44,1666667
11	6	416	#NV	#NV	#NV	-82,1666667
12	7	616	#NV	#NV	#NV	117,833333
13	8	537	#NV	#NV	#NV	38,8333333
14	9	551	#NV	#NV	#NV	52,8333333
15	10	556	#NV	#NV	#NV	57,8333333
16	11	612	#NV	#NV	#NV	113,833333
17	12	332	#NV	493	0,2	-166,166667
18	13	496	490,033333	494,393333	0,558	-2,212
19	14	450	439,784667	496,9944	1,17092	-53,5322133
20	15	352	363,998653	495,765589	0,4510008	-136,086451
21	16	621	600,049923	500,406605	1,70800539	107,185346
22	17	438	457,947944	498,1[illegible]5022	0,51112874	-47,3583377
23	18	383	416,469484	491,942254	-1,4970403	-87,5217841
24	19	649	608,278547	498,589504	0,94624688	124,348766
25	20	532	538,369084	498,261934	0,56410181	37,8142798
26	21	531	551,659369	494,694162	-0,67546035	49,5278342
27	22	596	551,852035	502,848295	1,97341754	64,8970077
28	23	646	618,655046	510,290703	3,6141148	118,208526
29	24	343	347,738151	512,957188	3,32982572	-166,924771
30	25	505	514,075013	514,472011	2,78532492	-3,66400215

$$L_t = \alpha * (y_t - S_{t-p}) + (1-\alpha) * (L_{t-1} + T_{t-1})$$

=$C$1*(B18-F6)+(1-$C$1)*(D17+E17)

$$T\_t = \beta * (L\_t - L_{t-1}) + (1-\beta) * T_{t-1}$$

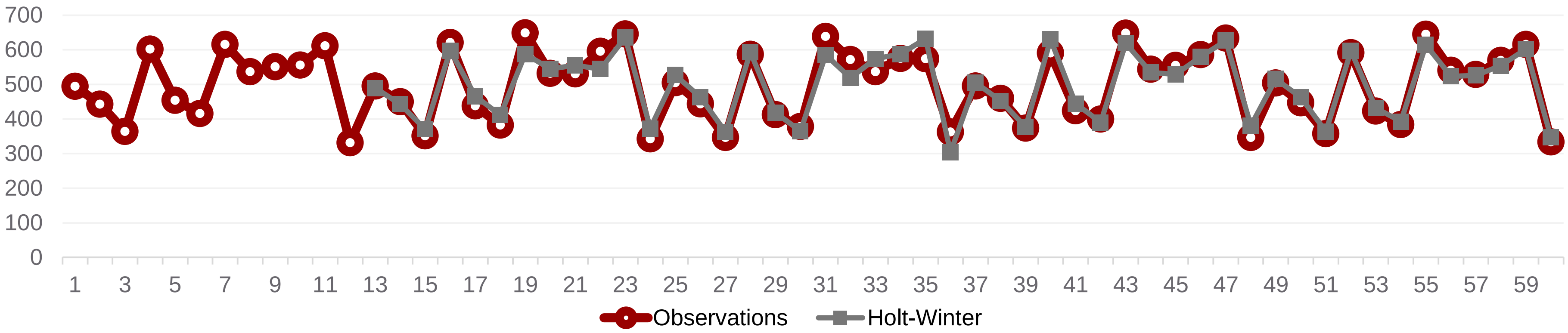
=$D$2*(D26-D25)+(1-$D$2)*E25

$$S_t = \gamma * (y_t - L_t) + (1-\gamma) * S_{t-p}$$

=$D$3*(B30-D30)+(1-$D$3)*F18

Solution 11-3 III/III

- The Seasonal Model achieves significant improvements in forecasting accuracy and costs
  - MSE: 520,8296592
  - MAPE: 3,943688326
  - Costs per period: 231,06



Problem 11-4

- You will receive the orders received from an anonymous UK-based online shop (which primarily supplies wholesalers). You should estimate the incoming orders for the weekdays of next week, so that the team supervisors can plan the necessary number of employees. To do this, proceed as follows:

a) Clean up the data set of purchase orders with negative order quantities. (In Excel: Select all negative orders and use the delete key. deleting the cells leads to a crash with the amount of data)

b) Aggregate the orders per day (Excel: Pivot tables of the date automatically aggregate to months, therefore proceed as follows:)

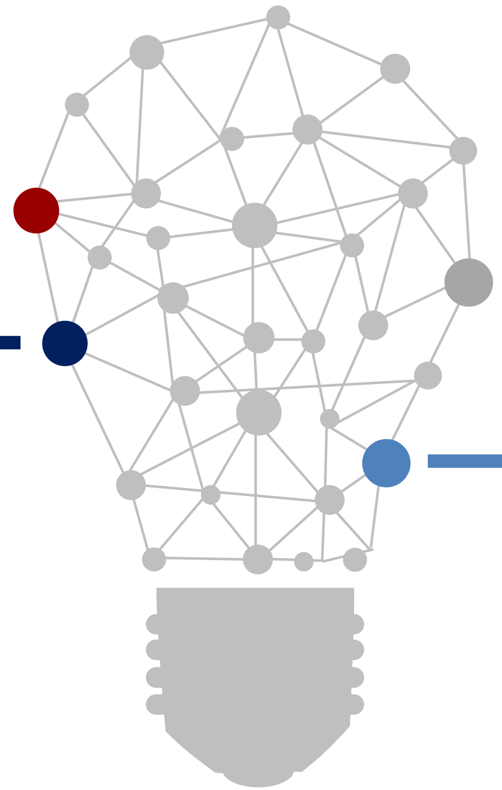
   (1) Create the columns Year, Calendar Week and Weekday. Use the appropriate Excel functions.

   (2) Use text concatenation of the new columns in a new fourth column

   (3) Aggregate the purchase orders in a pivot table with the new column

c) Copy the aggregation and create a suitable model for the problem in question.

Thank you
For your attention



Comments and directions
Will be greatly appreciated

Questions
Are welcome

Stay in contact
herden@logistik.tu-berlin.de
Tino Herden

References

- Baesens, B. (2014): Analytics in a Big Data World.
- Camm, J.D.; Cochran, J.J.; Fry, M.J.; Ohlmann, J.W.; Anderson, D.R.; Sweeney, D.J.; Williams, T.A. (2015): Essentials of Business Analytics.
- Chopra, S.; Meindl, P. (2015): Supply Chain Management - Strategy, Planning, and Operation. 6^{th} Ed.
- Das, S. (2014): Computational Business Analytics.
- Franks, B. (2014): The analytics revolution: how to improve your business by making analytics operational in the big data era.
- Provost, F., Fawcett, T. (2013): Data Science for Business.
- Siegel, E. (2013): Predictive Analytics.