TIMESERIES FORECASTING (A)

ALEXANDROS PATELIS (CR/AEU2)

BOSCH

Timeseries Forecasting (A)
Recap from 1st session

1. Introduction:
   - Definitions, Where to forecast, Timeseries examples.
2. Timeserie Characteristics & Components
3. Timeserie Statistical Analysis
4. Timeserie Preparation & Analysis:
   - Volume, Missing values, Adjustments, Special Events and Actions.
5. Forecasting Types
6. Forecasting Errors & Accuracy
7. Statistical forecasting methods – Basic:
   - Naïve simple methods, Moving Averages, Decomposition, Linear Regression, Exponential Smoothing.
8. Method selection:
   - Before selecting a method, finding the proper method, Experts adjustments, Update Forecasts.
9. Interesting use cases

© Robert Bosch GmbH 2018. All rights reserved, also regarding any disposal, exploitation, reproduction, editing, distribution, as well as in the event of applications for industrial property rights.

Timeseries Forecasting (A)
Agenda

10. Statistical Methods – Advanced:
    - Theta method, Alternative Theta, Averaging, Arima, Multiple Regression.
    - Neural Networks.
11. Judgmental forecasting methods:
    - Simple Judgment, Delphi method, Analogy, Structured Analogy, Scenario forecasting.
12. Special topic:
    - Intermitted demand
    - Croston, Synetos & Boylan, ADIDA.
13. Special topic:
    - Structural changes detection.
14. Interesting use cases
    - Inventory Demand Forecasting, Energy Management, Social Media, Battery data analysis, etc.
15. Proposed Links & Literature

© Robert Bosch GmbH 2018. All rights reserved, also regarding any disposal, exploitation, reproduction, editing, distribution, as well as in the event of applications for industrial property rights.

BOSCH

10. STATISTICAL FORECASTING METHODS

ADVANCED

THETA METHOD

Winner of the M3 Forecasting Competition (1999)

Held by INSEAD (Prof. S. Makridakis)

Forecasting Methods
Theta method: Model (1)

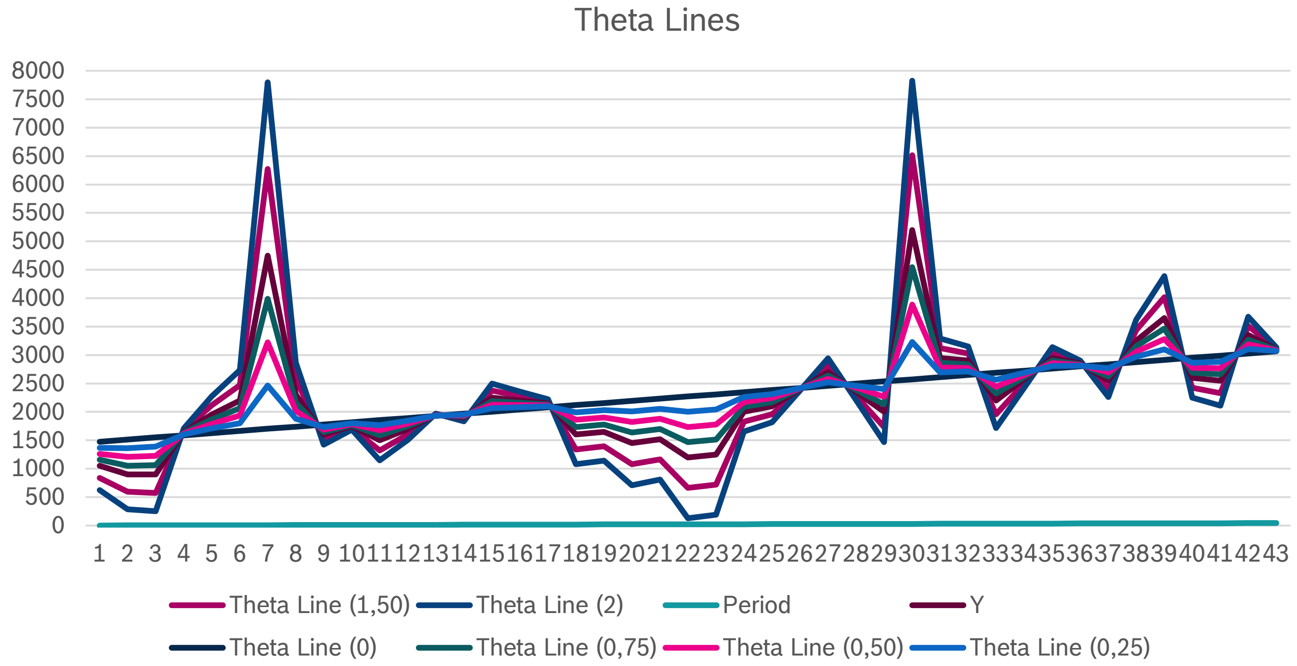
- A new ***univariate forecasting method***, introduced in 1999 (*Assimakopoulos et. al.*), based:
  - on the decomposition model, and
  - on the concept of modifying ***the local curvatures*** of the timeseries through a coefficient "***Theta***" (the Greek letter θ).
  - The initial timeserie is spitted into ***2 or more*** Theta Lines, by using a θ parameter.
  - The resulting series, the "Theta-lines" maintain the mean and the slope of the original data, but not their local curvatures.
  - The ***basic characteristic*** of this method is:
    - a better approach of the long-term behaviour, or
    - the emphasis of short characteristics,

    depending on the value of the Θ parameter (< 1 or > 1).

© Robert Bosch GmbH 2018. All rights reserved, also regarding any disposal, exploitation, reproduction, editing, distribution, as well as in the event of applications for industrial property rights.
BOSCH

Forecasting Methods
Theta method: Model (2)

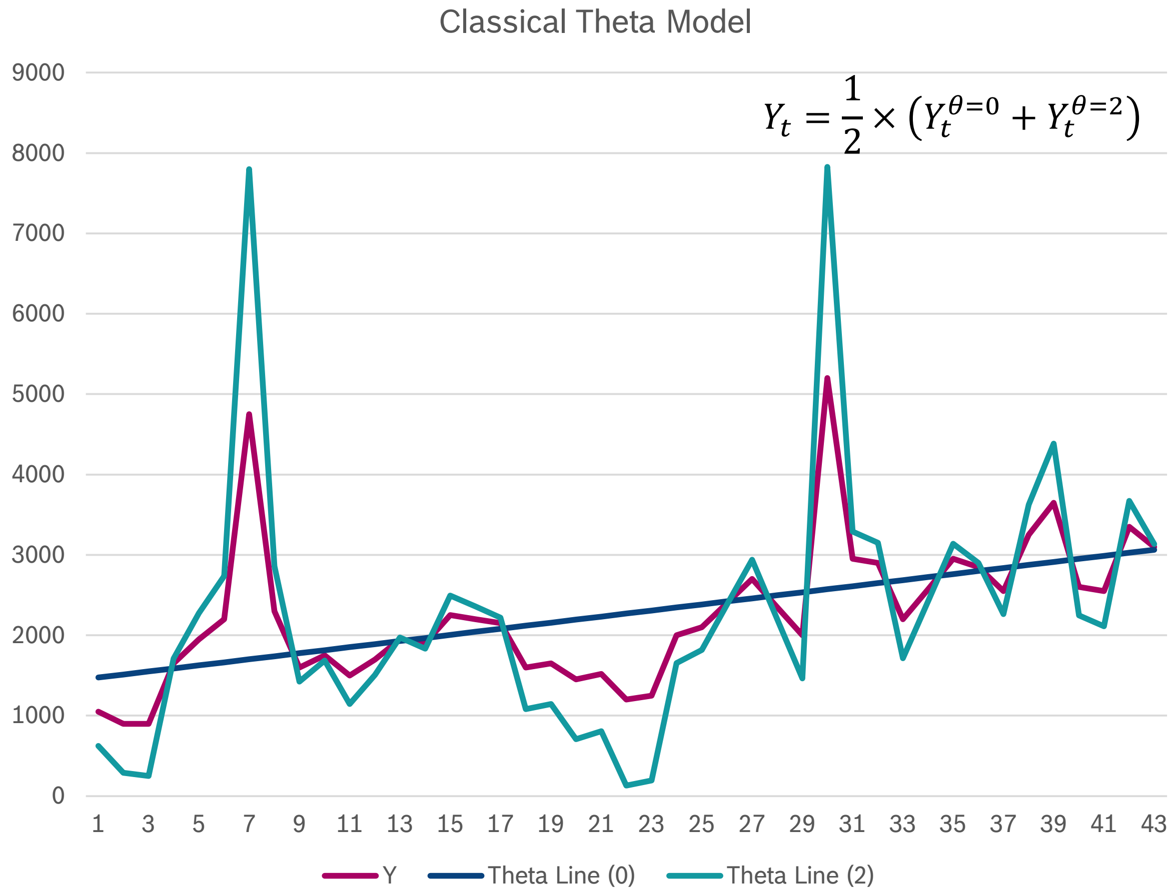
- The Theta method has introduced a different approach in the decomposition theory.
- The challenge for this method was ***to increase the degree of utilization*** of the useful information, which is hidden within the data ***before applying*** an extrapolative model.
- Acts like a magnifying lens through which the fluctuations of timeseries magnified or diminished.
- The linear combination of forecasts of the components, through this process, is more efficient.



Internal | CR/AEU2 - Patelis | 2018-03-14
© Robert Bosch GmbH 2018. All rights reserved, also regarding any disposal, exploitation, reproduction, editing, distribution, as well as in the event of applications for industrial property rights.

Forecasting Methods
Theta method: Basic Model (1)

- Simplest combination of 2 Theta lines, where:
  - Θ = 0 (straight line)
  - Θ = 2 (doubling of local curvatures)
- Each of these ***extends separately*** and their estimations are ***combined***.
- This combination has been used for estimating forecasts for the ***M3 forecasting competition*** (Makridakis et. al. 2000), and it has won the competition.



© Robert Bosch GmbH 2018. All rights reserved, also regarding any disposal, exploitation, reproduction, editing, distribution, as well as in the event of applications for industrial property rights.

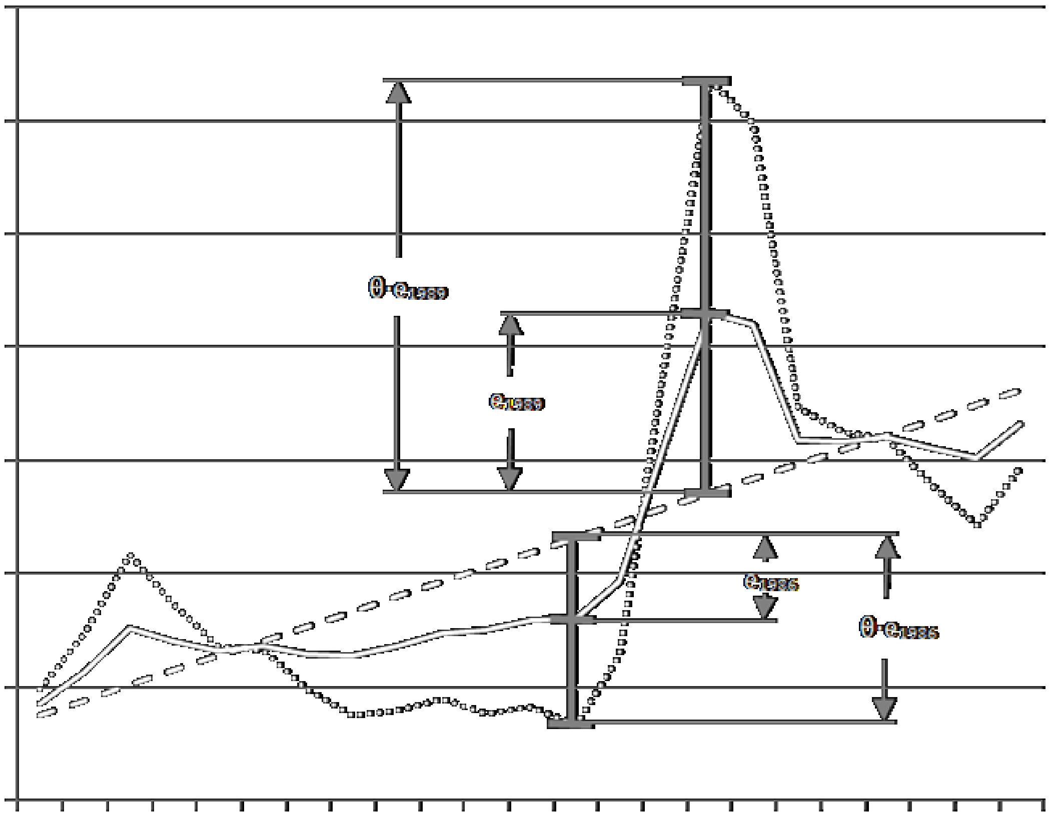
Forecasting Methods
Theta method: Equations

- ***Simplest model***:
  - Theta line (0) = Linear Regression (SLR)
  - Theta line (2) = 2 x Data – SLR
- In General:
  - Theta line (θ) = (θ x Data) + (1-θ) x SLR

or

  - Theta line (θ) = SLR + (θ x e_{SLR})

$$Y_t = \frac{1}{\theta} \times Y_t^{\theta} + (1-\theta) \times Y_t^{0}$$



Internal | CR/AEU2 - Patelis | 2018-03-14

© Robert Bosch GmbH 2018. All rights reserved, also regarding any disposal, exploitation, reproduction, editing, distribution, as well as in the event of applications for industrial property rights.

Forecasting Methods
Theta method: How to use

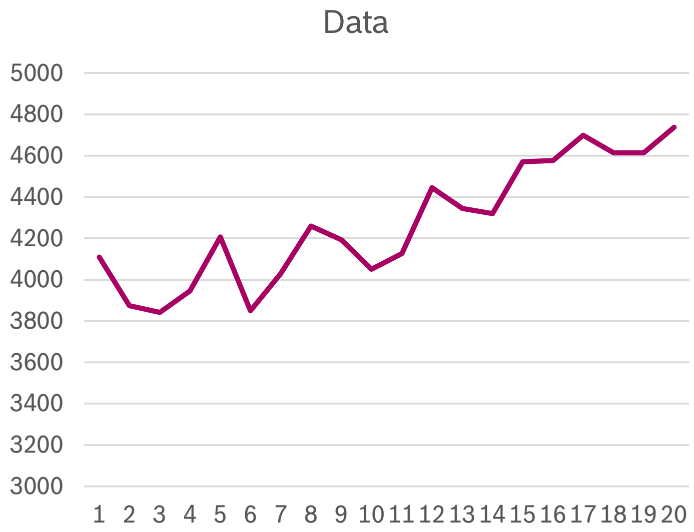
- Step 1: ***Deseasonalization***
  - Test timeserie for seasonal behaviour.
  - The timeserie is de-seasonalized, by using the classical decomposition method.
- Step 2: ***Decomposition & Theta lines***
  - The timeserie is decomposed into two Theta lines, for $\theta=0$ and $\theta=2$.
- Step 3: ***Forecast***
  - The first theta line ($\theta=0$) is extrapolated with simple linear regression (SLR) and the second line ($\theta=2$) with a selected forecasting model (example: simple exponential smoothing).
- Step 4: ***Synthesis***
  - The forecasts for the previous step are combined with equal weights, for estimating the final forecasts.
- Step 5: ***Seasonalization***
  - The final forecasts are seasonalized.

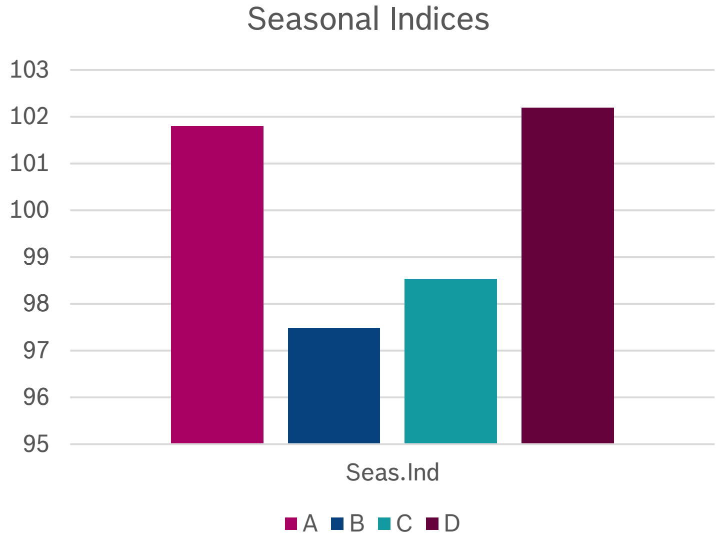
© Robert Bosch GmbH 2018. All rights reserved, also regarding any disposal, exploitation, reproduction, editing, distribution, as well as in the event of applications for industrial property rights.
BOSCH

Forecasting Methods
Theta method: Example (1)

▶ Step 1: ***Deseasonalization***

Period	Data	TxC	SxRx100	I=SxR=S	TxCxR
1	4109			101,79	4036,68
2	3874			97,48	3974,13
3	3842	3955	97,14	98,54	3899,03
4	3946	3964,25	99,54	102,19	3861,43
5	4207	3984,75	105,58	101,79	4132,95
6	3850	4047,5	95,12	97,48	3949,51
7	4030	4085	98,65	98,54	4089,82
8	4260	4108,38	103,69	102,19	4168,71
9	4193	4145,5	101,15	101,79	4119,2
10	4051	4180,63	96,9	97,48	4155,7
11	4126	4222,63	97,71	98,54	4187,24
12	4445	4275	103,98	102,19	4349,74
13	4344	4364,13	99,54	101,79	4267,54
14	4319	4436,13	97,36	97,48	4430,63
15	4571	4496,88	101,65	98,54	4638,85
16	4576	4578,13	99,95	102,19	4477,93
17	4699	4620,25	101,7	101,79	4616,3
18	4614	4645,75	99,32	97,48	4733,25
19	4613			98,54	4681,47
20	4738			102,19	4636,46





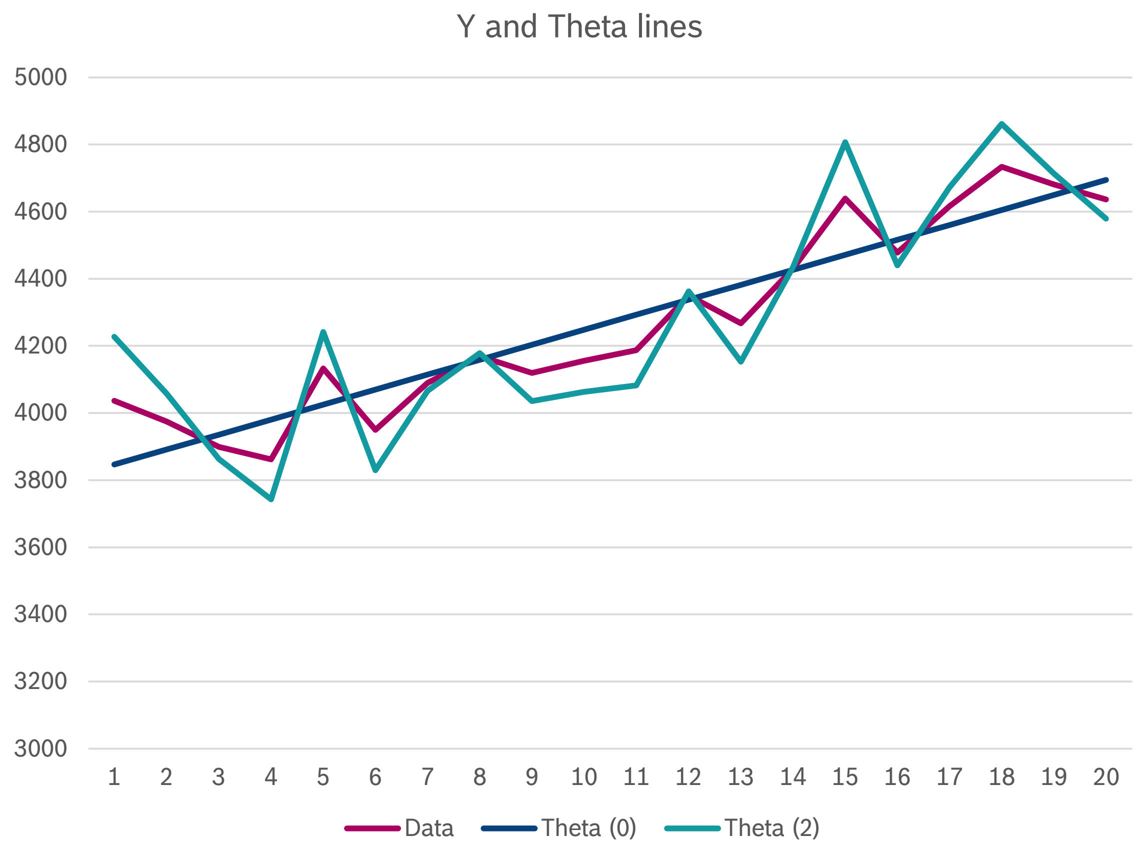
						min	max		average w/o min max	Seas.Ind
A		105,58	101,15	99,54	101,7	99,54	105,58		101,43	101,79
B		95,12	96,9	97,36	99,32	95,12	99,32		97,13	97,48
C	97,14	98,65	97,71	101,65		97,14	101,65		98,18	98,54
D	99,54	103,69	103,98	99,95		99,54	103,98		101,82	102,19
								sum	398,555	400
								Cor	99,639%	

Internal | CR/AEU2 - Patelis | 2018-03-14
© Robert Bosch GmbH 2018. All rights reserved, also regarding any disposal, exploitation, reproduction, editing, distribution, as well as in the event of applications for industrial property rights.
BOSCH

Forecasting Methods
Theta method: Example (2)

▶ Step 2: ***Decomposition & Theta lines***

		Nominator			Denominator		
		A	B			SLR	
Period	Data	X-meax(X)	Y=mean(Y)	A x B	(X-Mean(X))^2	Theta (0)	Theta (2)
1	4036,68	-9,5	-233,65	2219,66	90,25	3846,41	4226,95
2	3974,13	-8,5	-296,20	2517,69	72,25	3891,03	4057,23
3	3899,03	-7,5	-371,30	2784,74	56,25	3935,65	3862,41
4	3861,43	-6,5	-408,90	2657,84	42,25	3980,28	3742,58
5	4132,95	-5,5	-137,38	755,58	30,25	4024,90	4241,00
6	3949,51	-4,5	-320,82	1443,68	20,25	4069,52	3829,50
7	4089,82	-3,5	-180,51	631,78	12,25	4114,15	4065,49
8	4168,71	-2,5	-101,62	254,05	6,25	4158,77	4178,65
9	4119,2	-1,5	-151,13	226,69	2,25	4203,39	4035,01
10	4155,7	-0,5	-114,63	57,31	0,25	4248,02	4063,38
11	4187,24	0,5	-83,09	-41,54	0,25	4292,64	4081,84
12	4349,74	1,5	79,41	119,12	2,25	4337,26	4362,22
13	4267,54	2,5	-2,79	-6,97	6,25	4381,89	4153,19
14	4430,63	3,5	160,30	561,06	12,25	4426,51	4434,75
15	4638,85	4,5	368,52	1658,35	20,25	4471,13	4806,57
16	4477,93	5,5	207,60	1141,81	30,25	4515,76	4440,10
17	4616,3	6,5	345,97	2248,81	42,25	4560,38	4672,22
18	4733,25	7,5	462,92	3471,91	56,25	4605,00	4861,50
19	4681,47	8,5	411,14	3494,70	72,25	4649,63	4713,31
20	4636,46	9,5	366,13	3478,25	90,25	4694,25	4578,67
				sum	sum		
				29674,52	665,00		
average			SLR	Slope	44,62		
10,5	4270,329			constant	3801,78		

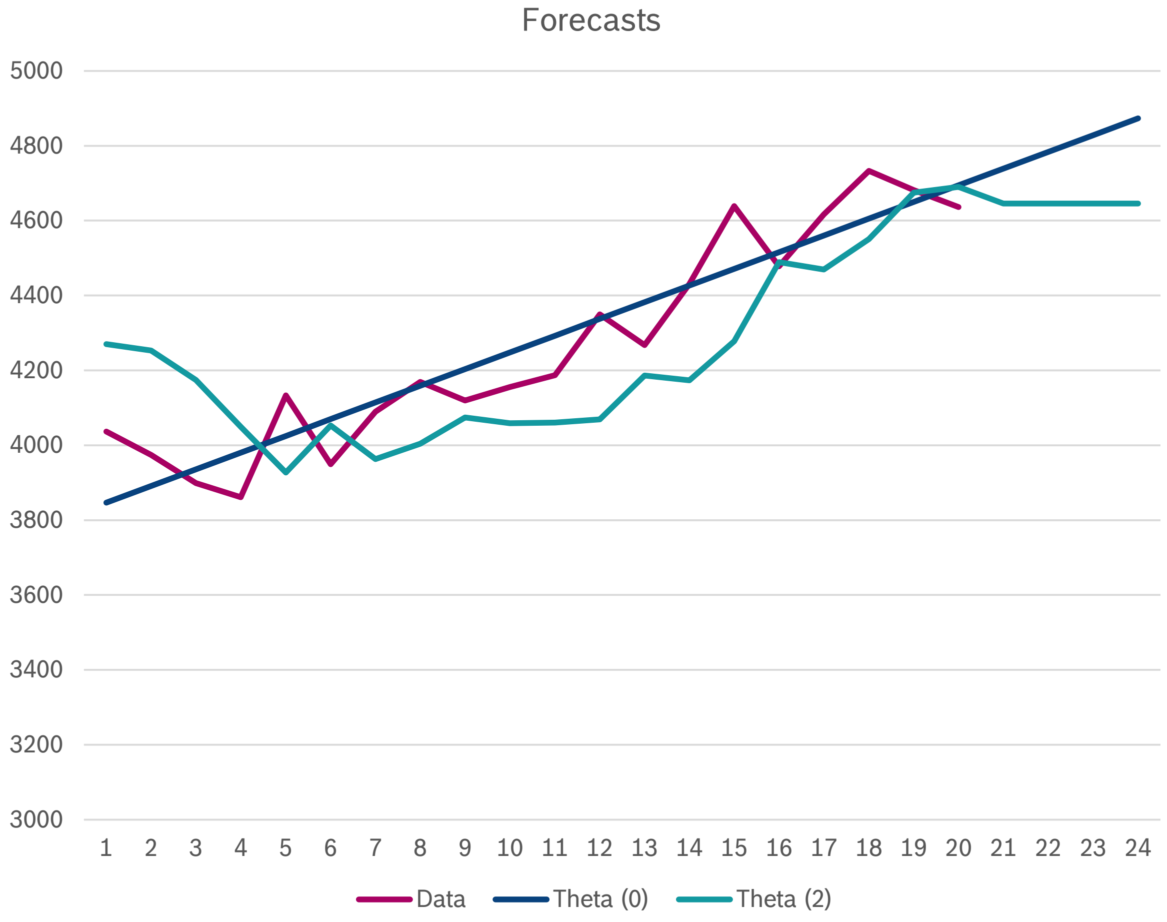


© Robert Bosch GmbH 2018. All rights reserved, also regarding any disposal, exploitation, reproduction, editing, distribution, as well as in the event of applications for industrial property rights.

Forecasting Methods
Theta method: Example (3)

▶ Step 3: ***Forecast***

		SLR	S(0) =	4270,33		α=0.4	
Period	Data	Theta (0)	Period	Theta (2)	Forecasts	Error e	Level S
1	4036,68	3846,41	1	4226,95	4270,3	-43,4	4253,0
2	3974,13	3891,03	2	4057,23	4253,0	-195,7	4174,7
3	3899,03	3935,65	3	3862,41	4174,7	-312,3	4049,8
4	3861,43	3980,28	4	3742,58	4049,8	-307,2	3926,9
5	4132,95	4024,90	5	4241,00	3926,9	314,1	4052,5
6	3949,51	4069,52	6	3829,50	4052,5	-223,0	3963,3
7	4089,82	4114,15	7	4065,49	3963,3	102,2	4004,2
8	4168,71	4158,77	8	4178,65	4004,2	174,5	4074,0
9	4119,2	4203,39	9	4035,01	4074,0	-39,0	4058,4
10	4155,7	4248,02	10	4063,38	4058,4	5,0	4060,4
11	4187,24	4292,64	11	4081,84	4060,4	21,5	4069,0
12	4349,74	4337,26	12	4362,22	4069,0	293,2	4186,3
13	4267,54	4381,89	13	4153,19	4186,3	-33,1	4173,0
14	4430,63	4426,51	14	4434,75	4173,0	261,7	4277,7
15	4638,85	4471,13	15	4806,57	4277,7	528,8	4489,3
16	4477,93	4515,76	16	4440,10	4489,3	-49,2	4469,6
17	4616,3	4560,38	17	4672,22	4469,6	202,6	4550,6
18	4733,25	4605,00	18	4861,50	4550,6	310,9	4675,0
19	4681,47	4649,63	19	4713,31	4675,0	38,3	4690,3
20	4636,46	4694,25	20	4578,67	4690,3	-111,6	4645,7
21		4738,87	21		4645,7		
22		4783,50	22		4645,7		
23		4828,12	23		4645,7		
24		4872,74	24		4645,7		

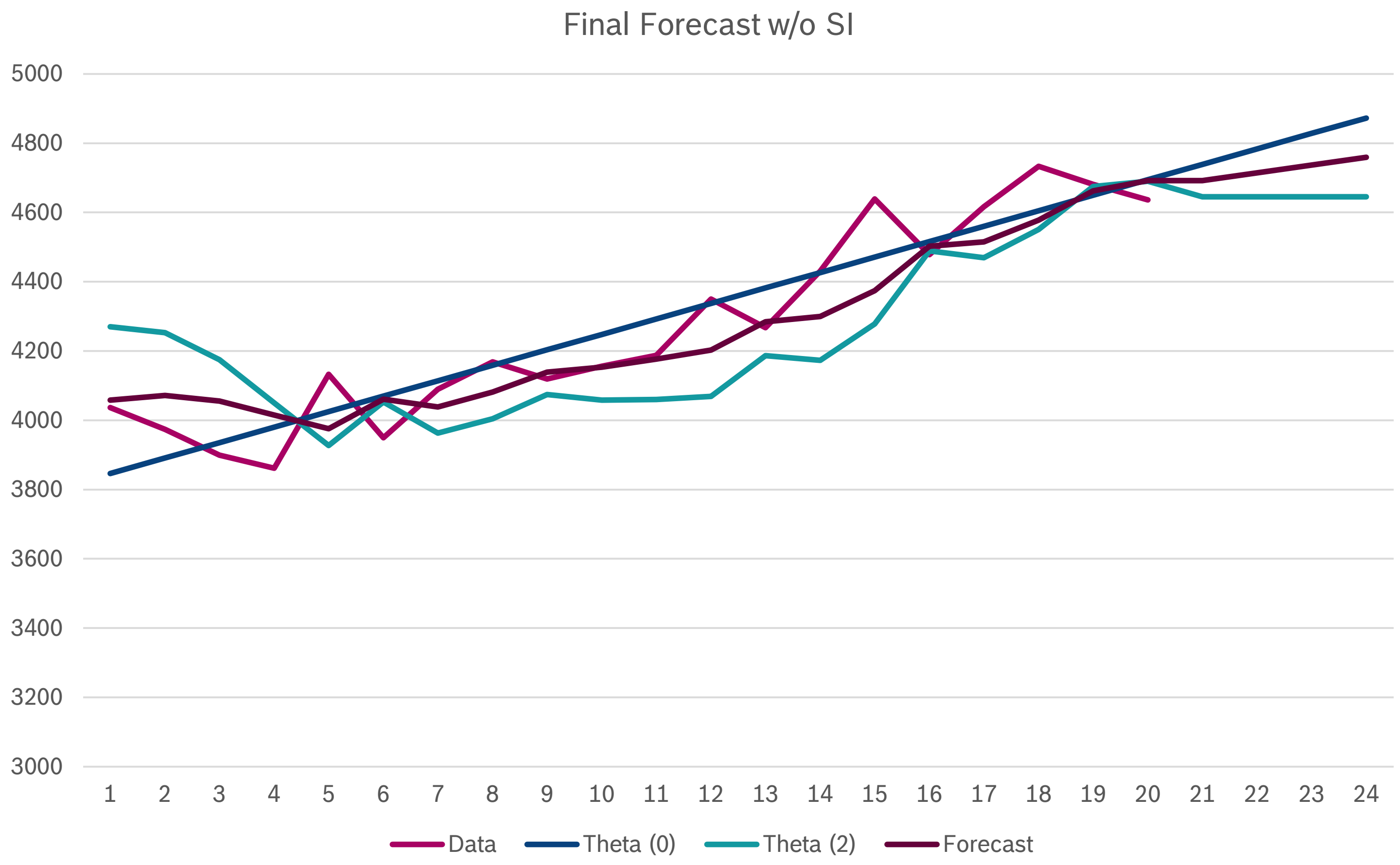


© Robert Bosch GmbH 2018. All rights reserved, also regarding any disposal, exploitation, reproduction, editing, distribution, as well as in the event of applications for industrial property rights.
BOSCH

Forecasting Methods
Theta method: Example (4)

▶ Step 4: ***Synthesis***

Period	Data	Theta (0)	Theta (2)	Forecast
1	4036,68	3846,41	4270,33	4058,37
2	3974,13	3891,03	4252,98	4072,00
3	3899,03	3935,65	4174,68	4055,17
4	3861,43	3980,28	4049,77	4015,02
5	4132,95	4024,90	3926,90	3975,90
6	3949,51	4069,52	4052,54	4061,03
7	4089,82	4114,15	3963,32	4038,73
8	4168,71	4158,77	4004,19	4081,48
9	4119,2	4203,39	4073,97	4138,68
10	4155,7	4248,02	4058,39	4153,20
11	4187,24	4292,64	4060,39	4176,51
12	4349,74	4337,26	4068,97	4203,12
13	4267,54	4381,89	4186,27	4284,08
14	4430,63	4426,51	4173,04	4299,77
15	4638,85	4471,13	4277,72	4374,43
16	4477,93	4515,76	4489,26	4502,51
17	4616,3	4560,38	4469,60	4514,99
18	4733,25	4605,00	4550,65	4577,82
19	4681,47	4649,63	4674,99	4662,31
20	4636,46	4694,25	4690,32	4692,28
21		4738,87	4645,66	4692,27
22		4783,50	4645,66	4714,58
23		4828,12	4645,66	4736,89
24		4872,74	4645,66	4759,20

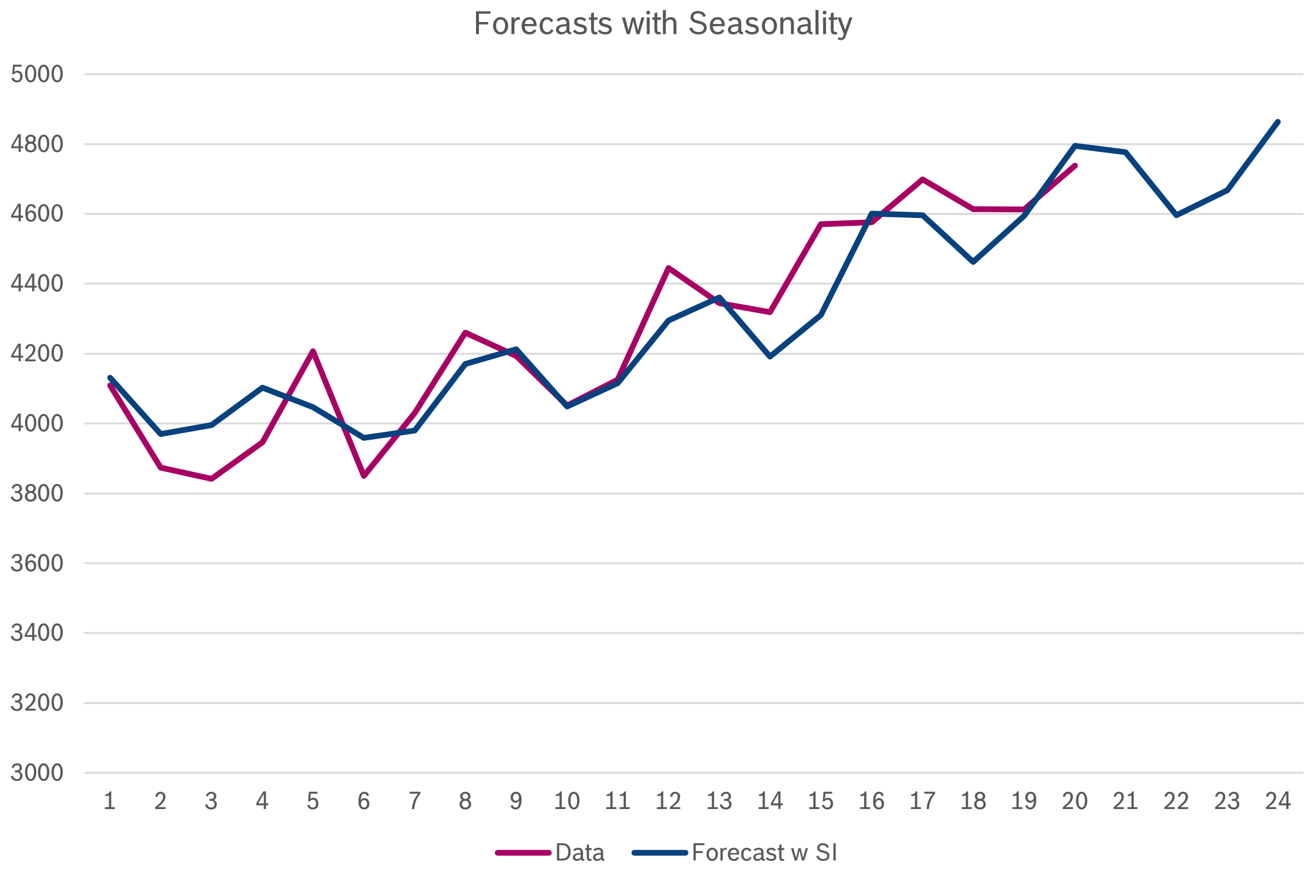


© Robert Bosch GmbH 2018. All rights reserved, also regarding any disposal, exploitation, reproduction, editing, distribution, as well as in the event of applications for industrial property rights.
BOSCH

Forecasting Methods
Theta method: Example (5)

▶ Step 5: ***Seasonalization***

Period	Data	Forecast w/o SI	Seas.Ind	Forecast w SI
1	4109	4058,37	101,79	4131,12
2	3874	4072,00	97,48	3969,48
3	3842	4055,17	98,54	3995,80
4	3946	4015,02	102,19	4102,92
5	4207	3975,90	101,79	4047,17
6	3850	4061,03	97,48	3958,78
7	4030	4038,73	98,54	3979,61
8	4260	4081,48	102,19	4170,83
9	4193	4138,68	101,79	4212,88
10	4051	4153,20	97,48	4048,63
11	4126	4176,51	98,54	4115,37
12	4445	4203,12	102,19	4295,13
13	4344	4284,08	101,79	4360,88
14	4319	4299,77	97,48	4191,51
15	4571	4374,43	98,54	4310,38
16	4576	4502,51	102,19	4601,08
17	4699	4514,99	101,79	4595,93
18	4614	4577,82	97,48	4462,56
19	4613	4662,31	98,54	4594,05
20	4738	4692,28	102,19	4795,01
21		4692,27	101,79	4776,39
22		4714,58	97,48	4595,87
23		4736,89	98,54	4667,54
24		4759,20	102,19	4863,39



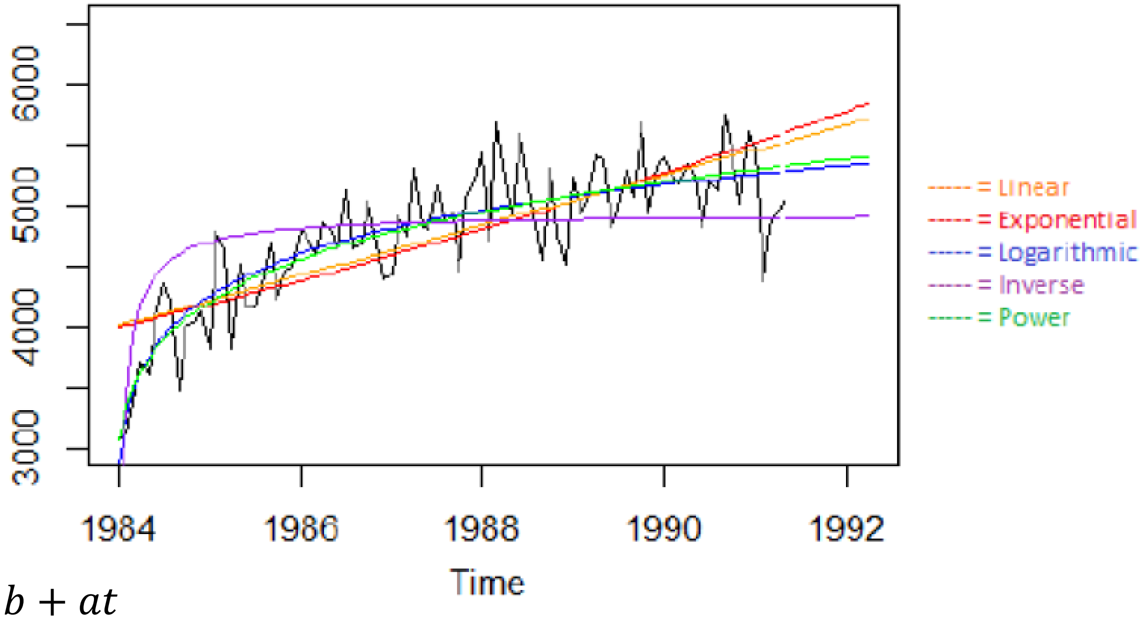
© Robert Bosch GmbH 2018. All rights reserved, also regarding any disposal, exploitation, reproduction, editing, distribution, as well as in the event of applications for industrial property rights.

ALTERNATIVE THETA METHOD

Forecasting Methods
Theta method: Considering nonlinear trends

- Generalize the method:

$$Y_t = \frac{1}{\theta} \times Y_t^{\theta} + (1-\theta) \times Y_t^{0}$$



- by using ***not only linear trends*** for Y^0:
  - Linear regression: $Y_t^0 = b + a \times t$
  - Exponential curve: $Y_t^0 = b \times e^{at}$, or $\log(Y_t^0) = \log b + at$
  - Logarithmic curve: $Y_t^0 = b + a \log t$
  - Inverse curve: $Y_t^0 = b + a\frac{1}{t}$
  - Power curve: $Y_t^0 = bt^a$, or $\log(Y_t^0) = \log b + a \log t$

Internal | CR/AEU2 - Patelis | 2018-03-14
© Robert Bosch GmbH 2018. All rights reserved, also regarding any disposal, exploitation, reproduction, editing, distribution, as well as in the event of applications for industrial property rights.

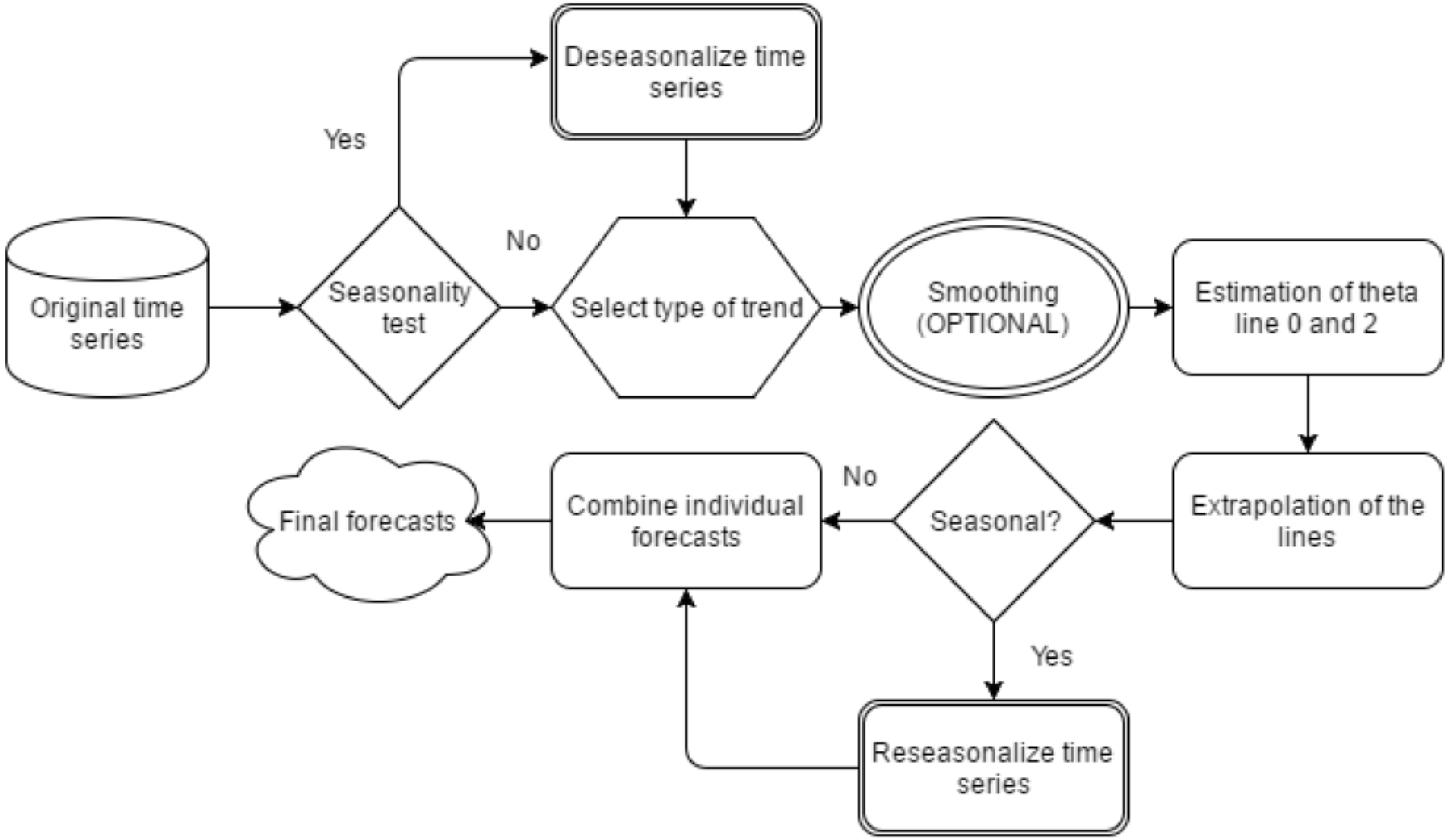
Forecasting Methods
Theta method: Smoothing the data

- ***Randomness and Outliers*** can affect the performance, in two ways:
  - ***Ineffective parameterization*** when calculating *a* and *b* for Y^0, leading to wrong trend pattern.
  - ***Poor performance*** of SES used for extrapolating Y^θ.

- Three ways for solving the problem:
  - ***Transformation***: (+) Limit the variations, (-) affect the trend pattern
  - ***Outlier detection***: (+) removes outliers, (-) increases uncertainty, (-) increases complexity
  - ***Smoothing***: (+) Limit the variations, (+) simple to implement
    - Example: Using a non-linear smoothing process (*Assimakopoulos V., 'A successive filtering technique for identifying long-term trends, IJF 14, 1995*)

Internal | CR/AEU2 - Patelis | 2018-03-14
© Robert Bosch GmbH 2018. All rights reserved, also regarding any disposal, exploitation, reproduction, editing, distribution, as well as in the event of applications for industrial property rights.
BOSCH

Forecasting Methods
Theta method: Selecting the best



Source: *“Transforming the Theta model into a flexible decomposition method by considering nonlinear trends, FSU, 36th International Symposium on Forecasting, Spain, 2016*

© Robert Bosch GmbH 2018. All rights reserved, also regarding any disposal, exploitation, reproduction, editing, distribution, as well as in the event of applications for industrial property rights.

AVERAGING

“Can anything beat the simple average?”

Michael Clements, Chief editor International Journal of Forecasting

BOSCH

Forecasting Methods
Averaging method (1)

- Averaging method: ***Combination of two or more*** simple statistical methods, in equal or unequal weights.
- The choice of methods to be involved (and their weights) are determined by:
  - the specific characteristics of each method.
  - the characteristics of the timeseries.
  - the ***forecasting horizon*** (plays a significant factor).

© Robert Bosch GmbH 2018. All rights reserved, also regarding any disposal, exploitation, reproduction, editing, distribution, as well as in the event of applications for industrial property rights.
BOSCH

Forecasting Methods
Averaging method (2)

- If method X has smaller forecasting error E_x that the errors E_y of the method Y (in terms of MAPE or MSE), the combination of methods X and Y will ***not have as error the average*** of E_x and E_y.
- In general the combination of methods:
  - can lead to ***better forecasts*** and smaller forecasting errors.
  - allows us to be ***more accurate*** in our predictions.
  - ***reduces our uncertainty*** for the future, when we are not sure if the data will keep the same characteristics in the future.

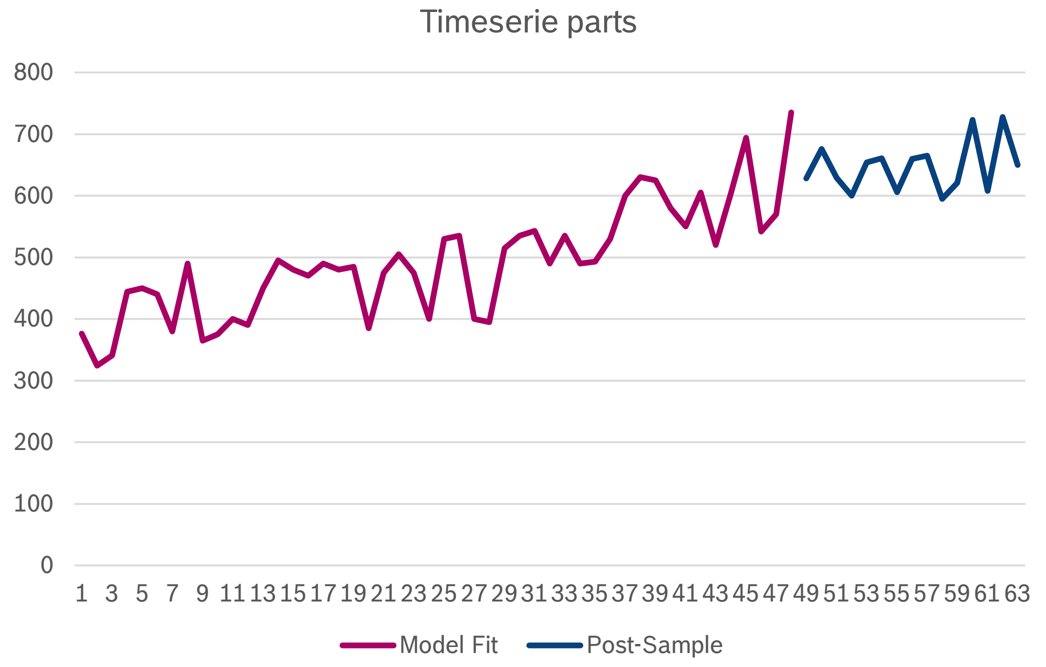
➢ The benefits accrue because they allow the forecast to draw on a ***wider range of information***.

  ➢ If X tends to be to high and Y to low, the combined forecast will tend to ***cancel out these biases***.

© Robert Bosch GmbH 2018. All rights reserved, also regarding any disposal, exploitation, reproduction, editing, distribution, as well as in the event of applications for industrial property rights.

Forecasting Methods
Averaging method: Example (1)

- We will estimate forecasts for a timeserie with:
  - 48 data points in model-fit, 15 data points in post-sample
  - With ***3 smoothing methods***:
    - Simple (SES)
    - Holt
    - Damped
  - With ***2 different set of weights***:
    - Simple weights, thus same weights for all 3 methods.
    - Advanced weights, thus the weights will be calculated with an inverse function of MAPE.

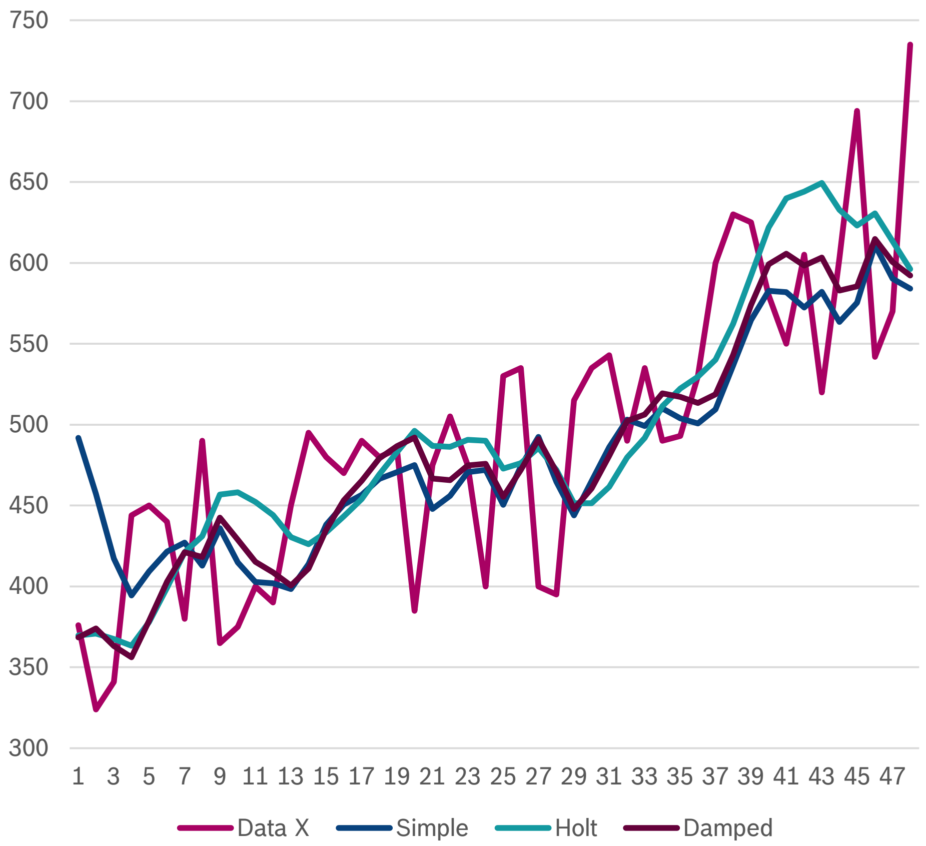


© Robert Bosch GmbH 2018. All rights reserved, also regarding any disposal, exploitation, reproduction, editing, distribution, as well as in the event of applications for industrial property rights.

Forecasting Methods
Averaging method: Example (2)

Time (t)	Data X	Simple	Holt	Damped
1	376	492	369,6	368,6
2	324	457	370,9	374,0
3	341	417	367,4	363,1
4	444	394	363,3	356,2
5	450	409	378,0	378,8
6	440	421	399,1	402,9
7	380	427	421,2	421,1
8	490	413	431,0	418,3
9	365	436	456,7	442,6
10	375	415	458,2	428,9
11	400	403	452,2	415,2
12	390	402	444,1	408,7
13	450	398	430,4	400,6
14	495	414	426,0	411,0
15	480	438	433,5	434,9
16	470	451	443,4	453,2
17	490	457	454,0	465,3
18	480	467	469,2	479,2
19	485	471	483,0	486,7
20	385	475	496,1	492,0
21	475	448	486,8	466,6
22	505	456	486,2	465,7
23	475	471	490,6	474,7
24	400	472	490,0	475,8
25	530	450	472,9	455,3
26	535	474	476,2	472,1
27	400	493	485,6	491,1
28	395	465	472,0	470,8
29	515	444	451,5	447,9
30	535	465	451,4	460,5
31	543	486	461,6	480,7
32	490	503	479,8	502,3
33	535	499	491,8	506,2
34	490	510	511,4	519,4
35	493	504	522,4	517,2
36	530	501	529,7	513,4
37	600	509	540,1	518,8
38	630	537	562,4	543,2
39	625	565	592,2	573,9
40	580	583	621,9	599,1
41	550	582	639,9	605,7
42	605	572	644,1	598,4
43	520	582	649,5	603,4
44	603	564	632,9	582,9
45	694	575	623,2	585,6
46	542	611	630,7	614,7
47	570	590	613,4	600,4
48	735	584	596,2	592,2

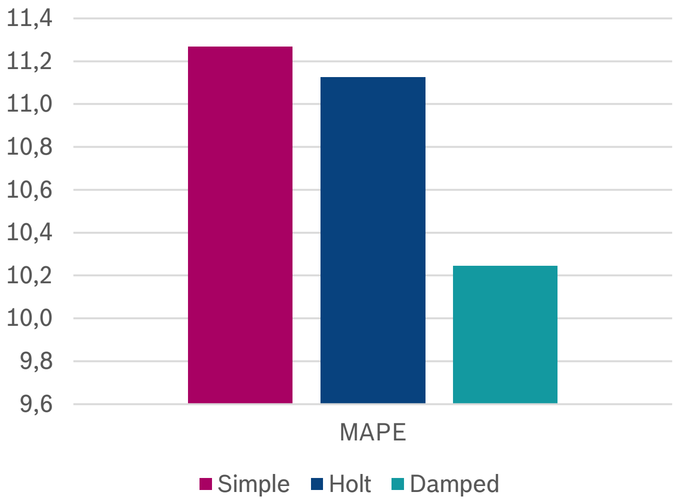
Fit-model Forecasts



▶ Best method: ***Damped***

	Simple	Holt	Damped
MAPE	11,268	11,126	10,245

Fit-model MAPE

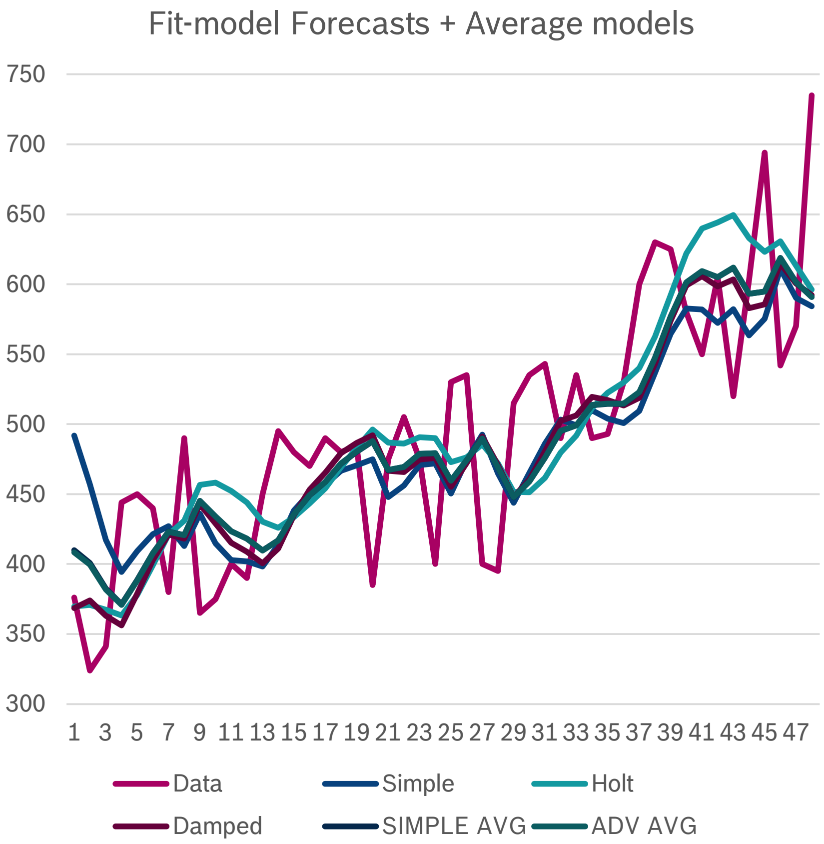


© Robert Bosch GmbH 2018. All rights reserved, also regarding any disposal, exploitation, reproduction, editing, distribution, as well as in the event of applications for industrial property rights.

Forecasting Methods

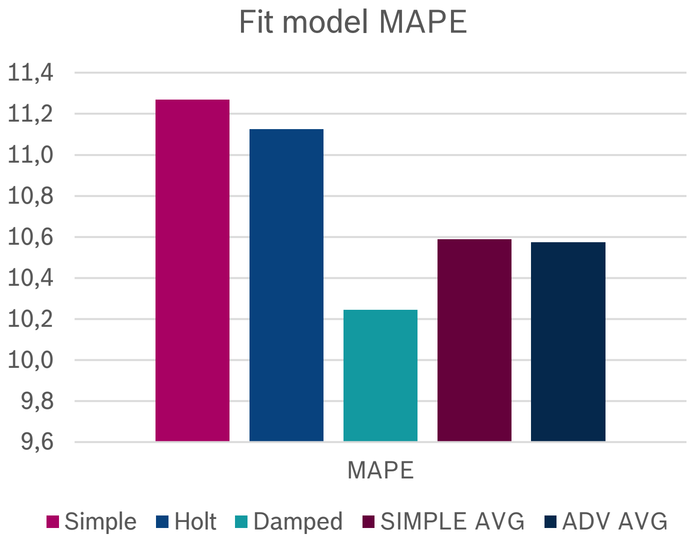
Averaging method: Example (3)

Time (t)	Data X	Simple	Holt	Damped	SIMPLE AVG	ADV AVG
1	376	491,875	369,6125	368,572	410	408,4
2	324	457,1125	370,89	374,02	401	399,6
3	341	417,1788	367,4	363,12	383	381,8
4	444	394,3251	363,3	356,22	371	370,8
5	450	409,2276	378	378,84	389	388,3
6	440	421,4593	399,1	402,86	408	407,6
7	380	427,0215	421,2	421,1	423	423,0
8	490	412,9151	431	418,26	421	420,8
9	365	436,0405	456,7	442,6	445	445,2
10	375	414,7284	458,2	428,86	434	434,0
11	400	402,8099	452,2	415,22	423	423,5
12	390	401,9669	444,1	408,68	418	418,2
13	450	398,3768	430,4	400,58	410	409,7
14	495	413,8638	426	410,98	417	416,9
15	480	438,2046	433,5	434,92	436	435,5
16	470	450,7433	443,4	453,18	449	449,2
17	490	456,5203	454	465,3	459	458,8
18	480	466,5642	469,2	479,24	472	471,9
19	485	470,5949	483	486,68	480	480,3
20	385	474,9165	496,1	491,98	488	487,9
21	475	447,9415	486,8	466,6	467	467,3
22	505	456,0591	486,2	465,74	469	469,4
23	475	470,7413	490,6	474,72	479	478,7
24	400	472,0189	490	475,76	479	479,3
25	530	450,4133	472,9	455,32	460	459,6
26	535	474,2893	476,2	472,06	474	474,1
27	400	492,5025	485,6	491,08	490	489,7
28	395	464,7517	472	470,82	469	469,3
29	515	443,8262	451,5	447,94	448	447,8
30	535	465,1784	451,4	460,52	459	459,0
31	543	486,1248	461,6	480,68	476	476,1
32	490	503,1874	479,8	502,3	495	495,2
33	535	499,2312	491,8	506,2	499	499,2
34	490	509,9618	511,4	519,44	514	513,8
35	493	503,9733	522,4	517,2	515	514,7
36	530	500,6813	529,7	513,36	515	514,7
37	600	509,4769	540,1	518,78	523	522,8
38	630	536,6338	562,4	543,16	547	547,4
39	625	564,6437	592,2	573,94	577	577,0
40	580	582,7506	621,9	599,08	601	601,4
41	550	581,9254	639,9	605,7	609	609,4
42	605	572,3478	644,1	598,44	605	605,2
43	520	582,1434	649,5	603,4	612	611,8
44	603	563,5004	632,9	582,94	593	593,2
45	694	575,3503	623,2	585,56	595	594,7
46	542	610,9452	630,7	614,72	619	618,8
47	570	590,2616	613,4	600,44	601	601,5
48	735	584,1831	596,2	592,16	591	590,9



- Best method: ***Damped***
  - 2nd best: ***Advanced Average***

	Simple	Holt	Damped	SIMPLE AVG	ADV AVG
MAPE	11,268	11,126	10,245	10,589	10,575

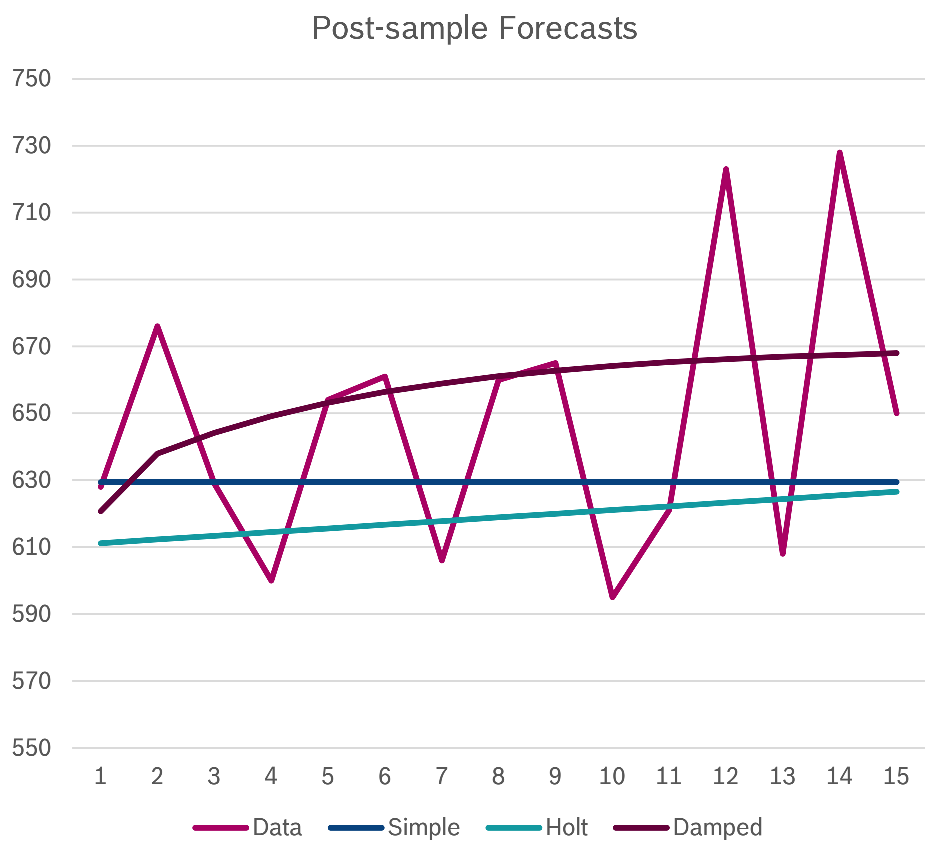


Internal | CR/AEU2 - Patelis | 2018-03-14
© Robert Bosch GmbH 2018. All rights reserved, also regarding any disposal, exploitation, reproduction, editing, distribution, as well as in the event of applications for industrial property rights.
BOSCH

Forecasting Methods

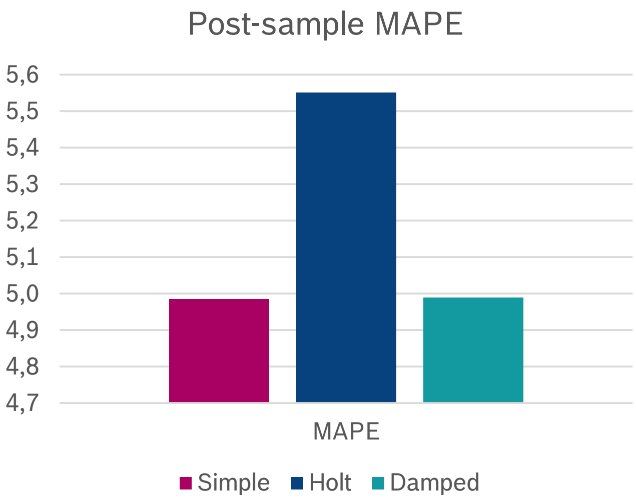
Averaging method: Example (4)

Time (t)	Data	Simple	Holt	Damped
49	628	629	611,2	620,7
50	676	629	612,3	638,0
51	629	629	613,4	644,2
52	600	629	614,5	649,2
53	654	629	615,6	653,2
54	661	629	616,7	656,4
55	606	629	617,8	659,0
56	660	629	618,9	661,1
57	665	629	620,0	662,8
58	595	629	621,1	664,2
59	621	629	622,2	665,3
60	723	629	623,3	666,2
61	608	629	624,4	666,9
62	728	629	625,5	667,5
63	650	629	626,6	668,0



▶ Best method: ***Simple***

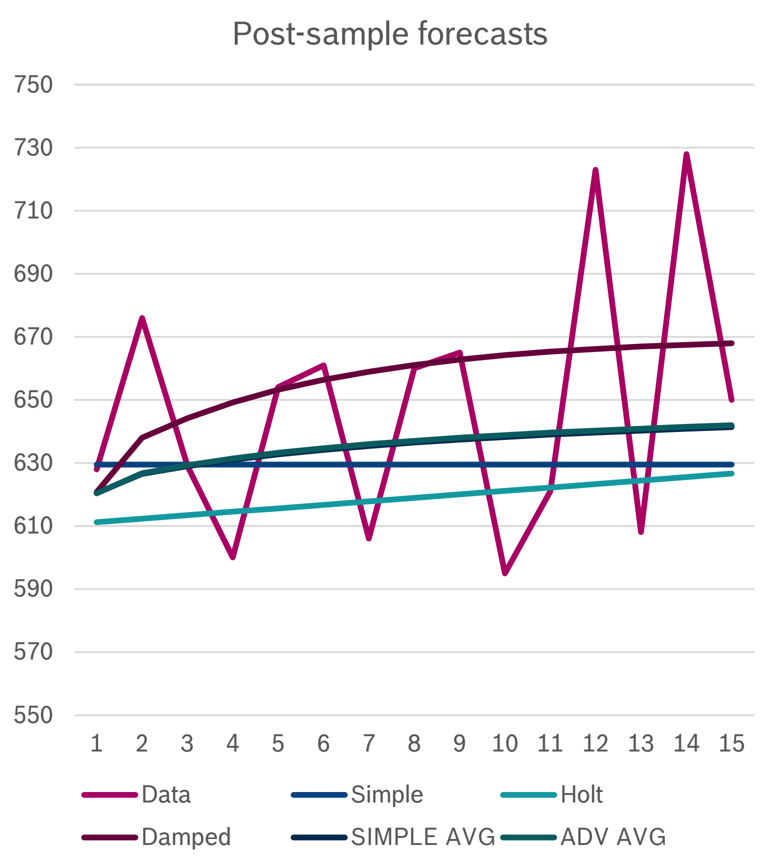
	Simple	Holt	Damped
MAPE	4,986	5,550	4,989



Internal | CR/AEU2 - Patelis | 2018-03-14
© Robert Bosch GmbH 2018. All rights reserved, also regarding any disposal, exploitation, reproduction, editing, distribution, as well as in the event of applications for industrial property rights.

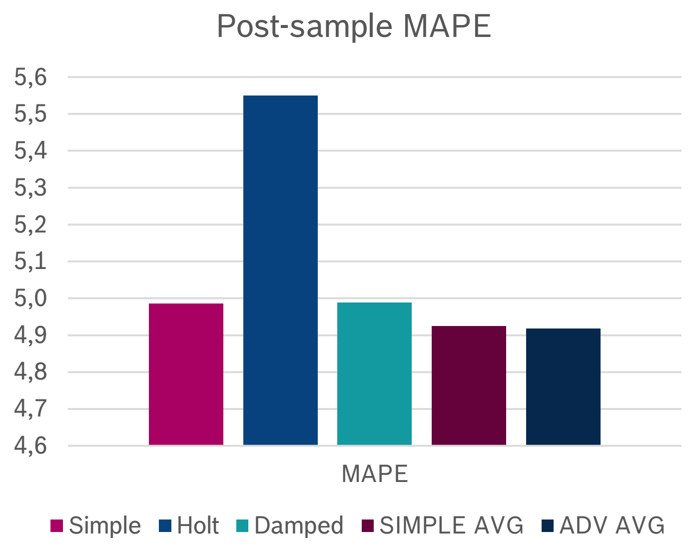
Forecasting Methods
Averaging method: Example (5)

Time (t)	Data	Simple	Holt	Damped	SIMPLE AVG	ADV AVG
49	628	629,4	611,2	620,7	620,4	620,4
50	676	629,4	612,3	638,0	626,6	626,8
51	629	629,4	613,4	644,2	629,0	629,3
52	600	629,4	614,5	649,2	631,0	631,4
53	654	629,4	615,6	653,2	632,7	633,2
54	661	629,4	616,7	656,4	634,2	634,7
55	606	629,4	617,8	659,0	635,4	635,9
56	660	629,4	618,9	661,1	636,5	637,0
57	665	629,4	620,0	662,8	637,4	638,0
58	595	629,4	621,1	664,2	638,2	638,8
59	621	629,4	622,2	665,3	639,0	639,6
60	723	629,4	623,3	666,2	639,6	640,3
61	608	629,4	624,4	666,9	640,2	640,9
62	728	629,4	625,5	667,5	640,8	641,4
63	650	629,4	626,6	668,0	641,3	642,0



- Best method: ***Adv. Average***
  - 2^{nd} best: ***Simple***

	Simple	Holt	Damped	SIMPLE AVG	ADV AVG
MAPE	4,986	5,550	4,989	4,925	4,918



Internal | CR/AEU2 - Patelis | 2018-03-14
© Robert Bosch GmbH 2018. All rights reserved, also regarding any disposal, exploitation, reproduction, editing, distribution, as well as in the event of applications for industrial property rights.

ARIMA

Forecasting Methods
ARIMA

- ARIMA: ***A**uto **R**egressive **I**ntegrated **M**oving **A**verage*
  - A stochastic model that tries to describe the ***evolution of a measure*** over time.
  - Introduced by ***Box & Jenkins*** (1971).
  - They can be applied in cases where data shows ***evidence of non-stationary***, where an initial differencing step (corresponding to the "integrated" part of the model) can be applied to remove the non-stationary.
- The model is generally referred to as an ***ARIMA (p, d, q)*** model where:
  - **p**, **d**, and **q** are non-negative integers that refer to the order of the ***autoregressive, integrated, and moving average parts*** of the model respectively.
    - If d = 0, then the data could be used without a differencing step, thus in a ARMA(p, q) model.

Internal | CR/AEU2 - Patelis | 2018-03-14
© Robert Bosch GmbH 2018. All rights reserved, also regarding any disposal, exploitation, reproduction, editing, distribution, as well as in the event of applications for industrial property rights.

Forecasting Methods
ARIMA: Equations

- ***AR(p): Autoregressive model***
  - A data point value depends on the previous data points ***values***.

$Y_t = c + \varphi_1 \times \Upsilon_{t-1} + \varphi_2 \times \Upsilon_{t-2} + \cdots + \varphi_p \times \Upsilon_{t-p} + e_t,$ where $\mathrm{c} = \bar{Y} \times (1 - \varphi_1 - \varphi_2 - \cdots - \varphi_p)$

- ***MA(q): Moving Average model***
  - A data point value depends on the previous data points ***errors***.

$Y_t = c - \theta_1 \times \mathrm{e}_{t-1} - \theta_2 \times \mathrm{e}_{t-2} - \cdots - -\theta_q \times \mathrm{e}_{t-q} + e_t$ where $\mathrm{c} = \bar{Y}$

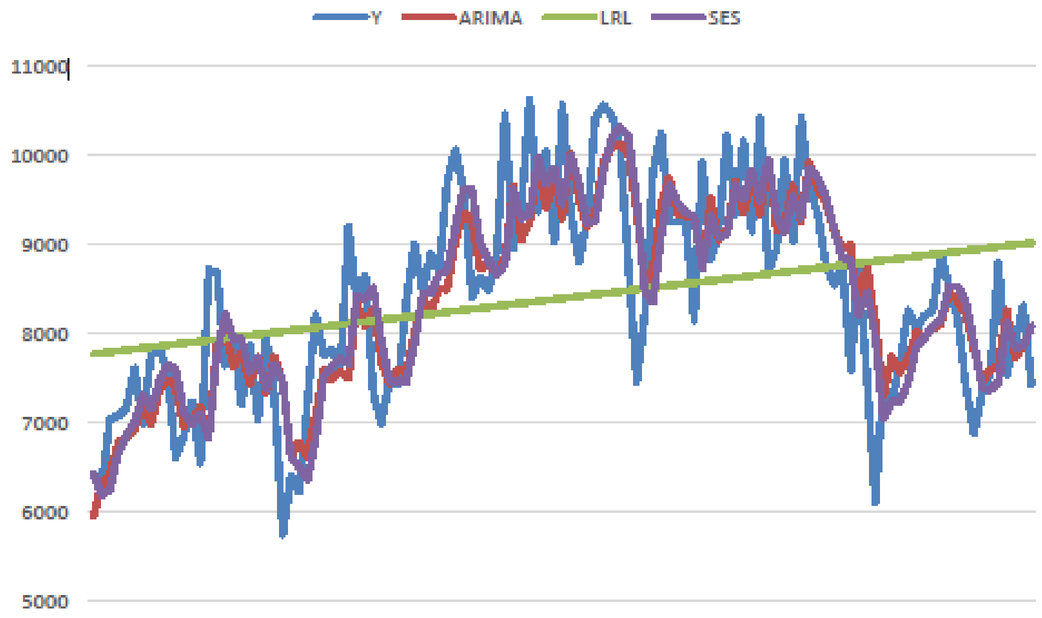
- ***ARMA(p, q)***

$$Y_t = c + \varphi_1 \times \Upsilon_{t-1} + \varphi_2 \times \Upsilon_{t-2} + \cdots + \varphi_p \times \Upsilon_{t-p} - \theta_1 \times \mathrm{e}_{t-1} - \theta_2 \times \mathrm{e}_{t-2} - \cdots - -\theta_q \times \mathrm{e}_{t-q} + e_t$$

© Robert Bosch GmbH 2018. All rights reserved, also regarding any disposal, exploitation, reproduction, editing, distribution, as well as in the event of applications for industrial property rights.

Forecasting Methods
ARIMA: Model

- They have a logic ***similar to regression (SLR) and smoothing (SES)*** methods, but:
- they express each timeserie value as a linear sum of previous data points values.
  - ***SLR***: $\widehat{Y_{t+1}} = 7768 + 11 \times (t+1)$
  - ***SES***: $\widehat{Y_{t+1}} = \widehat{Y_t} + 0.48 \times e_t$
  - ***ARIMA***: $\widehat{Y_{t+1}} = 1.31 \times Y_{t-1} - 0.31 \times Y_{t-2} + 0.8 \times e_t$



Internal | CR/AEU2 - Patelis | 2018-03-14
© Robert Bosch GmbH 2018. All rights reserved, also regarding any disposal, exploitation, reproduction, editing, distribution, as well as in the event of applications for industrial property rights.
BOSCH

Forecasting Methods
ARIMA: Requirements & Limitations

- For using an ARIMA model:
  - The timeserie should be ***distinct***: data points in equal time frames.
  - The timeserie should be ***static***: mean value, variance, and autocorrelation must be static within the sample.
  - ***Short-term*** forecasts must be estimated.
  - ***Granularity***: Forecast depends on ***p*** historical observations which as known.
  - ***Stagnation***: ARMA models can ***not be applied to data with trend*** or to data with ***periodical fluctuations.***
  - ***Long-term*** forecasts are based in short-term forecast, therefore the forecast error increases significantly.

© Robert Bosch GmbH 2018. All rights reserved, also regarding any disposal, exploitation, reproduction, editing, distribution, as well as in the event of applications for industrial property rights.
BOSCH

Forecasting Methods
ARIMA: Model Selection

- How to select proper ARIMA(p, d, q) model: ***A four step approach***
  - ***Data Preparation***: Remove seasonality and trend
    - By using integration, transformation, etc.
  - ***Recognition***: Detect potential representative models
    - by using statistical analysis, data autocorrelation, etc.
  - ***Evaluation***: Estimation of the p, d, q, parameters
    - by using expected likelihood, least square methods, etc.
  - ***Diagnostics***: Check the statistical significance of the parameters
    - by using t-test, error autocorrelation, etc.

© Robert Bosch GmbH 2018. All rights reserved, also regarding any disposal, exploitation, reproduction, editing, distribution, as well as in the event of applications for industrial property rights.

Forecasting Methods
ARIMA: Data Preparation (1)

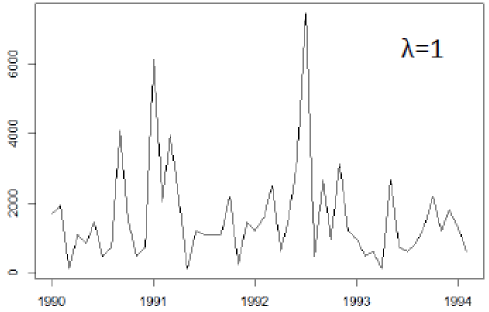
- ***Data transformations***:
  - Why? For limitation of variation and for removing abnormal values.
  - When? If timeserie has strong variance or the forecasting error is high.
- ***Data integration***:
  - Why? To remove (or to minimize) trend and seasonality.
  - When? If timeserie data have significant trend, or significant seasonality.

➢ ***We don’t need to deseasonalize*** the data (by using a decomposition method) before applying an ARIMA model. These models can handle seasonality though data transformations.

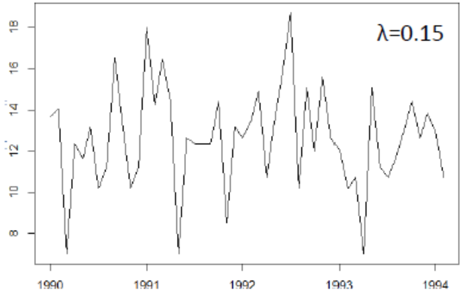
© Robert Bosch GmbH 2018. All rights reserved, also regarding any disposal, exploitation, reproduction, editing, distribution, as well as in the event of applications for industrial property rights.

Forecasting Methods
ARIMA: Data Preparation (2)

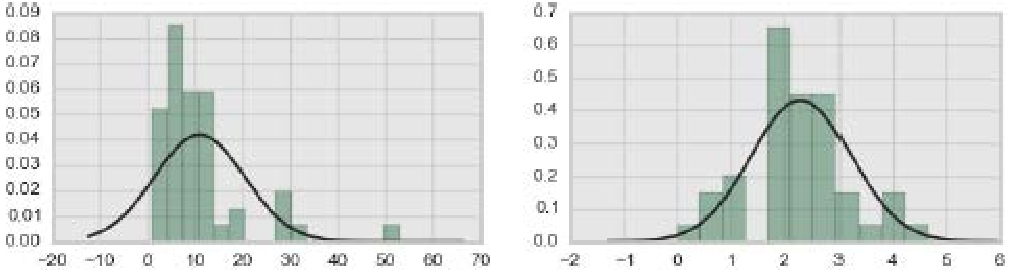
- The ARIMA models can not interpret periodical variations.
  - forecasting accuracy is limited when the variance of the data is high (high level of noisy data).
  - In case of white noise, forecasting accuracy is not affected.
- For removing the noise, we need to ***normalize the data*** by using logarithm or other transformations
  - Example: Box-Cox transformation $x_t = \frac{{y_t}^{\lambda}-1}{\lambda}$



Sd/mean=0.89



Sd/mean=0.2



Marginal distribution: before & after applying Box-Cox transformation

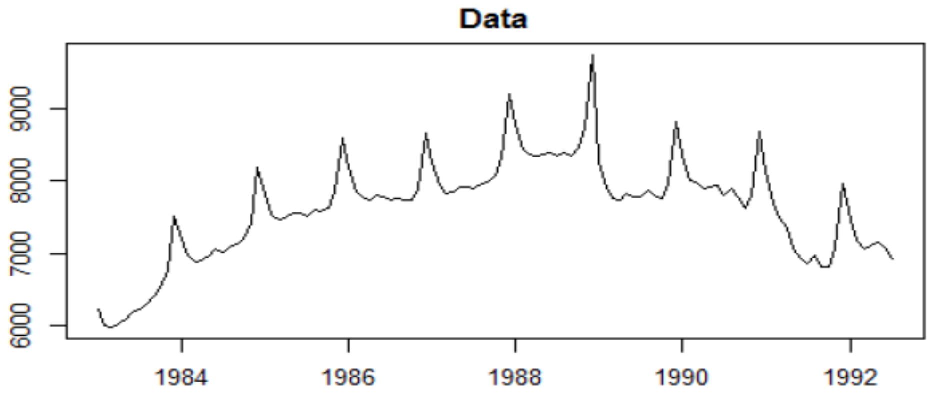
© Robert Bosch GmbH 2018. All rights reserved, also regarding any disposal, exploitation, reproduction, editing, distribution, as well as in the event of applications for industrial property rights.

Forecasting Methods
ARIMA: Data Preparation (3)

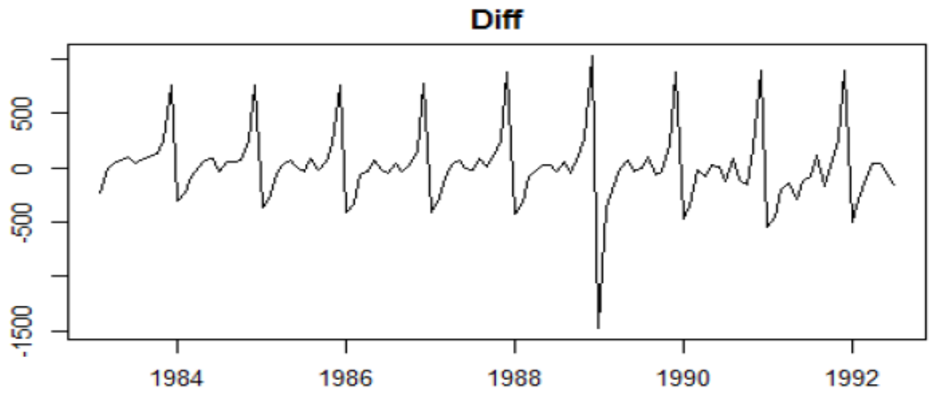
- The ***I(d)*** part of the ARIMA(p, d, q) model.
  - For removing a ***trend***, we create a new timeserie based on data point values differences
    - For a ***linear*** trend, we use ***1st level***, thus d=1: $Y_t{}' = Y_t - Y_{t-1}$
    - For a ***non-linear*** trend, we use ***2nd level***, thus d=2: $Y_t{}'' = Y_t{}' - Y_{t-1}{}' = Y_t - 2 \times Y_{t-1} - Y_{t-2}$
    - Parameter could be: $1 \leq d \leq n - 1$, where n is the number of timeseries' data points.
  - For removing strong ***seasonality***, we use seasonal differences:
    - For an m-seasonality: $Y_t{}' = Y_t - Y_{t-m}$

Internal | CR/AEU2 - Patelis | 2018-03-14
© Robert Bosch GmbH 2018. All rights reserved, also regarding any disposal, exploitation, reproduction, editing, distribution, as well as in the event of applications for industrial property rights.

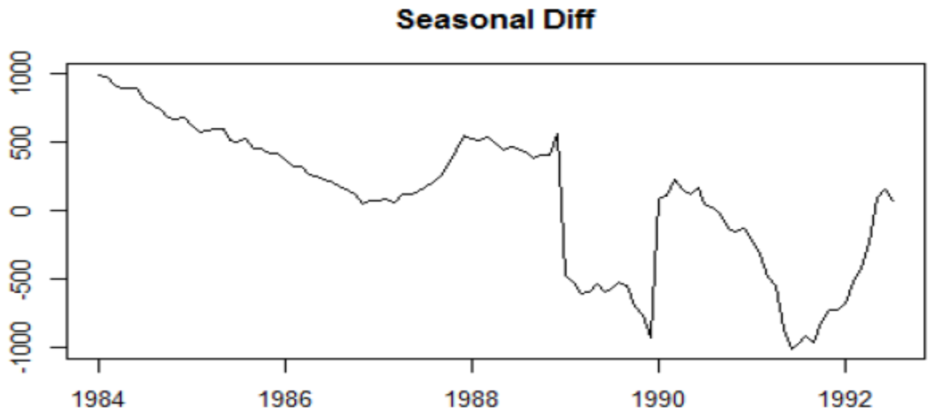
Forecasting Methods
ARIMA: Data Preparation Example



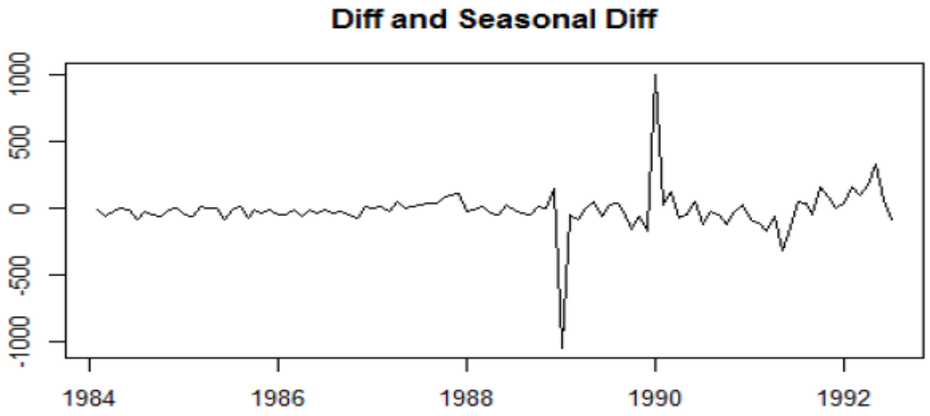
Data have significant trend and monthly seasonality



1st level integration can't remove seasonality → no stationary data



Seasonal integration can't remove trend → no stationary data



Solution: We must apply both ***1st level integration and seasonal integration***

 Internal | CR/AEU2 - Patelis | 2018-03-14
© Robert Bosch GmbH 2018. All rights reserved, also regarding any disposal, exploitation, reproduction, editing, distribution, as well as in the event of applications for industrial property rights.

BOSCH

Forecasting Methods
ARIMA: Recognition (1)

- ***ACF***: Indicates if a timeserie data point value depends from k past data points values

$$p_k = \frac{\sum_{t=k+1}^{T}(Y_t - \bar{Y}) \times (Y_{t-k} - \bar{Y})}{\sum_{t=1}^{T}(Y_t - \bar{Y})^2}$$

- ***PACF***: Indicates if a timeserie data point value depends from a past data point values, without taking into account the intermediate data points.

$$\varphi_{11} = p_1, \varphi_{kk} = p_k - \sum_{j=1}^{k-1} \frac{\varphi_{k-1,j} \times p_{k-j}}{1 - \sum_{j=1}^{k-1} \varphi_{k-1,j} \times p_j} \, when \; k = 2{,}3 \; and \; \varphi_{k,j} = \varphi_{k-1,j} - \varphi_{k,k} \times \varphi_{k-1,k-j} \; when \; k = 3{,}4 \; and \; j = 1{,}2, \dots$$

- Usually, we use them with ***k<3***. Greater values can lead to very complicated models without any gain in forecasting accuracy.
- These checks must be performed into stationary data, thus ***after applying integration***.

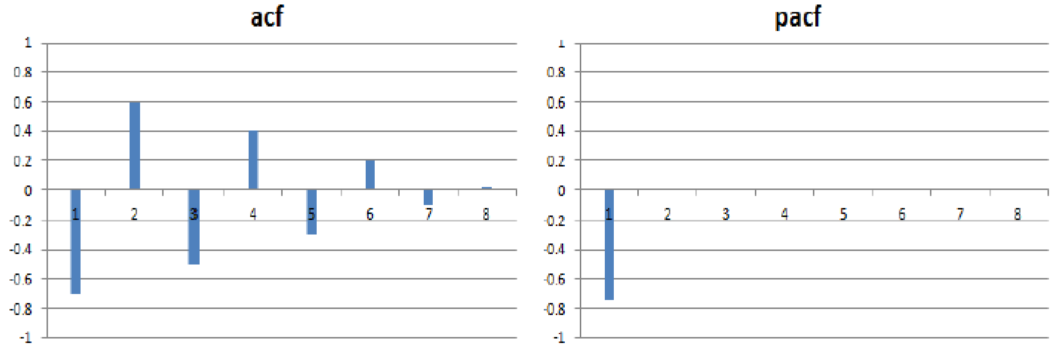
© Robert Bosch GmbH 2018. All rights reserved, also regarding any disposal, exploitation, reproduction, editing, distribution, as well as in the event of applications for industrial property rights.
BOSCH

Forecasting Methods
ARIMA: Recognition (2)

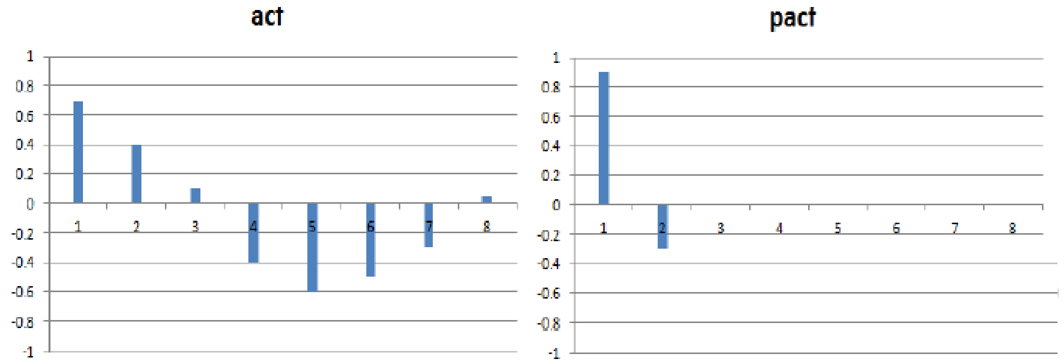
$$Y_t = c + \varphi_1 \times Y_{t-1} + \varphi_2 \times Y_{t-2} + \cdots + \varphi_p \times Y_{t-p} + e_t$$

- For a stationary model ***AR(p):***
  - The ACF values must ***decrease*** (exponentially or sin) ***towards 0***.
  - The PACF values must be 0 ***immediately*** after p lag-periods.

AR(1)



AR(2)



AR(1): $-1 < \varphi_1 < 1$

AR(2): $-1 < \varphi_2 < 1,\ \varphi_1 + \varphi_2 < 1, \varphi_2 - \varphi_1 < 1$

AR(p): $\varphi_1 + \varphi_2 + \cdots + \varphi_p < 1$

© Robert Bosch GmbH 2018. All rights reserved, also regarding any disposal, exploitation, reproduction, editing, distribution, as well as in the event of applications for industrial property rights.
BOSCH

Forecasting Methods
ARIMA: Recognition (3)

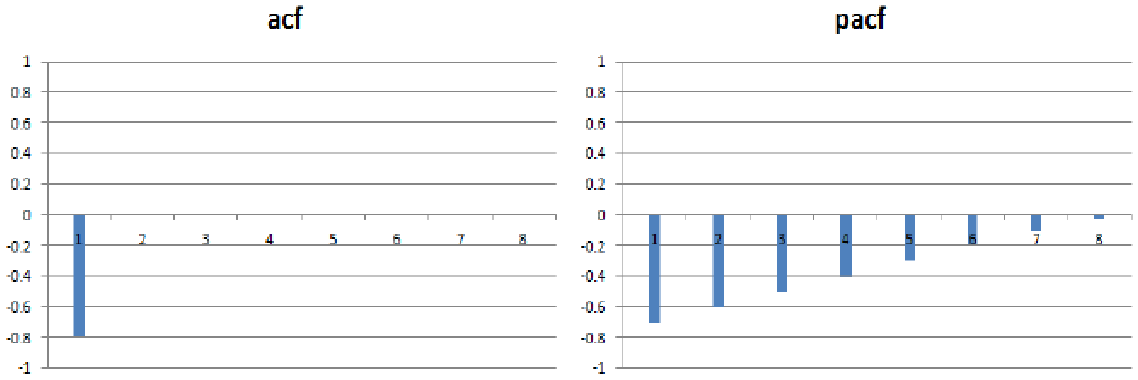
$$Y_t = c - \theta_1 \times e_{t-1} - \theta_2 \times e_{t-2} - \cdots - -\theta_q \times e_{t-q} + e_t$$

- For a stationary model ***MA(q):***
  - The ACF values must be 0 ***immediately*** after q lag-periods.
  - The PACF values must ***decrease*** (exponentially or sin) ***towards 0***.

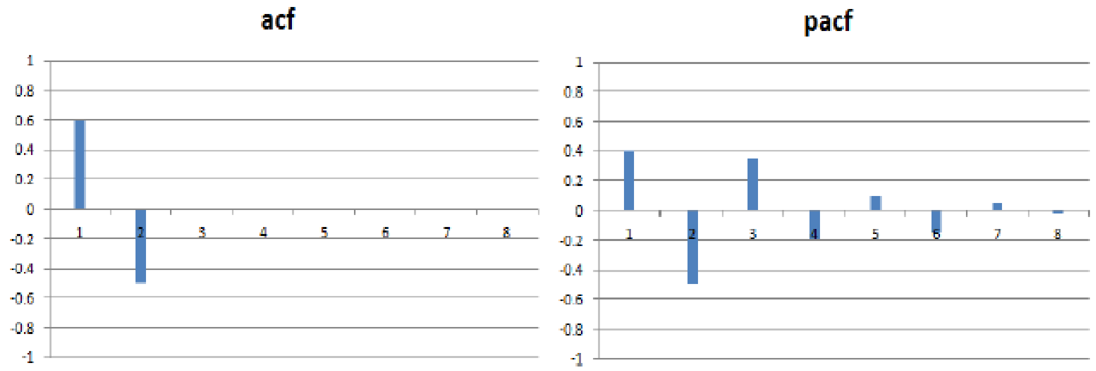
MA(1): $-1 < \theta_1 < 1$

MA(2): $-1 < \theta_2 < 1,\ \theta_1 + \theta_2 > 1, \theta_1 - \theta_2 < 1$

MA(1)



MA(2)



© Robert Bosch GmbH 2018. All rights reserved, also regarding any disposal, exploitation, reproduction, editing, distribution, as well as in the event of applications for industrial property rights.

Forecasting Methods
ARIMA: Recognition (4)

- A ***practical guide*** for estimating $\boldsymbol{\varphi}$ and $\boldsymbol{\theta}$:

ARMA(p, q)	ρ_1	ρ_2
AR(1)	φ_1	
MA(1)	$\frac{-\theta_1}{1+{\theta_1}^2}$	
AR(2)	$\frac{\varphi_1}{1-\varphi_2}$	$\frac{{\varphi_1}^2}{1-\varphi_2}+\varphi_2$
MA(2)	$\frac{-\theta_1(1-\theta_2)}{1+{\theta_1}^2+{\theta_2}^2}$	$\frac{-\theta_2}{1+{\theta_1}^2+{\theta_2}^2}$
ARMA(1,1)	$\frac{(1-\varphi_1\theta_1)(\varphi_1-\theta_1)}{1+{\theta_1}^2-2\varphi_1\theta_1}$	$\rho_1\varphi_1$

Internal | CR/AEU2 - Patelis | 2018-03-14
© Robert Bosch GmbH 2018. All rights reserved, also regarding any disposal, exploitation, reproduction, editing, distribution, as well as in the event of applications for industrial property rights.

Forecasting Methods
ARIMA: Recognition (5)

- Using a constant ***c*** in a ARIMA(p, d, q) model:
  - If ***d=0***, the initial data are ***stationary***.
    - We may need to add a constant **c**, in order to better define the level.
  - If ***d=1***, the initial data have ***constant trend***.
    - We may need to add a constant **c**, in order to fine-tune the level of the timeserie.
    - In practice, this is avoided.
  - If ***d=2***, the initial data have ***non-constant trend***.
    - No constant **c** should be used in such cases.

- Constant **c** has an ***important effect on long-term*** forecasting:
  - c=0, d=0 → forecasts towards zero
  - c=0, d=1 → forecasts towards non-zero constant
  - c=0, d=2 → forecasts follow straight line
  - c<>0, d=0 → forecasts towards mean of data
  - c<>0, d=1 → forecasts follow straight line
  - c<>0, d=2 → forecasts follow quadratic trend

© Robert Bosch GmbH 2018. All rights reserved, also regarding any disposal, exploitation, reproduction, editing, distribution, as well as in the event of applications for industrial property rights.

BOSCH

Forecasting Methods
ARIMA: Evaluation (1)

- ***Likelihood***: the logarithm of the probability of the observed data

$$-2logL = n \times \left[log(2\pi) + 1 + log\left(\frac{RSS}{n}\right)\right]$$

- ***Akaike's Information Criterion (AIC)***:

$$AIC = -2logL + 2(p + q + k + 1)$$

$$AIC_c = AIC + \frac{2(p + q + k + 1)(p + q + k + 2)}{n - p - q - k - 2}$$

- ***Bayesian Information Criterion (BIC)***:

$$BIC = AIC + (logT - 2)(p + q + k + 1)$$

where n = number of observations, RSS=sum of squared errors, k=1 if the model has a constant (else 0)

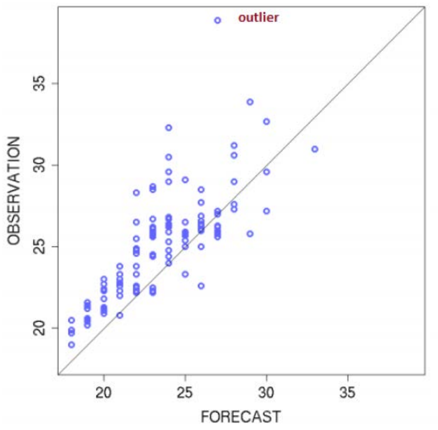
➢ Good models are obtained by ***minimizing*** either the AIC, AIC_c or BIC.

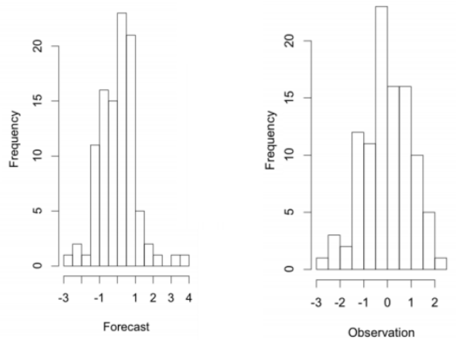
Internal | CR/AEU2 - Patelis | 2018-03-14
© Robert Bosch GmbH 2018. All rights reserved, also regarding any disposal, exploitation, reproduction, editing, distribution, as well as in the event of applications for industrial property rights.

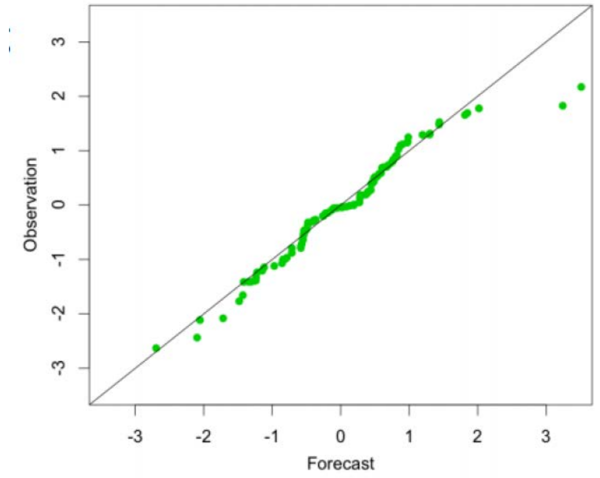
Forecasting Methods
ARIMA: Diagnostics (1)

- ***Scatter plot***
  - Real values *vs* Forecasted values
  - Perfect forecast: when all point are in the diagonal
  - Detects: ***Errors linear correlation, outlies, bias***

- ***Quantile-Quantile plot***
  - Real values quantiles *vs* Forecasted values quantiles
    - Quantile: the fraction (or %) of points below a given value.
  - Perfect forecast: when all point are in the diagonal
  - Detects: ***Errors linear correlation, outlies, bias***



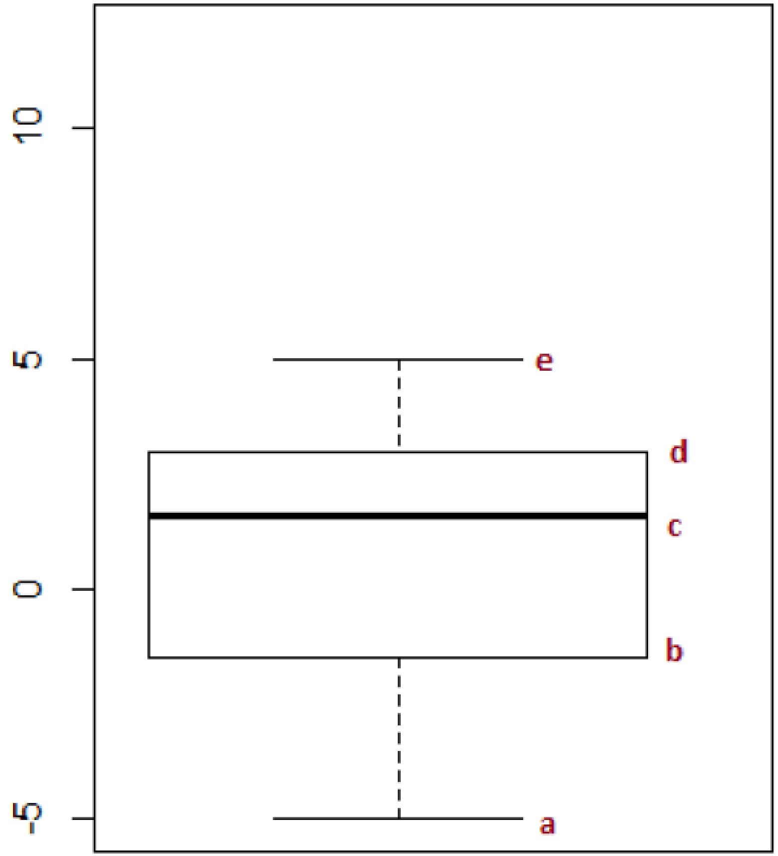




Internal | CR/AEU2 - Patelis | 2018-03-14
© Robert Bosch GmbH 2018. All rights reserved, also regarding any disposal, exploitation, reproduction, editing, distribution, as well as in the event of applications for industrial property rights.
BOSCH

Forecasting Methods
ARIMA: Diagnostics (2)

- ***Box plot***
  - Visualize ***error distribution***, where:
    - a = min
    - b = 25% of the observations
    - c = median
    - d = 25% of the observations
    - e = max
- Best forecasts when:
  - ***All distances are equal*** (a-b = b-c = c-d = d-e)
  - ***Median is zero*** (c = 0)

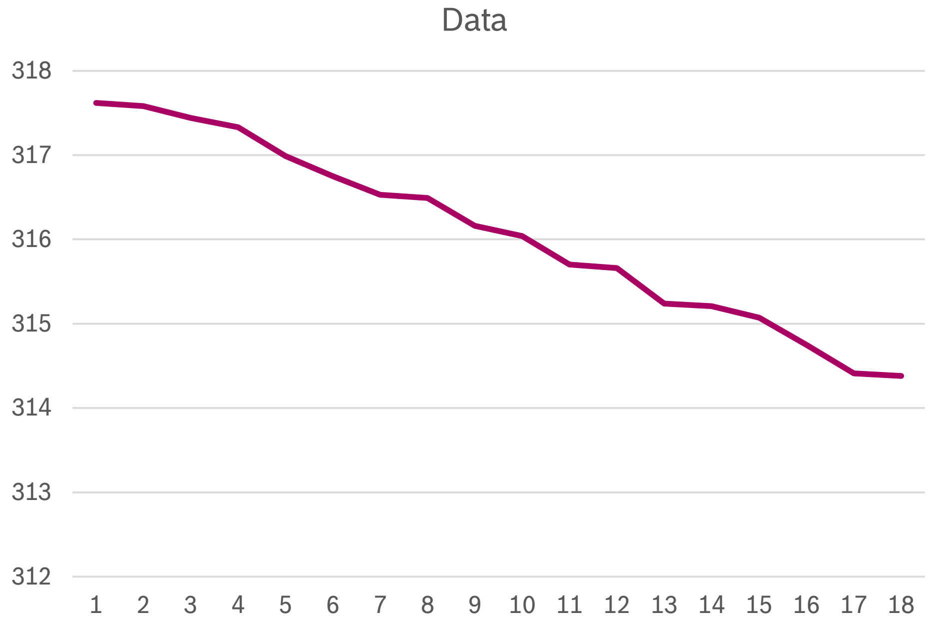


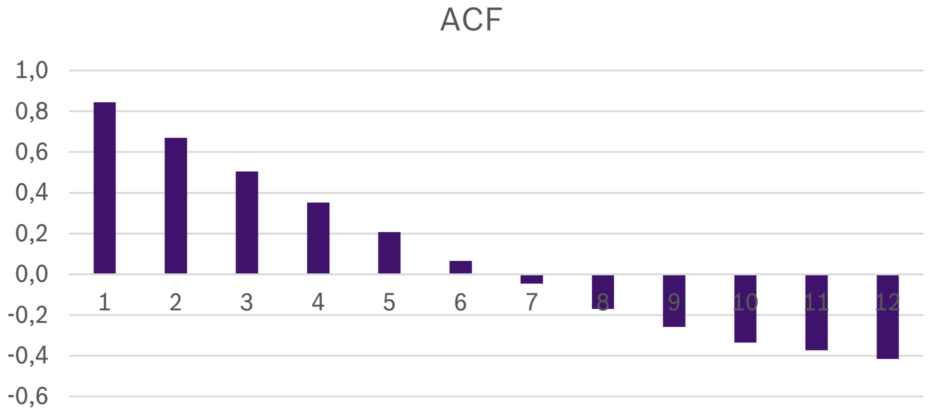
Internal | CR/AEU2 - Patelis | 2018-03-14
© Robert Bosch GmbH 2018. All rights reserved, also regarding any disposal, exploitation, reproduction, editing, distribution, as well as in the event of applications for industrial property rights.

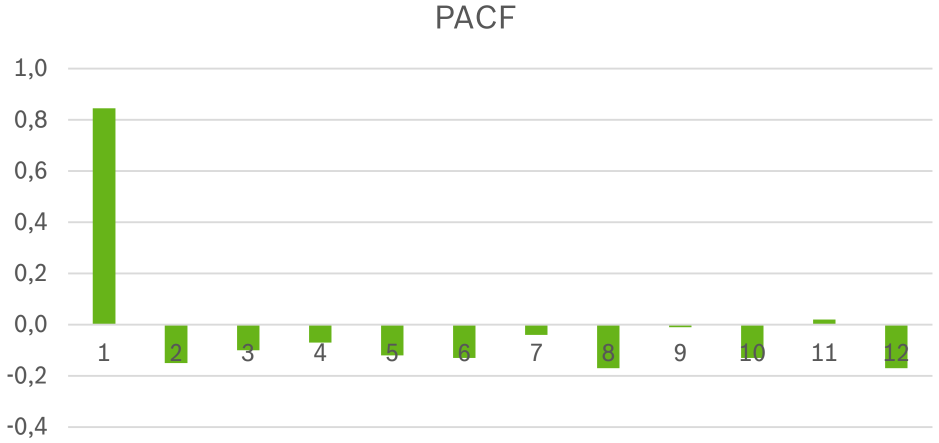
Forecasting Methods
ARIMA: Example (1)

- Estimate forecasts, for horizon = 3
  - Timeserie has a clear trend.

	Data
1	317,62
2	317,58
3	317,44
4	317,33
5	316,99
6	316,75
7	316,53
8	316,49
9	316,16
10	316,04
11	315,70
12	315,66
13	315,24
14	315,21
15	315,07
16	314,75
17	314,41
18	314,38





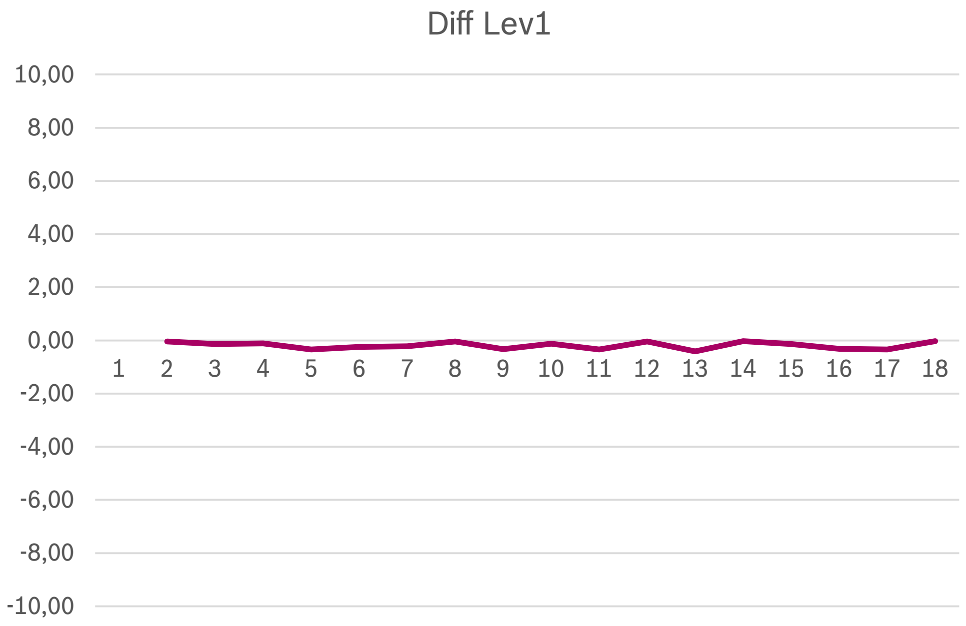


Internal | CR/AEU2 - Patelis | 2018-03-14
© Robert Bosch GmbH 2018. All rights reserved, also regarding any disposal, exploitation, reproduction, editing, distribution, as well as in the event of applications for industrial property rights.

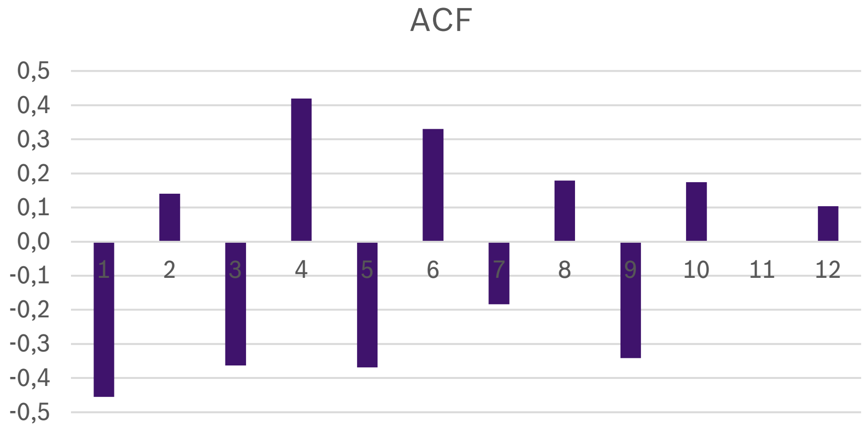
Forecasting Methods
ARIMA: Example (2)

- We need to do ***1st level integration***

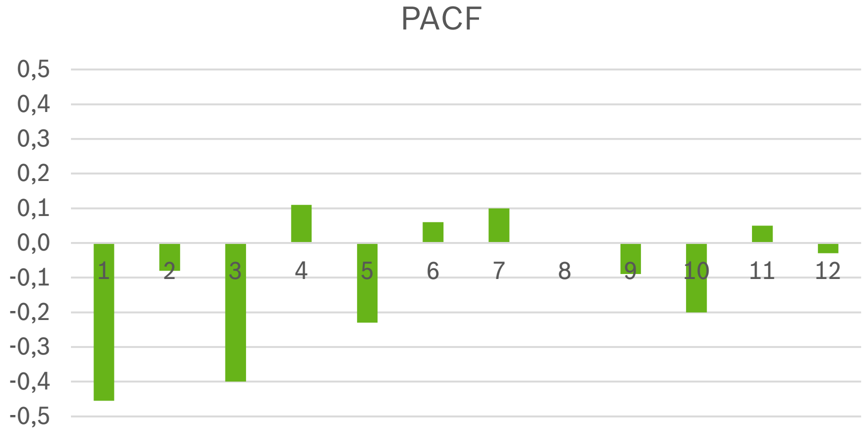
	Diff Lev1
1	
2	-0,04
3	-0,14
4	-0,11
5	-0,34
6	-0,24
7	-0,22
8	-0,04
9	-0,33
10	-0,12
11	-0,34
12	-0,04
13	-0,42
14	-0,03
15	-0,14
16	-0,32
17	-0,34
18	-0,03



- Data are ***stationary***, since average = -0,19



- ACF decreases (sin)



- PACF has large value for lag 1

© Robert Bosch GmbH 2018. All rights reserved, also regarding any disposal, exploitation, reproduction, editing, distribution, as well as in the event of applications for industrial property rights.

Forecasting Methods
ARIMA: Example (3)

- Selected model: AR(1), over 1st level integration data. Thus ***ARIMA(1,1,0)***

			P1	
	Data	Diff Lev1	Nom.	Denom.
1	317,62		-	-
2	317,58	-0,04	-	0,023
3	317,44	-0,14	0,008	0,003
4	317,33	-0,11	0,004	0,006
5	316,99	-0,34	-0,012	0,022
6	316,75	-0,24	0,007	0,002
7	316,53	-0,22	0,001	0,001
8	316,49	-0,04	-0,004	0,023
9	316,16	-0,33	-0,021	0,019
10	316,04	-0,12	-0,010	0,005
11	315,70	-0,34	-0,011	0,022
12	315,66	-0,04	-0,022	0,023
13	315,24	-0,42	-0,035	0,053
14	315,21	-0,03	-0,037	0,026
15	315,07	-0,14	0,008	0,003
16	314,75	-0,32	-0,007	0,017
17	314,41	-0,34	0,019	0,022
18	314,38	-0,03	-0,024	0,026
sum			-0,134	0,295
average	316,075	-0,191	P1=	**-0,455**

$$\varphi_1 = p_1 = -0{,}455$$
$$c = \bar{\acute{Y}} \times (1 - \varphi_1) = -0{,}277$$

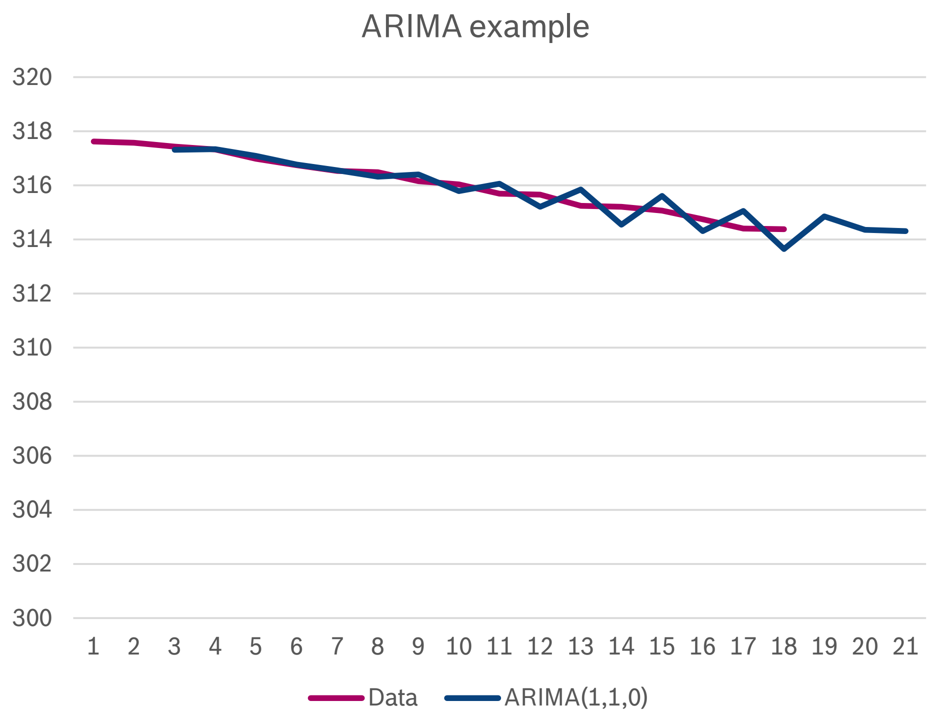
$$\acute{Y}_t = c + \varphi_1 \times \acute{Y}_{t-1} + e_t$$
$$Y_t - Y_{t-1} = c + \varphi_1 \times (Y_{t-1} - Y_{t-2}) + e_t$$
$$Y_t = c + \varphi_1 \times Y_{t-1} - \varphi_1 \times Y_{t-2} + e_t + Y_{t-1}$$
$$Y_t = -0{,}277 + 0{,}545 \times Y_{t-1} + 0{,}455 \times Y_{t-2} + e_t$$

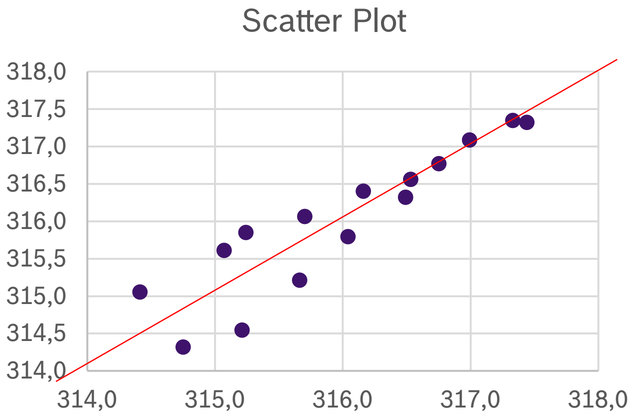
© Robert Bosch GmbH 2018. All rights reserved, also regarding any disposal, exploitation, reproduction, editing, distribution, as well as in the event of applications for industrial property rights.
BOSCH

Forecasting Methods
ARIMA: Example (4)

- Estimate forecasts

	Data	ARIMA(1,1,0)	Error
1	317,62		
2	317,58		
3	317,44	317,32	0,12
4	317,33	317,35	-0,02
5	316,99	317,09	-0,10
6	316,75	316,77	-0,02
7	316,53	316,56	-0,03
8	316,49	316,32	0,17
9	316,16	316,40	-0,24
10	316,04	315,79	0,25
11	315,70	316,06	-0,36
12	315,66	315,21	0,45
13	315,24	315,85	-0,61
14	315,21	314,55	0,66
15	315,07	315,61	-0,54
16	314,75	314,32	0,43
17	314,41	315,05	-0,64
18	314,38	313,65	0,73
19		314,85	
20		314,36	
21		314,31	





- Mean Error = 0,016
- Median error = -0,018

Internal | CR/AEU2 - Patelis | 2018-03-14
© Robert Bosch GmbH 2018. All rights reserved, also regarding any disposal, exploitation, reproduction, editing, distribution, as well as in the event of applications for industrial property rights.

MULTIPLE REGRESSION

Forecasting Methods
Multiple Regression

- ***Multiple Regression (MR)***: When several independent variables are required

$$Y = b_0 + b_1 * X_1 + b_2 * X_2 + \cdots + b_k * X_k + e$$

- For estimating the weights, we must minimize the error:

$$Y_i = b_0 + b_1 * X_1 + b_2 * X_2 + e_i = \widehat{Y}_i + e_i$$

$$(b_0, b_1, b_2) \mid \min \left[\sum_{i=1}^{n} e_i{}^2 \right]$$

$$\sum_{i=1}^{n} e_i{}^2 = \sum_{i=1}^{n} \left(Y_i - \widehat{Y}_i\right)^2 = \sum_{i=1}^{n} \left(Y_i - b_0 - b_1 * X_{1,i} - b_2 * X_{2,i}\right)^2$$

- ***For 2 independent variables, we need to estimate 3 parameters***
  - calculate the partial derivatives for each one of the weights, assign the calculated derivatives equals with zero, and then solve a three-equation problem with three parameters.

- ***Important issue on multiple regression***: the stability of the regression weights is depending on the correlation of the independent variables. For two independent variables $\boldsymbol{X_1}$ and $\boldsymbol{X_2}$, the greater the correlations between them, the more unstable are the weights $\boldsymbol{b_1}$ and $\boldsymbol{b_2}$.

© Robert Bosch GmbH 2018. All rights reserved, also regarding any disposal, exploitation, reproduction, editing, distribution, as well as in the event of applications for industrial property rights.
BOSCH

Forecasting Methods
Regression: How to define Predictors (1)

- ***Dummy variables***
  - In most cases, the predictor variables are numerical.
  - But, what if a predictor is a ***categorical variable*** (true/false, yes/no)? Such a variable might arise in a forecasting problem.
    - This predictor can be handled by creating a dummy variable taking values on 1 and 0.
    - If the predictor has more values, then the variable can be coded using several dummy variables.
  - Example:
    - ***6 dummy variables are needed*** in order to use the ***day of week*** as a predictor. That is because the last week day (Sunday) will be specified when all 6 dummy variables are set to zero. Many beginners will try to add a seventh dummy variable for the seventh category.
    - This is known as the "***dummy variable trap***" because it will cause the regression to fail. There will be too many parameters to estimate. The general rule is to use ***one fewer dummy variables than categories***.

© Robert Bosch GmbH 2018. All rights reserved, also regarding any disposal, exploitation, reproduction, editing, distribution, as well as in the event of applications for industrial property rights.
BOSCH

Forecasting Methods
Regression: How to define Predictors (2)

- ***Outliers***
  - If there is an outlier in the data, rather than omit it, it can be used as a dummy variable to remove its effect.
    - In this case, the dummy variable takes value one for that observation and zero everywhere else.
- ***Public holidays***
  - ***Static***: For daily data, the effect of public holidays can be accounted for by:
    - including a dummy variable predictor taking value one on public holidays and zero elsewhere.
  - ***Moving***: Easter is different from most holidays because it is not held on the same date each year and the effect can last for several days.
    - a dummy variable can be used with value one where any part of the holiday falls in the particular time period and zero otherwise.

Internal | CR/AEU2 - Patelis | 2018-03-14
© Robert Bosch GmbH 2018. All rights reserved, also regarding any disposal, exploitation, reproduction, editing, distribution, as well as in the event of applications for industrial property rights.
BOSCH

Forecasting Methods
Regression: How to define Predictors (3)

- ***Working (trading) days***
  - The number of trading days in a month can vary considerably and can have a substantial effect on sales data.
    - To allow for this, the number of trading days in each month can be included as a predictor.
- **Intervention variables**
  - It is often necessary to model ***interventions*** that may have affected the variable to be forecast.
    - For example, competitor activity, advertising expenditure, industrial action, and so on, can all have an effect.
  - When the effect lasts ***only for one period***, a ***spike variable*** can be used.
    - Thus, a dummy variable taking value one in the period of the intervention and zero elsewhere.
    - A spike variable is equivalent to a dummy variable for handling an outlier.
  - When the effect is ***immediate and permanent*** (i.e. a level shift, the value of the series changes suddenly and permanently from the time of intervention), a ***step variable*** can be used.
    - A step variable takes value zero before the intervention and one from the time of intervention onwards.

© Robert Bosch GmbH 2018. All rights reserved, also regarding any disposal, exploitation, reproduction, editing, distribution, as well as in the event of applications for industrial property rights.

BOSCH

Forecasting Methods
Multiple Regression: Correlation Coefficient

- For MR case, we need to take into account the number of independent variables ***k*** and the number of observations ***n***. Thus, to use an ***"adjusted" coefficient***.

$$R^2 = \frac{\sum_{i=1}^{n}\left(\hat{Y}_i - \bar{Y}\right)^2}{(Y_i - \bar{Y})^2}$$

$$\bar{R}^2 = 1 - \ (1 - R^2) * \frac{n-1}{n-k-1}$$

- This "adjusted" coefficient ***R²*** represents the percentage of the ***dispersion*** of the variable *Y*, which is justified by the independent variables.
  - the different ***(n-1)*** is the total degrees of freedom of the total variance of the model.
  - the ***(n-k-1)*** is the degrees of freedom of the interpreted variance.

Internal | CR/AEU2 - Patelis | 2018-03-14
© Robert Bosch GmbH 2018. All rights reserved, also regarding any disposal, exploitation, reproduction, editing, distribution, as well as in the event of applications for industrial property rights.
BOSCH

Forecasting Methods
Multiple Regression: Statistical Indexes (1)

- Index ***F***, for MR case:

$$F = \frac{\frac{\Sigma\left(\hat{Y}_t - \bar{Y}\right)^2}{k}}{\frac{\Sigma\left(Y_t - \hat{Y}_t\right)^2}{n - k - 1}} \qquad F = \frac{\frac{R^2}{k}}{\frac{1 - R^2}{n - k - 1}}$$

- The value of index ***F*** depends in the size of the numerator and denominator:
  - If the non-interpreted variance (error variance) is large, then the denominator is large and the F index gets smaller. ***That means that the regression model is not successful***.
  - If the interpreted variance is relatively large, then the nominator is large and the F index gets larger. That means that there is a close relation between the coefficient $\boldsymbol{R^2}$ and the statistical index ***F***.

Internal | CR/AEU2 - Patelis | 2018-03-14
© Robert Bosch GmbH 2018. All rights reserved, also regarding any disposal, exploitation, reproduction, editing, distribution, as well as in the event of applications for industrial property rights.

Forecasting Methods
Multiple Regression: Statistical Indexes (2)

- Index ***t***, for MR case:
  - sometimes useful to examine the significance of each of the regression weights
  - the statistical index ***t***, for a particular rate, is an estimate of the significance of this factor within the presence of all other independent variables. For each regression weight $\boldsymbol{b_j}$, a standard error can be set (as a measure of the stability of the weight) and, based on the normality of the regression model, the index ***t*** follows the t-distribution with ***(n-k-1)*** degrees of freedom (as given in the following equation).

$$t_{b_j} = \frac{b_j}{SE_{b_j}}$$

  - By using this equation for each weight of the regression model, we can calculate its significance, through the comparison of the value and the value of 0, thus with the value of which the corresponding independence variable does not contribute to the prediction of ***Y***, given the presence of all other independent variables.

Internal | CR/AEU2 - Patelis | 2018-03-14
© Robert Bosch GmbH 2018. All rights reserved, also regarding any disposal, exploitation, reproduction, editing, distribution, as well as in the event of applications for industrial property rights.
BOSCH

Forecasting Methods
Regression: Basic assumptions (1)

1. The existence of a ***linear relationship*** between the dependent and the independent variables.
   - When this assumption is not valid, the independent variables must be transformed to other variables which have linear relationship with the dependent variable *Y*.
2. ***Constant variance of regression errors*** ("homoscedasticity", "heteroscedasticity").
   - This assumption indicates that the forecast errors should be stable throughout the range of observations.
3. The ***residual errors are independent*** of one another.
   - This means that the price of each residual error is independent of the values of the previous and the next.
   - When this assumption is not fulfilled, there is a serial correlation (or a auto-correlation) between consecutive values of residual errors. **Thus, a significant independent variable has been skipped!**.
     - Alternative ways of recognizing the independence of the residual errors is the graphical representation of their values and the calculation of the Durbin-Watson statistical index.

© Robert Bosch GmbH 2018. All rights reserved, also regarding any disposal, exploitation, reproduction, editing, distribution, as well as in the event of applications for industrial property rights.
BOSCH

Forecasting Methods
Regression: Basic assumptions (2)

4. An important issue in MR is the probability of multi-collinearity.
   - The multi-collinearity arises ***when two or more independent variables are highly correlated***.
   - This is a frequent problem in financial and operational data, due to the high degree of correlation that exists between the different factors.
   - This fact should be taken into account, when selecting the independent variables and the data collection.
   - The goal is the use of independent variables which are not strongly correlated
     - a rule of thumb is that the correlation ***should not exceed the value 0.7 or less than -0.7***.

   - If the independent variables are strongly correlated, then they provide ***redundant information*** which does not improve the explanatory strength of the regression.

Internal | CR/AEU2 - Patelis | 2018-03-14
© Robert Bosch GmbH 2018. All rights reserved, also regarding any disposal, exploitation, reproduction, editing, distribution, as well as in the event of applications for industrial property rights.
BOSCH

Forecasting Methods
Regression: How NOT to select predictors (1)

- ***Visual inspection***
  - A common approach is to plot the forecast variable against the specific predictor and to visual examine their relationship.
    - This is an approach that can lead to faults, because it is not always possible to see the relationship from a graph. It should be avoided.
- ***MLR on all predictors***
  - Another common approach is to perform multiple linear regression on all the predictors and discharged the predictors with ***p*** value ***greater than 0.05***.
    - This can also lead to faults and should be avoided, since statistical significance does not always indicate predictive value.
    - Also, ***p*** values can be misleading when two or more predictors are correlated with each other.

© Robert Bosch GmbH 2018. All rights reserved, also regarding any disposal, exploitation, reproduction, editing, distribution, as well as in the event of applications for industrial property rights.
BOSCH

Forecasting Methods
Regression: How NOT to select predictors (2)

- ***R^2***
  - It is not a good approach to select predictors based on R^2 value.
    - If a model produces forecasts that are always 30% higher than the actual values, then R^2 is equal to 1 indicating a perfect correlation. But the forecasts are not close to the actual values.
    - Also, ***R^2 does not allow for degrees of freedom.*** Adding any variable tends to increase the R^2 value, even if the new variable is irrelevant.
- ***Mean Square Error (MSE)***
  - Another approach is to select predictors based on minimizing the mean of square errors (MSE) or sum of square errors (SSE).
    - But, minimizing MSE or SSE is equivalent to maximizing R^2 and will always close the model with the most variables.

© Robert Bosch GmbH 2018. All rights reserved, also regarding any disposal, exploitation, reproduction, editing, distribution, as well as in the event of applications for industrial property rights.
BOSCH

Forecasting Methods
Regression: How to select predictors (1)

- ***Adjusted R^2***
  - An alternative approach for selecting predictors in the adjusted R^2

$$\bar{R}^2 = 1 - (1 - R^2) * \frac{n-1}{n-k-1}$$

where n is the number of observations and k is the number of predictors.

  - This is improving R^2, since it will no longer increase with each added predictor. The best set of predictors is the one which **maximizes** the value of adjusted R^2.
  - This approach works quite well as a method for selecting predictors, although it does ***tent sometimes*** to err on the side of selecting too many predictors.
- ***Variance of errors***
  - An equivalent approach for selecting predictors, in the minimization of the variance of the forecasts errors

$$\hat{\sigma}^2 = \frac{Sum\ of\ Squared\ Errors\ (SSE)}{n-k-1}$$

© Robert Bosch GmbH 2018. All rights reserved, also regarding any disposal, exploitation, reproduction, editing, distribution, as well as in the event of applications for industrial property rights.

BOSCH

Forecasting Methods
Regression: How to select predictors (2)

- ***Cross-Validation***
  - a useful approach for examining the predictive ability of a forecast model. In general, leave-one-out cross-validation must follow the following steps:
    - Remove observation 1 for the data set, and fit the model using the remaining data. Then, forecast the value for observation 1 and estimate the ***cross-validation error*** (it is not the same as residual, since the 1st observation was not used for estimating forecasts).

$$e_1^* = Y_1 - \widehat{Y_1}$$

    - Remove the previous step for all observations of the data step, thus 1,2, ... n
    - Estimate the mean square error (MSE) from all cross-validation errors (CVe)
    - Repeat all previous steps for each model (for each predictors variables set).
  - ***The best model is the one with the minimum CVe.***

© Robert Bosch GmbH 2018. All rights reserved, also regarding any disposal, exploitation, reproduction, editing, distribution, as well as in the event of applications for industrial property rights.

Forecasting Methods

Regression: How to select predictors (3)

- ***Akaike's Information Criterion (AIC)***
  - Another approach is Akaike's information criterion, which can be estimated as: $AIC = n * \log\left(\frac{SSE}{n}\right) + 2 * (k+2)$

    , where n is the number of observations and k the number of predictors in the model.
  - The *k+2* part of the equation is because there are k+2 parameters in the model (the k coefficients for the predictors, the intercept and the variance of the residuals).
  - ***The model with the minimum AIC value is often the best model for forecasting.***
- ***Corrected AIC$_c$***
  - For small amount of observations n, the AIC tends to select too many predictors.
  - So, a corrected AIC can be used: $AIC_C = AIC + \frac{2*(k+2)*(k+3)}{n-k-3}$
  - ***The model with the minimum AIC$_c$ value is often the best model for forecasting.***

Internal | CR/AEU2 - Patelis | 2018-03-14
© Robert Bosch GmbH 2018. All rights reserved, also regarding any disposal, exploitation, reproduction, editing, distribution, as well as in the event of applications for industrial property rights.

Forecasting Methods
Regression: How to select predictors (4)

- ***Schwarz Bayesian Information Criterion (SBIC)***
  - Another approach is the Schwarz Bayesian information criterion, which can be calculated as:

$$SBIC = n * \log\left(\frac{SSE}{n}\right) + \ (k+2) * \log n$$

  - ***As with AIC, the best model is the one with the minimum SBIC.***
  - The model chosen by SBIC is either the same as that chosen by AIC, or one with fewer predictors.
  - This is because SBIC penalizes the SSE (sum of squared errors) more heavily that AIC.
    - Many statisticians like to use SBIC because it has the feature that ***if there is a true underlying model, then with enough data the SBIC will select that model.***
    - However, in reality there is rarely if ever a true underlying model, and even if there was a true underlying model, selecting that model will not necessarily give the best forecasts (since the predictors values must also be forecasted, and these forecasts may not be accurate!).

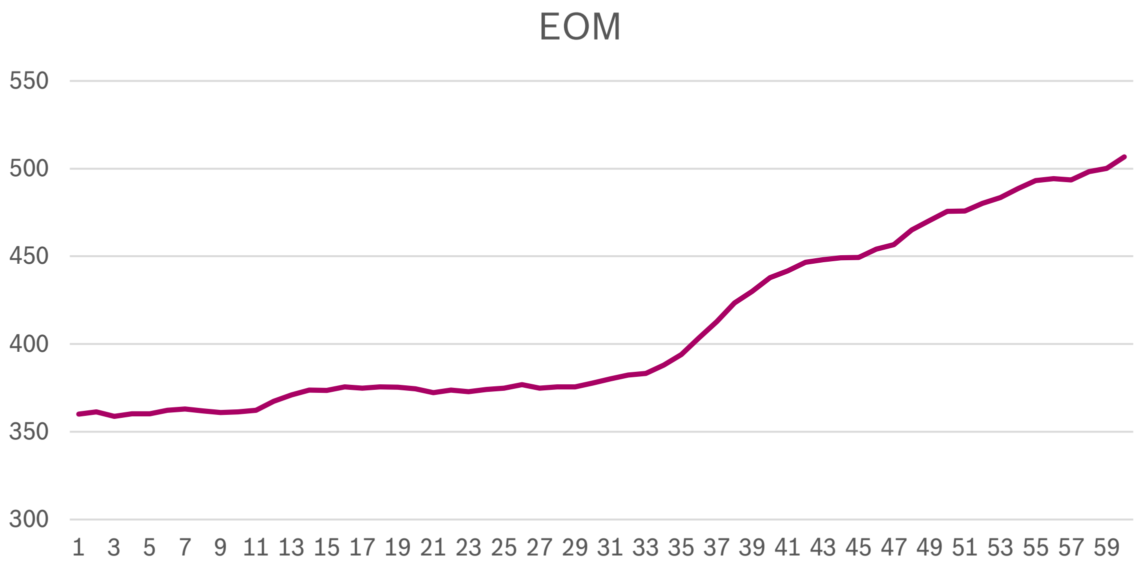
  - It must be mentioned that, SBIC ***does tent sometimes to err on the side of selecting too few predictors.***

© Robert Bosch GmbH 2018. All rights reserved, also regarding any disposal, exploitation, reproduction, editing, distribution, as well as in the event of applications for industrial property rights.

Forecasting Methods
Multiple Regression: Example (1)

- ***Mutual savings bank deposit in a large metropolitan area***
  - Problem: Monthly changes in deposits were getting smaller and monthly changes in withdrawals were getting bigger. It was of interest to develop a short-term forecasting model to forecast the changes in end-of-month.

Months	EOM	Months	EOM	Months	EOM	Months	EOM	Months	EOM
1	360,071	13	371,031	25	374,880	37	412,727	49	470,408
2	361,217	14	373,734	26	376,735	38	423,417	50	475,600
3	358,774	15	373,463	27	374,841	39	429,948	51	475,857
4	360,271	16	375,518	28	375,622	40	437,821	52	480,259
5	360,139	17	374,804	29	375,461	41	441,703	53	483,432
6	362,164	18	375,457	30	377,694	42	446,663	54	488,536
7	362,901	19	375,423	31	380,119	43	447,964	55	493,182
8	361,878	20	374,365	32	382,288	44	449,118	56	494,242
9	360,922	21	372,314	33	383,270	45	449,234	57	493,484
10	361,307	22	373,765	34	387,978	46	454,162	58	498,186
11	362,290	23	372,776	35	394,041	47	456,692	59	500,064
12	367,382	24	374,134	36	403,423	48	465,117	60	506,684

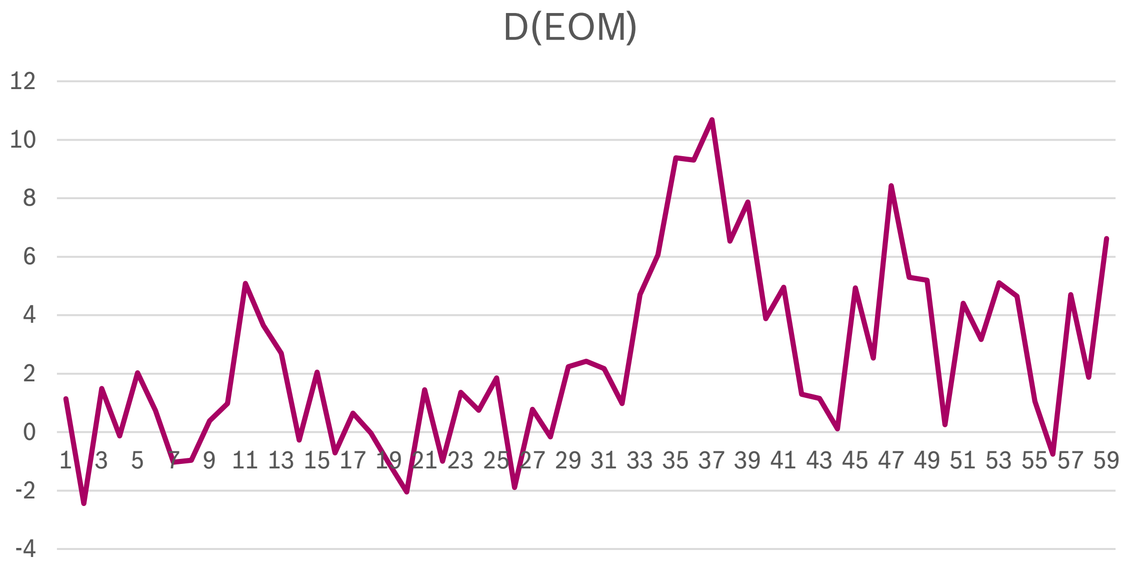


Source: "Forecasting: Methods and Applications", Spyros Makridakis, Steven C. Wheelwright

Internal | CR/AEU2 - Patelis | 2018-03-14
© Robert Bosch GmbH 2018. All rights reserved, also regarding any disposal, exploitation, reproduction, editing, distribution, as well as in the event of applications for industrial property rights.
BOSCH

Forecasting Methods
Multiple Regression: Example (2)

- The interest of the bank was the change in the EOM balance and so first differences of the EOM data are estimated.



- It is clear that the bank was facing a ***volatile situation*** in the last two years or so.

- The challenge is to forecast these rapidly changing EOM values.

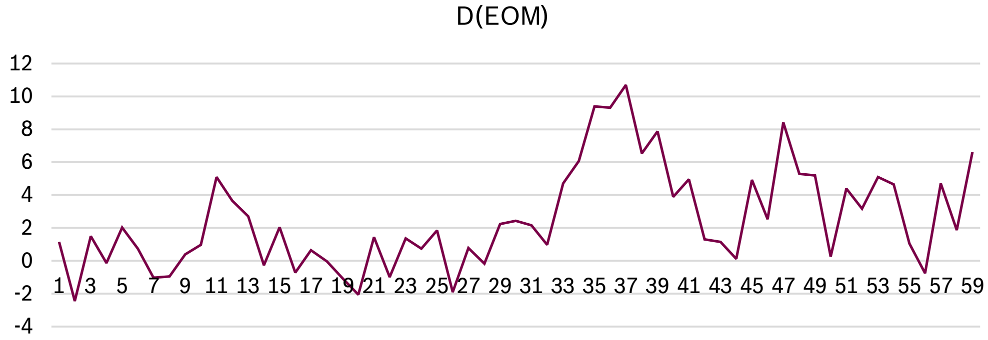
© Robert Bosch GmbH 2018. All rights reserved, also regarding any disposal, exploitation, reproduction, editing, distribution, as well as in the event of applications for industrial property rights.

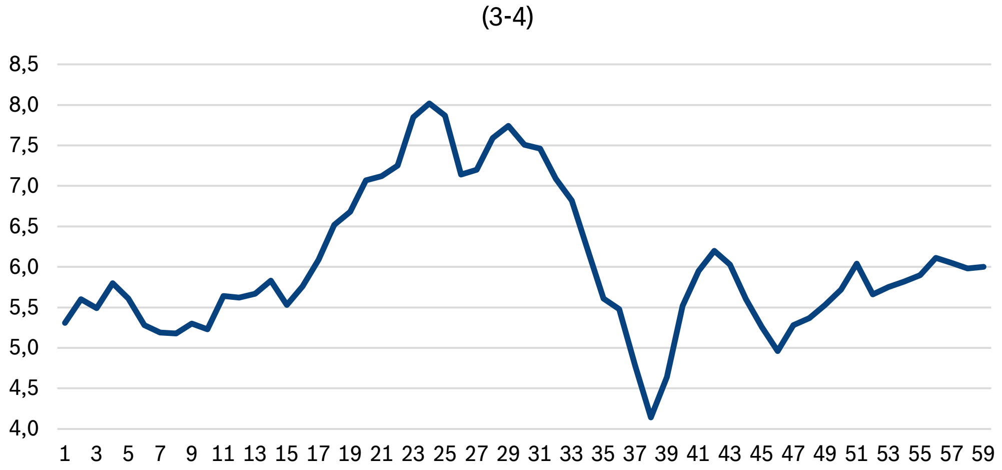
Forecasting Methods
Multiple Regression: Example (3)

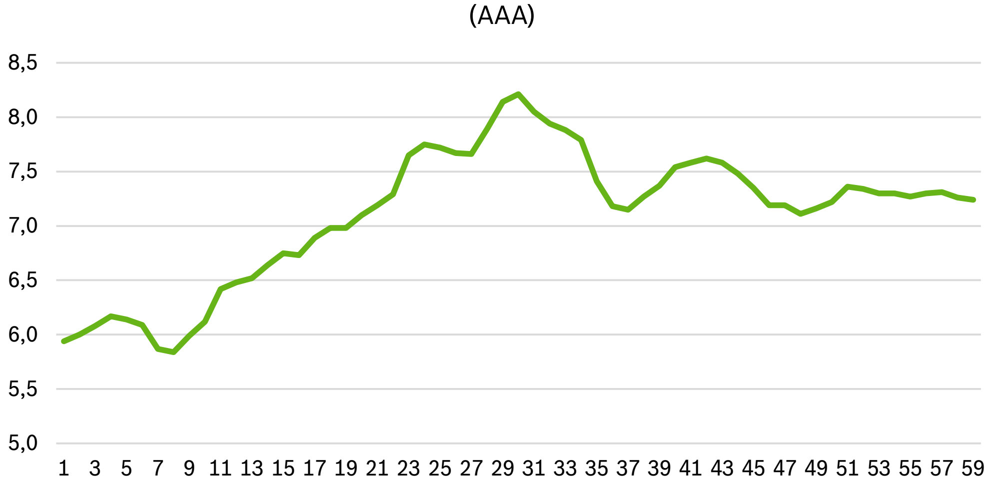
- It was hypothesized that two variables ***had an influence*** on the EOM balance figures:
  - Composite AAA bond rates, and
  - The rates on U.S. Government 3-4 year bonds.
- ***Three explanatory*** variables (regressors):
  - X_1: the AAA bond rates, but they are now shown leading the D(EOM) values
  - X_2: the rates on 3-4 year government bonds, but they are shown leading the D(EOM) values by 1 month.
  - X_3: The first differences of the 3-4 year bond rates, and the timing for this variable coincides with that of the D(EOM) variable.
- ***D(EOM) is the forecast value***

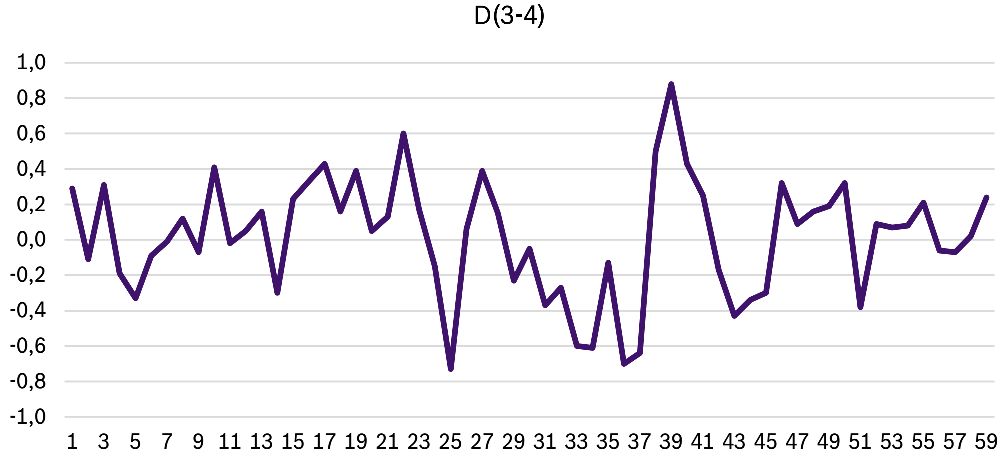
© Robert Bosch GmbH 2018. All rights reserved, also regarding any disposal, exploitation, reproduction, editing, distribution, as well as in the event of applications for industrial property rights.
BOSCH

Forecasting Methods
Multiple Regression: Example (4)





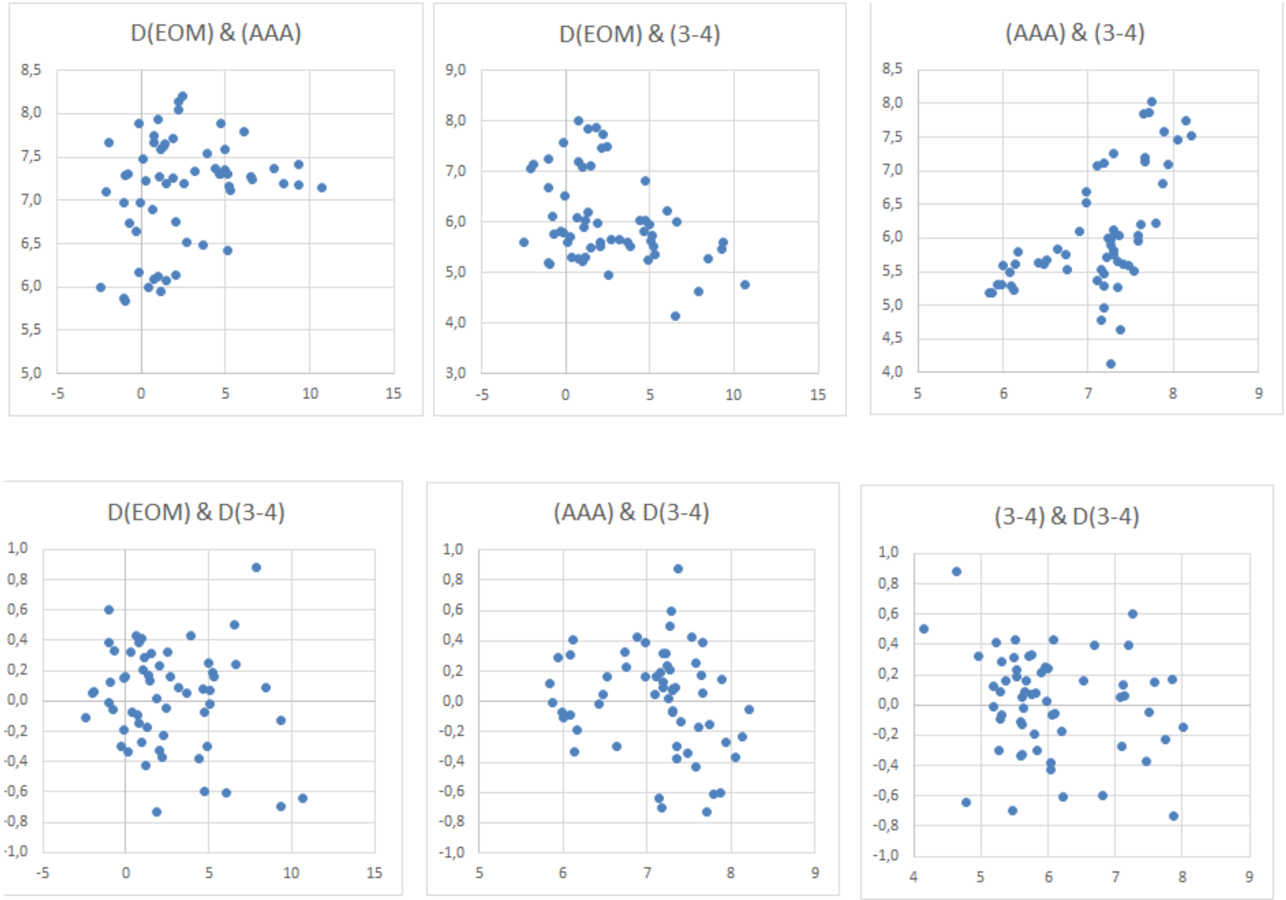




© Robert Bosch GmbH 2018. All rights reserved, also regarding any disposal, exploitation, reproduction, editing, distribution, as well as in the event of applications for industrial property rights.

Forecasting Methods
Multiple Regression: Example (5)

- ***Scatterplots***:
  - Visualize the relationship between each pair of values

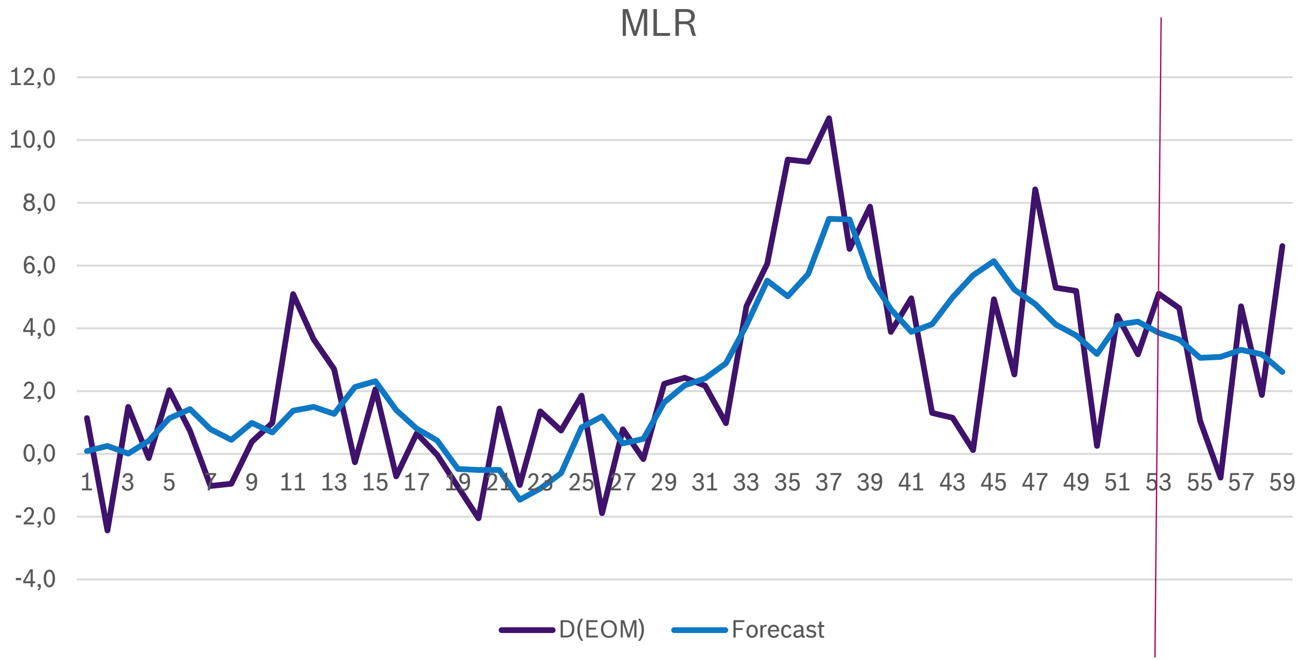


Internal | CR/AEU2 - Patelis | 2018-03-14
© Robert Bosch GmbH 2018. All rights reserved, also regarding any disposal, exploitation, reproduction, editing, distribution, as well as in the event of applications for industrial property rights.

Forecasting Methods
Multiple Regression: Example (6)

- ***Model Estimation***: By using a least square error, we can estimate:
  - $Y = -4.3391 + 3.3722 * X_1 - 2.8316 * X_2 - 1.9648 * X_3$

	Y	X			Y'
Months	D(EOM)	(AAA)	(3-4)	D(3-4)	Forecast
1	1,146	5,940	5,310	0,290	0,086
2	-2,443	6,000	5,600	-0,110	0,253
3	1,497	6,080	5,490	0,310	0,009
4	-0,132	6,170	5,800	-0,190	0,417
5	2,025	6,140	5,610	-0,330	1,129
6	0,737	6,090	5,280	-0,090	1,424
7	-1,023	5,870	5,190	-0,010	0,779
8	-0,956	5,840	5,180	0,120	0,451
9	0,385	5,990	5,300	-0,070	0,990
10	0,983	6,120	5,230	0,410	0,684
11	5,092	6,420	5,640	-0,020	1,379
12	3,649	6,480	5,620	0,050	1,501
13	2,703	6,520	5,670	0,160	1,278
14	-0,271	6,640	5,830	-0,300	2,134
15	2,055	6,750	5,530	0,230	2,313
16	-0,714	6,730	5,760	0,330	1,397
17	0,653	6,890	6,090	0,430	0,806
45	4,928	7,350	5,260	-0,300	6,142
46	2,530	7,190	4,960	0,320	5,234
47	8,425	7,190	5,280	0,090	4,779
48	5,291	7,110	5,370	0,160	4,117
49	5,192	7,160	5,530	0,190	3,774
50	0,257	7,220	5,720	0,320	3,183
51	4,402	7,360	6,040	-0,380	4,124
52	3,173	7,340	5,660	0,090	4,209
53	5,104	7,300	5,750	0,070	3,859



Internal | CR/AEU2 - Patelis | 2018-03-14
© Robert Bosch GmbH 2018. All rights reserved, also regarding any disposal, exploitation, reproduction, editing, distribution, as well as in the event of applications for industrial property rights.
BOSCH

Forecasting Methods
Multiple Regression: Example (7)

- ***Coefficient of Determination R²***:
  - For estimating the multiple regression coefficient, we can estimate the correlation between D(EOM) and Forecasts. Thus:
    - $R^2 = r_{Y\hat{Y}}{}^2 = 0.749^2 = 0.561$
  - Alternative:
    - $R^2 = \frac{\sum_{i=1}^{53}(\hat{Y}_i - \bar{Y})^2}{\sum_{i=1}^{53}(Y_i - \bar{Y})^2} = \frac{explained\ SumSquaredDev}{total\ SumSquaredDev} = \frac{SSR}{SST} = \frac{280.389}{499.986} = 0.561$

	Y	Y'	SSE		SSR		SST	
Months	D(EOM)	Forecast	Y' - Y	(Y' - Y)^2	Y' - AvgY	(Y' - AvgY)^2	Y-AvgY	(Y-AvgY)^2
1	1,146	0,086	1,060	1,123	-2,338	5,465	-1,278	1,633
2	-2,443	0,253	-2,696	7,270	-2,171	4,712	-4,867	23,686
3	1,497	0,009	1,488	2,213	-2,415	5,830	-0,927	0,859
4	-0,132	0,417	-0,549	0,302	-2,006	4,026	-2,556	6,532
5	2,025	1,129	0,896	0,802	-1,295	1,676	-0,399	0,159
6	0,737	1,424	-0,687	0,471	-1,000	1,001	-1,687	2,846
7	-1,023	0,779	-1,802	3,248	-1,645	2,704	-3,447	11,881
8	-0,956	0,451	-1,407	1,980	-1,973	3,892	-3,380	11,424
9	0,385	0,990	-0,605	0,367	-1,433	2,055	-2,039	4,157
44	0,116	5,696	-5,580	31,137	3,272	10,707	-2,308	5,326
45	4,928	6,142	-1,214	1,473	3,718	13,823	2,504	6,271
46	2,530	5,234	-2,704	7,309	2,810	7,894	0,106	0,011
47	8,425	4,779	3,646	13,291	2,355	5,548	6,001	36,014
48	5,291	4,117	1,174	1,378	1,693	2,867	2,867	8,220
49	5,192	3,774	1,418	2,011	1,350	1,822	2,768	7,663
50	0,257	3,183	-2,926	8,560	0,759	0,576	-2,167	4,695
51	4,402	4,124	0,278	0,077	1,700	2,891	1,978	3,913
52	3,173	4,209	-1,036	1,074	1,785	3,187	0,749	0,561
53	5,104	3,859	1,245	1,551	1,435	2,059	2,680	7,183
SUM	128.465	128.465	0,004	219,605	-0,004	280,389	0,000	499,986
AVG	2,424							

- The sums and the average value of D(EOM) and Forecasts are the same.
- The sum of residuals (errors) is zero, as it will always be for fitting of linear regression models.
- Also, the coefficient of determination is highly significant.

© Robert Bosch GmbH 2018. All rights reserved, also regarding any disposal, exploitation, reproduction, editing, distribution, as well as in the event of applications for industrial property rights.

Forecasting Methods
Multiple Regression: Example (8)

- ***Index F***:
  - Construct an overall F-test to check on the statistical significance of the regression model.

$$F = \frac{MSR}{MSE} = \frac{\sum(\hat{Y} - \bar{Y})^2 / k}{\sum(Y - \bar{Y})^2 / (n - k - 1)} = \frac{R^2 / k}{(1 - R^2)/(n - k - 1)} = \frac{(0.561)/3}{(1 - 0.561)/49} = 20.85$$

Where:

- k = number of variables
- n = number of observations
- R^2 = coefficient of determinator

➢ A high significant regression model.

➢ The explanatory variables explain a significant amount of the variability in the change in end-of-month deposits at this mutual savings bank.

© Robert Bosch GmbH 2018. All rights reserved, also regarding any disposal, exploitation, reproduction, editing, distribution, as well as in the event of applications for industrial property rights.
BOSCH

Forecasting Methods
Multiple Regression: Example (9)

- ***Index t***:
  - For each regression coefficient, we can estimate t-index for estimating the significance of the coefficient.

Term	Coeff.	Value	se	t	P-value
Constant	b_0	-4,3391	3,259	-1,3314	0,1892
(AAA)	b_1	3,3722	0,556	6,0649	0,0000
(3-4)	b_2	-2,8316	0,3895	-7,2694	0,0000
D(3-4)	b_3	-1,9648	0,8627	-2,2773	0,0272

- The b_1 coefficient is very significantly different from zero.
  - Thus the (AAA) variable is a significant explanatory variable in the presence of the other two explanatory variables.
- Similarly, b_2 and b_3 for variables (3-4) and D(3-4) are highly significant.

Internal | CR/AEU2 - Patelis | 2018-03-14
© Robert Bosch GmbH 2018. All rights reserved, also regarding any disposal, exploitation, reproduction, editing, distribution, as well as in the event of applications for industrial property rights.
BOSCH

Forecasting Methods
Multiple Regression: Example (10)

- Correlation between variables

COR	D(EOM)	(AAA)	(3-4)	D(3-4)
D(EOM)	1,000	0,257	-0,391	-0,195
(AAA)	0,257	1,000	0,587	-0,204
(3-4)	-0,391	0,587	1,000	-0,201
D(3-4)	-0,195	-0,204	-0,201	1,000

  - None of the explanatory variables has a particularly high correlation with the Y value D(EOM).
  - The explanatory variables themselves do not correlate very highly.
  - We suspect no multicollinearity problem (since 0,587 is the biggest between X_i).

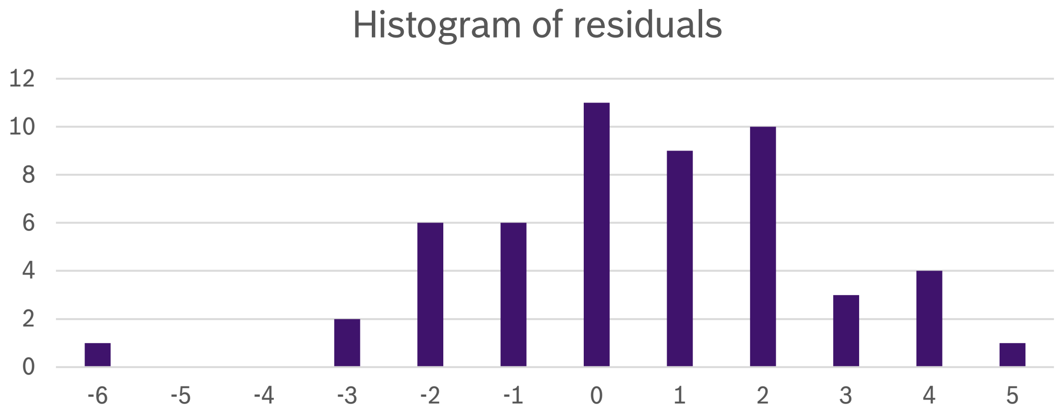
- The correlations of D(EOM) with X_i suggests that ***these three explanatory variables together will not be able to explain a lot of the variance in Y.***
  - They do combine to explain 56% (R^2), it is a significant contribution (F-test), and all three coefficients are significantly different from zero (t-test), but ***more can be done***.

Internal | CR/AEU2 - Patelis | 2018-03-14
© Robert Bosch GmbH 2018. All rights reserved, also regarding any disposal, exploitation, reproduction, editing, distribution, as well as in the event of applications for industrial property rights.

BOSCH

Forecasting Methods
Multiple Regression: Example (11)

- ***Residuals***
  - If plot shows any pattern, it indicates that the variable concerned contains some valuable predictive information and it should be included to the regression model.
  - There does not appear to be any problem with the normality assumption.
  - One residual lying away from the other values, but it not sufficient far from the others to warrant much closer attention.



© Robert Bosch GmbH 2018. All rights reserved, also regarding any disposal, exploitation, reproduction, editing, distribution, as well as in the event of applications for industrial property rights.

Forecasting Methods
Multiple Regression: Example (12)

- What more can be done?
  - Since the bank data are monthly, we can use 11 dummy variables for the ***months***.

	Y	X			D										
Months	D(EOM)	(AAA)	(3-4)	D(3-4)	D1	D2	D3	D4	D5	D6	D7	D8	D9	D10	D11
1	1,146	5,940	5,310	0,290	1	0	0	0	0	0	0	0	0	0	0
2	-2,443	6,000	5,600	-0,110	0	1	0	0	0	0	0	0	0	0	0
3	1,497	6,080	5,490	0,310	0	0	1	0	0	0	0	0	0	0	0
4	-0,132	6,170	5,800	-0,190	0	0	0	1	0	0	0	0	0	0	0
5	2,025	6,140	5,610	-0,330	0	0	0	0	1	0	0	0	0	0	0
6	0,737	6,090	5,280	-0,090	0	0	0	0	0	1	0	0	0	0	0
7	-1,023	5,870	5,190	-0,010	0	0	0	0	0	0	1	0	0	0	0
50	0,257	7,220	5,720	0,320	0	1	0	0	0	0	0	0	0	0	0
51	4,402	7,360	6,040	-0,380	0	0	1	0	0	0	0	0	0	0	0
52	3,173	7,340	5,660	0,090	0	0	0	1	0	0	0	0	0	0	0
53	5,104	7,300	5,750	0,070	0	0	0	0	1	0	0	0	0	0	0
54	4,646	7,300	5,820	0,080	0	0	0	0	0	1	0	0	0	0	0
55	1,060	7,270	5,900	0,210	0	0	0	0	0	0	1	0	0	0	0
56	-0,758	7,300	6,110	-0,060	0	0	0	0	0	0	0	1	0	0	0
57	4,702	7,310	6,050	-0,070	0	0	0	0	0	0	0	0	1	0	0
58	1,878	7,260	5,980	0,020	0	0	0	0	0	0	0	0	0	1	0
59	6,620	7,240	6,000	0,240	0	0	0	0	0	0	0	0	0	0	1

© Robert Bosch GmbH 2018. All rights reserved, also regarding any disposal, exploitation, reproduction, editing, distribution, as well as in the event of applications for industrial property rights.

Forecasting Methods
Multiple Regression: Example (13)

- If we run again a least square method (over the first 53 rows of data by using 3+11 explanatory variables), we can estimate that:

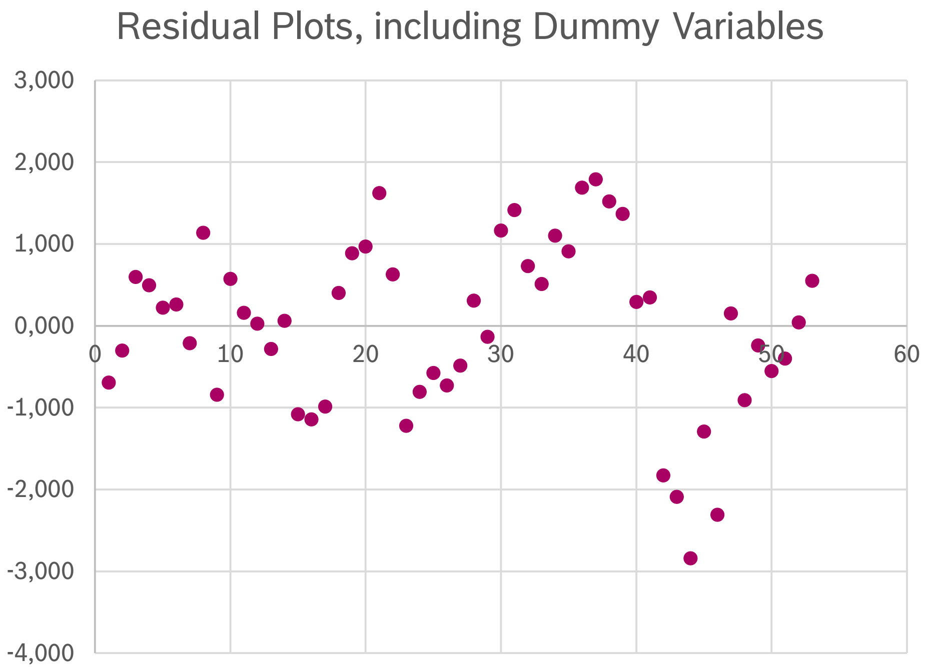
$$Y' = -2.1983 + 3.2988 * X_1 - 2.7524 * X_2 - 1.7308 * X_3$$
$$- 0.4403 * D_1 - 4.5125 * D_2 - 1.3130 * D_3 - 3.1493 * D_4 - 1.3833 * D_5 - 3.0397 * D_6$$
$$- 3.7102 * D_7 - 4.6966 * D_8 - 1.8684 * D_9 - 2.4762 * D_{10} + 1.4419 * D_{11}$$

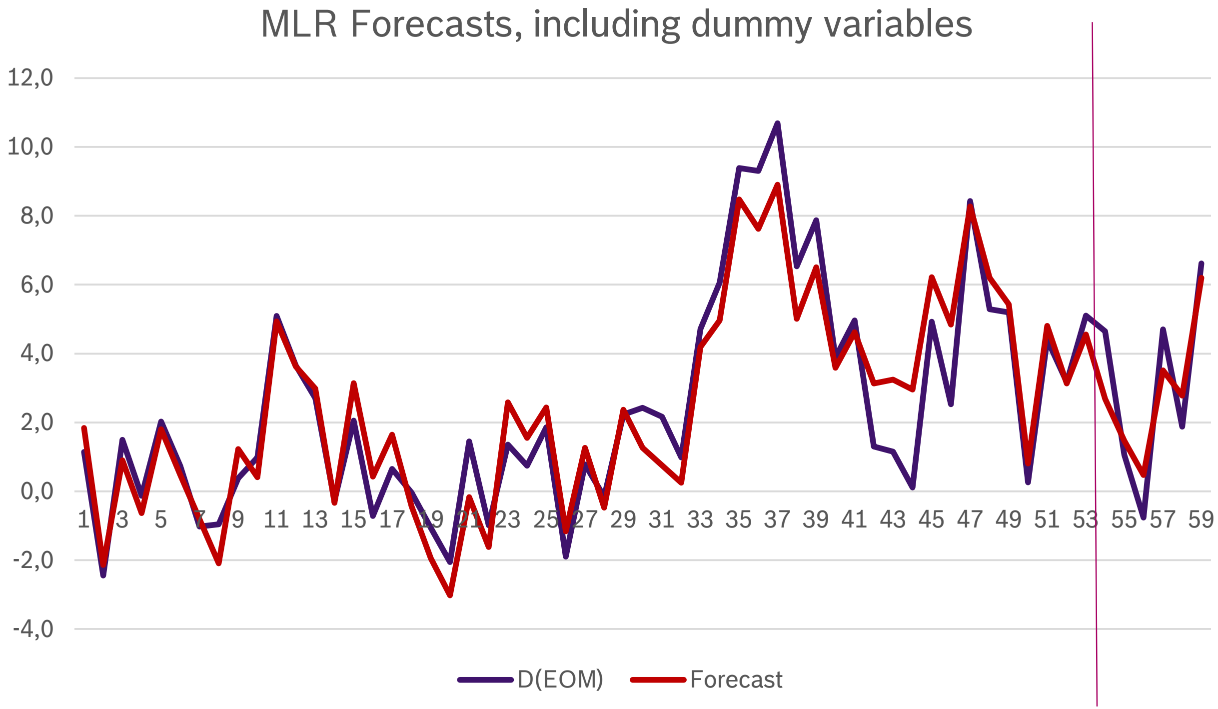
- Results:
  - ***R^2 = 0.887*** (previously 0.561)
    - The proportion of Y explained by regressing on these 14 explanatory variables is now 88.7%, instead on just 56.1%.
  - MSE = 1.49 (previously 4.48)
    - The mean squared error ***has dropped considerably***, from 4.48 to 1.49.
  - Index F = 21.27 (previously 20.85)
    - The F value is similar to the previous one, but there has been a shift in the degrees of freedom (from denominator to numerator), which makes ***the numerator more stable***.

© Robert Bosch GmbH 2018. All rights reserved, also regarding any disposal, exploitation, reproduction, editing, distribution, as well as in the event of applications for industrial property rights.

Forecasting Methods
Multiple Regression: Example (14)

- The residuals are generally smaller, and with no indication of a seasonality.
- The forecasts are significantly better.





© Robert Bosch GmbH 2018. All rights reserved, also regarding any disposal, exploitation, reproduction, editing, distribution, as well as in the event of applications for industrial property rights.

NEURAL NETWORKS

Forecasting Methods
Neural Networks (1)

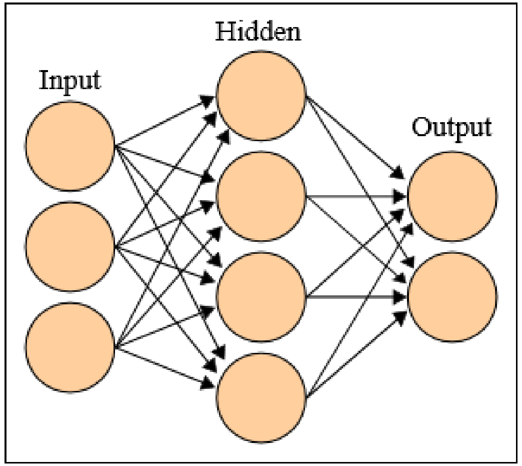
- ***Neural Network***:
  - A mathematical model inspired by biological neural networks.
  - A neural network consists of an ***interconnected group of artificial neurons***, and it processes information using a connectionist approach to computation.
  - In Artificial Intelligence are also referred to as ***Machine Learning*** (ML) or ***artificial neural networks*** (ANNs)
    - these are essentially simple mathematical models defining a function or a distribution over or both and, but sometimes models are also intimately associated with a particular learning algorithm or learning rule.
    - A common use of the phrase ANN model really ***means the definition of a class of such functions*** (where members of the class are obtained by varying parameters, connection weights, or specifics of the architecture such as the number of neurons or their connectivity).

© Robert Bosch GmbH 2018. All rights reserved, also regarding any disposal, exploitation, reproduction, editing, distribution, as well as in the event of applications for industrial property rights.

BOSCH

Forecasting Methods
Neural Networks (2)

- ***Network***: refers to the inter-connections between the neurons in the different layers of each system.
  - An example system has three layers.
    - The first layer has ***input neurons***, which send data via ***synapses*** to the second layer of neurons, and then via more synapses to the third layer of ***output neurons***.
    - More complex systems will have more layers of neurons with some having increased layers of input neurons and output neurons.
  - The synapses store parameters called "***weights***" that manipulate the data in the calculations.



Internal | CR/AEU2 - Patelis | 2018-03-14
© Robert Bosch GmbH 2018. All rights reserved, also regarding any disposal, exploitation, reproduction, editing, distribution, as well as in the event of applications for industrial property rights.

Forecasting Methods
Neural Networks (3)

- An ANN is typically defined by three types of parameters:
  - The ***interconnection*** pattern between different layers of neurons.
  - The ***learning process*** for updating the weights of the interconnections.
    - the process through which a neural network modifies itself to being able to produce a certain result with a given input.
  - The ***activation function*** that converts a neuron's weighted input to its output activation.
- Data:
  - NN is an adaptive system that changes its structure during a learning phase. Neural networks are used to model ***complex relationships between inputs and outputs*** or to find patterns in data.
  - They need a ***lot of data*** (in order to train the network) and they are usually ***time-consuming***.

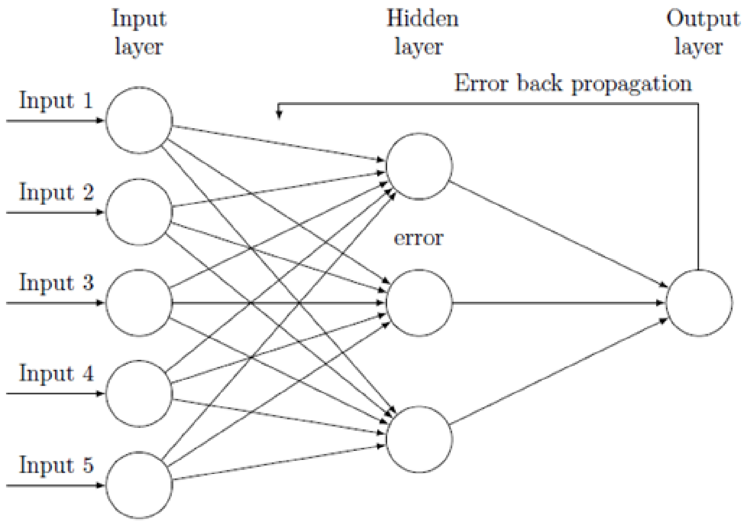
© Robert Bosch GmbH 2018. All rights reserved, also regarding any disposal, exploitation, reproduction, editing, distribution, as well as in the event of applications for industrial property rights.

Forecasting Methods
Neural Networks: Learning (1)

- ***Learning***:
  - Given a specific task to solve, and a class of functions F, learning means ***using a set of observations*** to find which ***solves the task*** in some optimal sense.
  - This entails defining a ***cost function*** such that, for the optimal solution, - i.e., no solution has a cost less than the cost of the optimal solution.
  - The ***cost function C*** is an important concept in learning, as it is a measure of how far away a particular solution is from an optimal solution to the problem to be solved.
  - Learning algorithms ***search through the solution space*** to find a function that has the smallest possible cost.

© Robert Bosch GmbH 2018. All rights reserved, also regarding any disposal, exploitation, reproduction, editing, distribution, as well as in the event of applications for industrial property rights.

Forecasting Methods
Neural Networks: Learning (2)

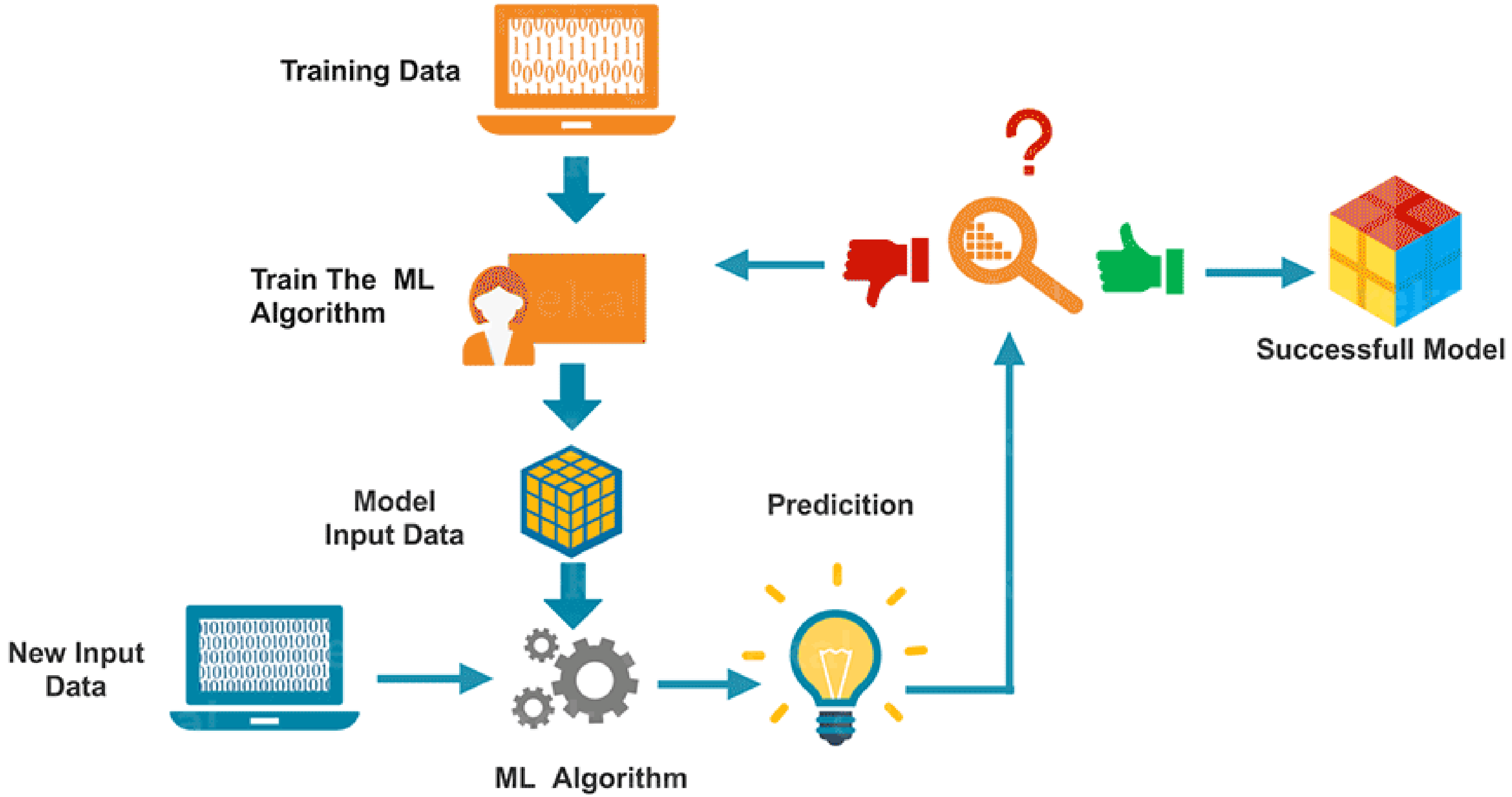


- ***Learning***:
  - To start this process, the initial weights are ***chosen randomly***. Then the learning begins.
  - The neural network processes the data in the "training set" one at a time, using the weights and functions in the hidden layers, then compares the resulting outputs against the desired outputs.
- ***Cost (Errors)*** are then ***propagated back through the system***, causing the system to adjust the weights for application to the next data.
- This process occurs repeatedly as the weights are tweaked. During the learning phase of a network, the same set of data is processed many times as the connection weights are continually refined.

Internal | CR/AEU2 - Patelis | 2018-03-14
© Robert Bosch GmbH 2018. All rights reserved, also regarding any disposal, exploitation, reproduction, editing, distribution, as well as in the event of applications for industrial property rights.

Source: Microsoft Social Technet

Forecasting Methods
Neural Networks: Learning (3)



Internal | CR/AEU2 - Patelis | 2018-03-14
© Robert Bosch GmbH 2018. All rights reserved, also regarding any disposal, exploitation, reproduction, editing, distribution, as well as in the event of applications for industrial property rights.

BOSCH

Source: Edureka: What is Machine

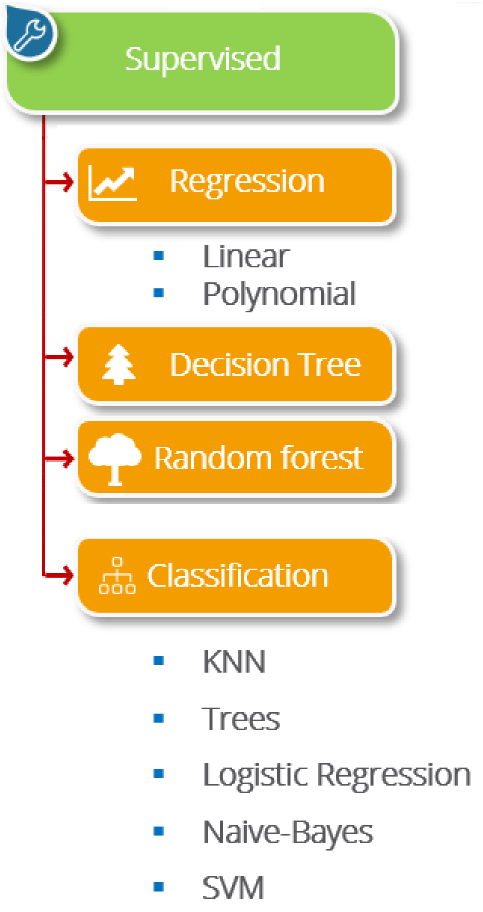
Forecasting Methods
Neural Networks: Supervised Learning (1)

- ***Supervised Learning***:
  - We have a set of example pairs (data, results) and the aim is to find a function (in the allowed class of functions) that matches the examples.
    - In other words, we wish ***to infer the mapping*** implied by the data.
    - the cost function is related to the ***mismatch between our mapping and the data*** and it implicitly contains prior knowledge about the problem domain.
  - A commonly used cost is the **mean-squared error**, which tries to minimize the average squared error between the network's output, *f(x)*, and the target value *y* over all the example pairs.
    - When one tries to minimize this cost using gradient descent for the class of neural networks called multilayer perceptrons, one obtains the common and well-known backpropagation algorithm for training neural networks.
  - The training process continues until the model achieves a ***desired level of accuracy*** on the training data.

© Robert Bosch GmbH 2018. All rights reserved, also regarding any disposal, exploitation, reproduction, editing, distribution, as well as in the event of applications for industrial property rights.
BOSCH

Forecasting Methods
Neural Networks: Supervised Learning (2)

- ***Supervised Learning***:
  - Tasks that fall within the paradigm of supervised learning are:
    - ***pattern recognition*** (also known as ***classification***) and
    - ***regression*** (also known as function ***approximation***).
  - The supervised learning paradigm is also applicable to sequential data (e.g., for speech and gesture recognition).
  - This can be thought of as learning with a "teacher," in the form of ***a function that provides continuous feedback on the quality of solutions*** obtained thus far.



Source: Edureka: What is Machine Learning

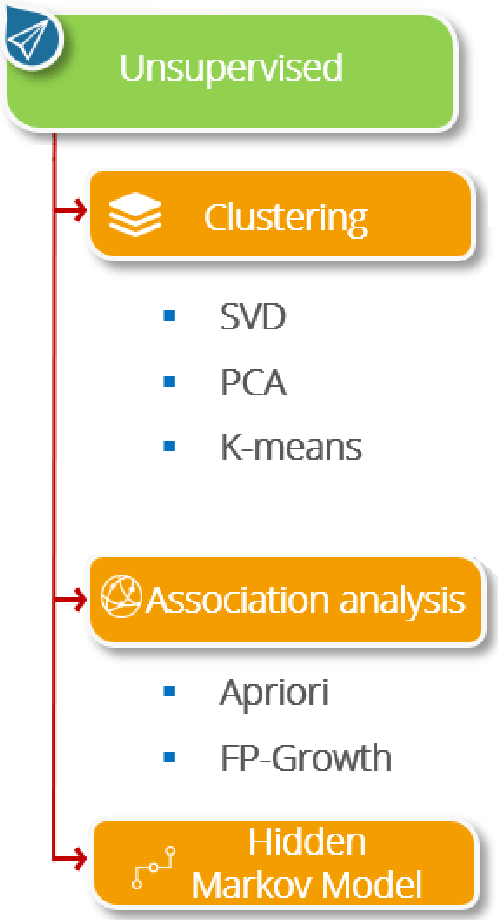
Internal | CR/AEU2 - Patelis | 2018-03-14
© Robert Bosch GmbH 2018. All rights reserved, also regarding any disposal, exploitation, reproduction, editing, distribution, as well as in the event of applications for industrial property rights.

BOSCH

Forecasting Methods

Neutral Networks: Unsupervised Learning

- ***Unsupervised Learning***:
  - Modifies the network's weights relying ***only on input data***, resulting in networks that will learn to group received informations with probabilistic methods.
  - The cost function is ***dependent on the task*** (what we are trying to model) and our ***a priori assumptions*** (the implicit properties of our model, its parameters and the observed variables).
  - Tasks that fall within the paradigm of unsupervised learning are:
    - in general ***estimation problems***
    - the applications include ***clustering***,
    - the estimation of statistical distributions,
    - ***compression*** and filtering.



Source: Edureka: What is Machine Learning

Internal | CR/AEU2 - Patelis | 2018-03-14
© Robert Bosch GmbH 2018. All rights reserved, also regarding any disposal, exploitation, reproduction, editing, distribution, as well as in the event of applications for industrial property rights.

Forecasting Methods
Neural Networks: Reinforcement Learning

- ***Reinforcement Learning***:
  - Data are usually not given, but generated by an agent's interactions with the environment.
    - It follows the concept of hit and trial method.
  - At each point in time, the agent performs an ***action*** and the environment generates an ***observation*** and an ***instantaneous cost***, according to some (usually unknown) dynamics.
  - The aim is to discover a ***policy for selecting*** actions that ***minimizes some measure*** of a long-term cost; i.e., the expected cumulative cost.
  - The environment's dynamics and the long-term cost for each policy are usually unknown, but can be estimated.

© Robert Bosch GmbH 2018. All rights reserved, also regarding any disposal, exploitation, reproduction, editing, distribution, as well as in the event of applications for industrial property rights.

BOSCH

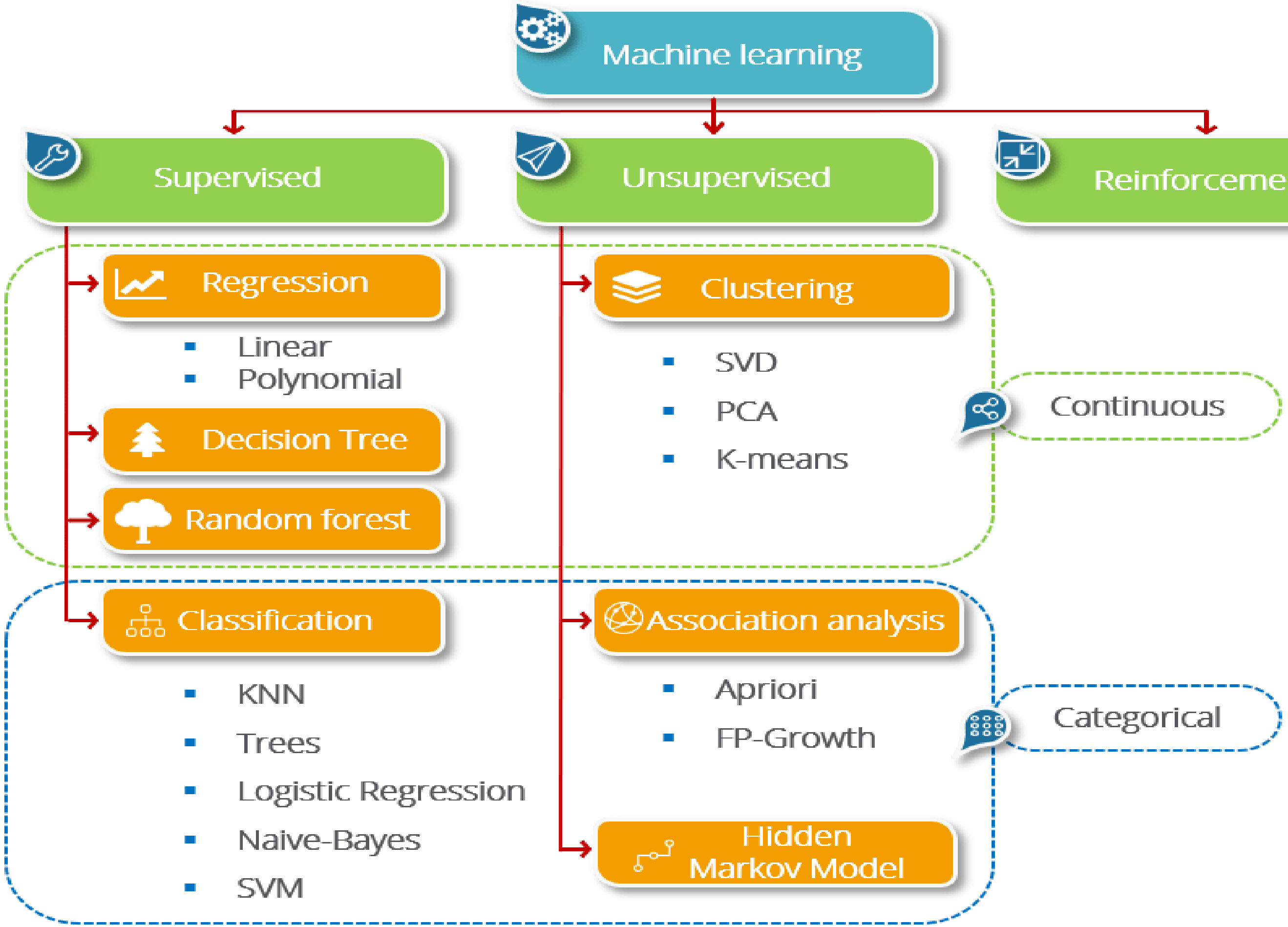
Forecasting Methods
Neural Networks: Deep Learning

- ***Deep Learning***:
  - A network with ***more than two hidden layers*** is commonly referred as "deep learning".
  - In practice it can be very effective.
  - They tend to converge to a solution ***faster*** during the fitting procedure.
- Keep in mind:
  - Any network with more than 1 hidden layer can be ***mimicked*** with only 1 hidden layer.
  - We can approximate any continuous function using a neural network with a single hidden layer, as per the Universal Approximation Theorem.

© Robert Bosch GmbH 2018. All rights reserved, also regarding any disposal, exploitation, reproduction, editing, distribution, as well as in the event of applications for industrial property rights.

BOSCH

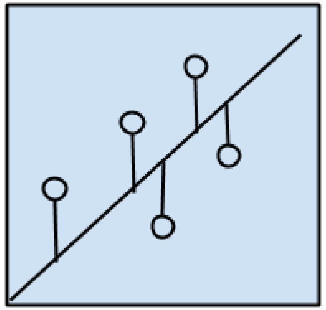
Forecasting Methods
Neural Networks: Categories

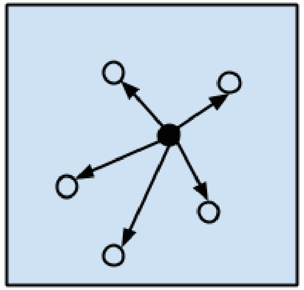


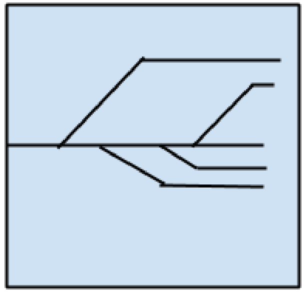
Internal | CR/AEU2 - Patelis | 2018-03-14
© Robert Bosch GmbH 2018. All rights reserved, also regarding any disposal, exploitation, reproduction, editing, distribution, as well as in the event of applications for industrial property rights.

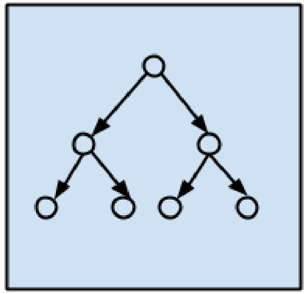
Source: Edureka: What is Machine Learning

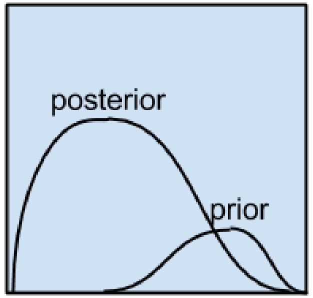
Forecasting Methods
Neural Networks: Types


Regression Algorithms

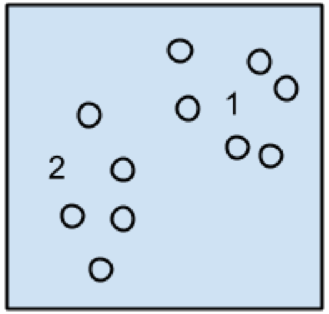

Instance-based Algorithms


Regularization Algorithms

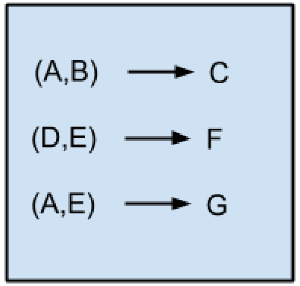

Decision Tree Algorithms



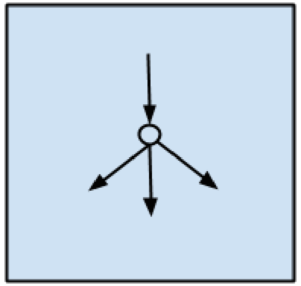
Bayesian Algorithms

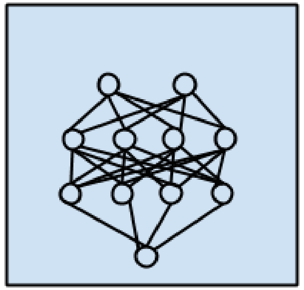


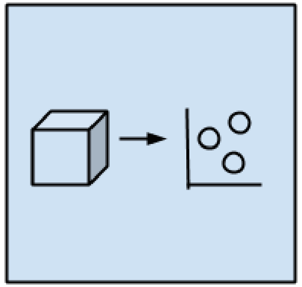
Clustering Algorithms

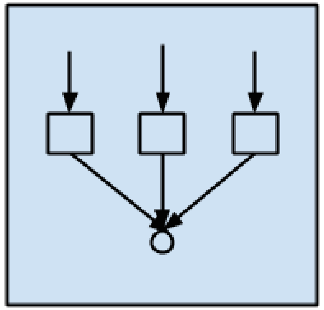


Association Rule Learning Algorithms


Artificial Neural Network Algorithms


Deep Learning Algorithms


Dimensional Reduction Algorithms


Ensemble Algorithms

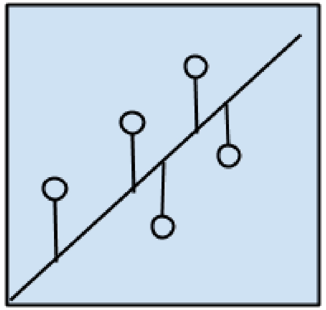
Internal | CR/AEU2 - Patelis | 2018-03-14
© Robert Bosch GmbH 2018. All rights reserved, also regarding any disposal, exploitation, reproduction, editing, distribution, as well as in the event of applications for industrial property rights.

BOSCH

Source: Machine Learning Mastery

Forecasting Methods
Neural Networks: Regression

- ***Regression***:
  - Modeling the ***relationship between variables*** that is iteratively refined using a measure of error in the predictions made by the mode.
  - Regression methods are a workhorse of statistics and have been co-opted into statistical machine learning.
    - Ordinary Least Squares Regression (OLSR)
    - Linear Regression
    - Logistic Regression
    - Stepwise Regression
    - Multivariate Adaptive Regression Splines (MARS)
    - Locally Estimated Scatterplot Smoothing (LOESS)



Regression Algorithms

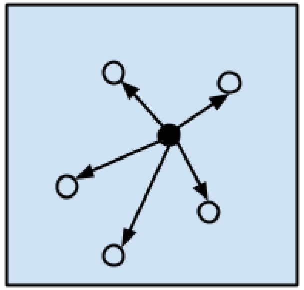
Internal | CR/AEU2 - Patelis | 2018-03-14
© Robert Bosch GmbH 2018. All rights reserved, also regarding any disposal, exploitation, reproduction, editing, distribution, as well as in the event of applications for industrial property rights.

BOSCH

Source: Machine Learning Mastery

Forecasting Methods
Neural Networks: Instance-based

- ***Instance-based***:
  - A decision problem with instances or examples of training data that are deemed important or required to the model.
  - Such methods typically build up a ***database of example data*** and compare new data to the database using a ***similarity measure*** in order to find the best match and make a prediction.
    - k-Nearest Neighbor (kNN)
    - Learning Vector Quantization (LVQ)
    - Self-Organizing Map (SOM)
    - Locally Weighted Learning (LWL)



Instance-based Algorithms

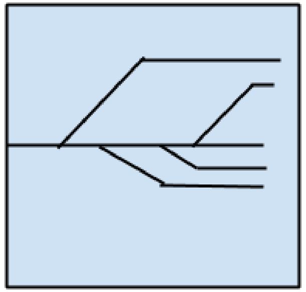
Internal | CR/AEU2 - Patelis | 2018-03-14
© Robert Bosch GmbH 2018. All rights reserved, also regarding any disposal, exploitation, reproduction, editing, distribution, as well as in the event of applications for industrial property rights.

BOSCH

Source: Machine Learning Mastery

Forecasting Methods
Neural Networks: Regulatization

- ***Regulatization***:
  - An extension made to another method (typically regression methods) that ***penalizes models based on their complexity***, favoring simpler models that are also better at generalizing.
    - Ridge Regression
    - Least Absolute Shrinkage and Selection Operator (LASSO)
    - Elastic Net
    - Least-Angle Regression (LARS)



Regularization Algorithms

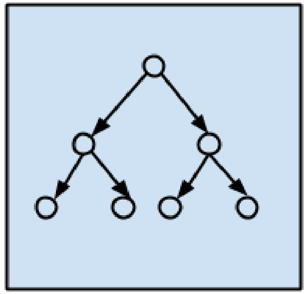
Internal | CR/AEU2 - Patelis | 2018-03-14
© Robert Bosch GmbH 2018. All rights reserved, also regarding any disposal, exploitation, reproduction, editing, distribution, as well as in the event of applications for industrial property rights.

BOSCH

Source: Machine Learning Mastery

Forecasting Methods
Neural Networks: Decision Tree

- ***Decision Tree***:
  - They construct a ***model of decisions*** made based on actual values of attributes in the data.
  - Decisions fork in tree structures until a prediction decision is made for a given record. They are trained on data for classification and regression problems. Decision trees are often fast and accurate and a big favorite in machine learning.
    - Classification and Regression Tree (CART)
    - Iterative Dichotomiser 3 (ID3)
    - C4.5 and C5.0 (different versions of a powerful approach)
    - Chi-squared Automatic Interaction Detection (CHAID)
    - Decision Stump
    - M5
    - Conditional Decision Trees



Decision Tree
Algorithms

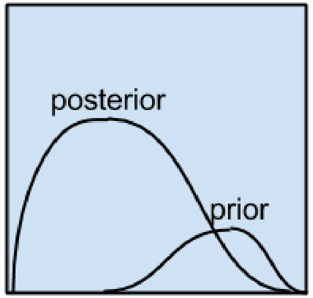
Internal | CR/AEU2 - Patelis | 2018-03-14
© Robert Bosch GmbH 2018. All rights reserved, also regarding any disposal, exploitation, reproduction, editing, distribution, as well as in the event of applications for industrial property rights.

BOSCH

Source: Machine Learning Mastery

Forecasting Methods
Neural Networks: Bayesian

- ***Bayesian***:
  - Bayesian methods are those that explicitly apply ***Bayes' Theorem*** for problems such as classification and regression.
    - Naive Bayes
    - Gaussian Naive Bayes
    - Multinomial Naive Bayes
    - Averaged One-Dependence Estimators (AODE)
    - Bayesian Belief Network (BBN)
    - Bayesian Network (BN)



Bayesian Algorithms

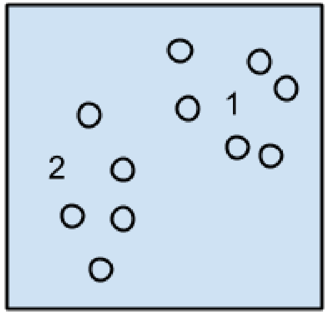
© Robert Bosch GmbH 2018. All rights reserved, also regarding any disposal, exploitation, reproduction, editing, distribution, as well as in the event of applications for industrial property rights.

BOSCH

Source: Machine Learning Mastery

Forecasting Methods
Neural Networks: Clustering

- ***Clustering***:
  - Clustering, like regression, describes the class of problem and the class of methods.
  - Clustering methods are typically organized by the modeling approaches such as centroid-based and hierarchal. All methods are concerned with using the inherent structures in the data to best ***organize the data into groups of maximum commonality***.
    - k-Means
    - k-Medians
    - Expectation Maximisation (EM)
    - Hierarchical Clustering



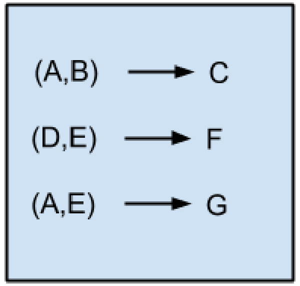
Clustering Algorithms

Internal | CR/AEU2 - Patelis | 2018-03-14
© Robert Bosch GmbH 2018. All rights reserved, also regarding any disposal, exploitation, reproduction, editing, distribution, as well as in the event of applications for industrial property rights.

Source: Machine Learning Mastery

Forecasting Methods
Neural Networks: Association Rule Learning

- ***Association Rule Learning***:
  - They extract rules that ***best explain observed relationships*** between variables in data.
  - These rules can discover important and ***commercially useful associations*** in large multidimensional datasets that can be exploited by an organization.
    - Apriori algorithm
    - Eclat algorithm



Association Rule Learning Algorithms

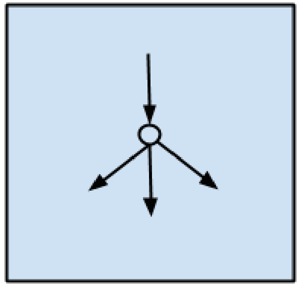
© Robert Bosch GmbH 2018. All rights reserved, also regarding any disposal, exploitation, reproduction, editing, distribution, as well as in the event of applications for industrial property rights.

BOSCH

Source: Machine Learning Mastery

Forecasting Methods
Neural Networks: Artificial NN

- ***Artificial NN***:
  - Models that are inspired by the structure and/or function of biological neural networks.
  - They are a ***class of pattern matching*** that are commonly used for ***regression*** and ***classification problems*** but are really an enormous subfield comprised of hundreds of algorithms and variations for all manner of problem types.
    - Perceptron
    - Back-Propagation
    - Hopfield Network
    - Radial Basis Function Network (RBFN)



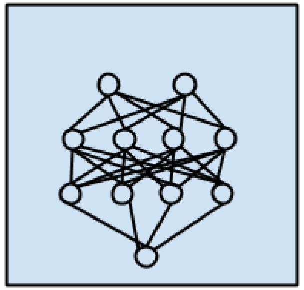
Artificial Neural Network Algorithms

© Robert Bosch GmbH 2018. All rights reserved, also regarding any disposal, exploitation, reproduction, editing, distribution, as well as in the event of applications for industrial property rights.

Source: Machine Learning Mastery

Forecasting Methods
Neural Networks: Deep Learning

- ***Deep Learning***:
  - A modern update to Artificial Neural Networks that exploit abundant cheap computation.
  - They are concerned with building much ***larger and more complex neural networks*** and, as commented on above, many methods are concerned with semi-supervised learning problems where large datasets contain very little labeled data.
    - Deep Boltzmann Machine (DBM)
    - Deep Belief Networks (DBN)
    - Convolutional Neural Network (CNN)
    - Stacked Auto-Encoders



Deep Learning Algorithms

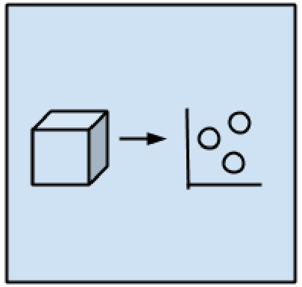
© Robert Bosch GmbH 2018. All rights reserved, also regarding any disposal, exploitation, reproduction, editing, distribution, as well as in the event of applications for industrial property rights.

BOSCH

Source: Machine Learning Mastery

Forecasting Methods
Neural Networks: Dimensionality Reduction

- ***Dimensionality Reduction***:
  - Seek and exploit the inherent structure in the data, but in this case in an unsupervised manner or order to summarize or ***describe data using less information***.
  - This can be useful to visualize dimensional data or to simplify data which can then be used in a supervised learning method.
    - Principal Component Analysis (PCA)
    - Principal Component Regression (PCR)
    - Partial Least Squares Regression (PLSR)
    - Sammon Mapping
    - Multidimensional Scaling (MDS)
    - Projection Pursuit
    - Linear Discriminant Analysis (LDA)
    - Mixture Discriminant Analysis (MDA)
    - Quadratic Discriminant Analysis (QDA)
    - Flexible Discriminant Analysis (FDA)

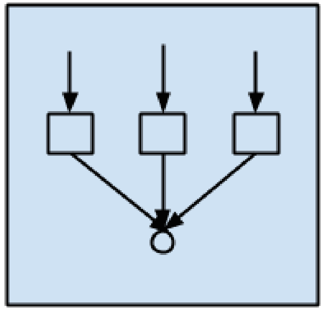


Dimensional Reduction Algorithms

Internal | CR/AEU2 - Patelis | 2018-03-14
© Robert Bosch GmbH 2018. All rights reserved, also regarding any disposal, exploitation, reproduction, editing, distribution, as well as in the event of applications for industrial property rights.
BOSCH
Source: Machine Learning Mastery

Forecasting Methods
Neural Networks: Ensemble

- ***Ensemble***:
  - Models composed of ***multiple weaker models*** that are ***independently trained*** and whose predictions are ***combined*** in some way to make the overall prediction.
  - Much effort is put into what types of weak learners to combine and the ways in which to combine them. This is a very powerful class of techniques and as such is very popular.
    - Boosting
    - Bootstrapped Aggregation (Bagging)
    - AdaBoost
    - Stacked Generalization (blending)
    - Gradient Boosting Machines (GBM)
    - Gradient Boosted Regression Trees (GBRT)
    - Random Forest



Ensemble Algorithms

© Robert Bosch GmbH 2018. All rights reserved, also regarding any disposal, exploitation, reproduction, editing, distribution, as well as in the event of applications for industrial property rights.

BOSCH

Source: Machine Learning Mastery

Forecasting Methods
Neural Networks: Employing

- ***Greatest advantage***: their ability to be used as an arbitrary function approximation mechanism that 'learns' from observed data.
  - However, using them is not so straightforward and a good understanding of the underlying theory is essential.
- ***Choice of model***: This will depend on the data representation and the application.
  - Overly complex models tend to lead to problems with learning.
- ***Learning algorithm***: There are numerous trade-offs between learning algorithms.
  - Almost any algorithm will work well with the hyper parameters for training on a particular fixed data set.
  - However selecting and tuning an algorithm for training on unseen data requires a significant amount of experimentation.
- ***Robustness***:
  - If the model, cost function and learning algorithm are selected appropriately the resulting ANN can be extremely robust.

© Robert Bosch GmbH 2018. All rights reserved, also regarding any disposal, exploitation, reproduction, editing, distribution, as well as in the event of applications for industrial property rights.
BOSCH

Forecasting Methods

Neural Networks: Theoretical Properties (1)

- ***Computational Power***:
  - The multi-layer perceptron (MLP) is a universal function approximator. However, the proof is not constructive regarding the number of neurons required or the settings of the weights.
- ***Capacity***:
  - Corresponds to their ability to model any given function. It is related to the amount of information that can be stored in the network and to the notion of complexity.
- ***Convergence***:
  - Depends on a number of factors (cost function, model, amount of data, parameters, etc.).
  - In practice, theoretical guarantees regarding convergence are an unreliable guide to practical application.

© Robert Bosch GmbH 2018. All rights reserved, also regarding any disposal, exploitation, reproduction, editing, distribution, as well as in the event of applications for industrial property rights.

BOSCH

Forecasting Methods
Neural Networks: Theoretical Properties (2)

- ***Generalization and statistics***:
  - The problem of over-training.
    - This arises in convoluted or over-specified systems when the capacity of the network significantly exceeds the needed free parameters.
  - There are two options for avoiding this problem:
    - Use ***cross-validation*** and similar techniques to check for the presence of overtraining and optimally select hyper parameters such as to minimize the generalization error.
    - ***Use some form of regularization***. This is a concept that emerges naturally in a probabilistic (Bayesian) framework, where the regularization can be performed by selecting a larger prior probability over simpler models.
    - Also in statistical learning theory, where the goal is to minimize over two quantities: the 'empirical risk' and the 'structural risk', which roughly corresponds to the error over the training set and the predicted error in unseen data due to overfitting.

© Robert Bosch GmbH 2018. All rights reserved, also regarding any disposal, exploitation, reproduction, editing, distribution, as well as in the event of applications for industrial property rights.
BOSCH

Forecasting Methods
Neural Networks & Timeseries?

- ML techniques have been ***gaining prominence*** over time as interest in AI has been rising.
- Although ML methods have been proposed in the academic literature as alternatives to statistical ones, only scant evidence is available about their performance in terms of accuracy and computational requirements.
  - ***Traditional statistical methods are considerably more accurate*** that ML ones.
  - ML methods have a long way to go to become more accurate, less time demanding and less of a black box.
  - Best fitted available data ***does not necessarily result*** in more accurate predictions.

Source: *"The accuracy of Machine Learning (ML) Forecasting methods versus Statistical ones: Extending the results of the M3-Competition", Makridakis et. al., March 2018*

© Robert Bosch GmbH 2018. All rights reserved, also regarding any disposal, exploitation, reproduction, editing, distribution, as well as in the event of applications for industrial property rights.

BOSCH

Forecasting Methods
Neural Networks & Timeseries: Train & Retrain (1)

- On NN we train a model, test it, retrain it if necessary until we have satisfactory results, and then evaluate it on a hold out data set. If we're satisfied with the performance, we then ***deploy it to production***.
  - Once in production, we score new data as it comes in. Eventually after a few months, we might want to update the model if a significant amount of new training data comes in.
  - Model training ***is a one time activity,*** or done at most at periodic intervals to maintain the model's performance.
    - Example: ***Classifying cat images***. The visual properties of cats are stable over time. Given enough data, the model we trained this week is good enough for the foreseeable future as well.

© Robert Bosch GmbH 2018. All rights reserved, also regarding any disposal, exploitation, reproduction, editing, distribution, as well as in the event of applications for industrial property rights.
BOSCH
Source: Towards Data Science

Forecasting Methods
Neural Networks & Timeseries: Train & Retrain (2)

- On Statistical Methods ***this is not the case***!
  - We have to ***retrain our model every time*** we want to generate a new forecast, or new actuals come in.
  - The development data set and the production data set are not the same (real world).
  - We are actually retraining a new model from scratch every time we want to generate a new forecast.

- ***In practice***: deploying forecasting algorithms to production is very different from deploying NN models.

© Robert Bosch GmbH 2018. All rights reserved, also regarding any disposal, exploitation, reproduction, editing, distribution, as well as in the event of applications for industrial property rights.

Source: Towards Data Science

Forecasting Methods
Neural Networks & Timeseries: Test/Train Split

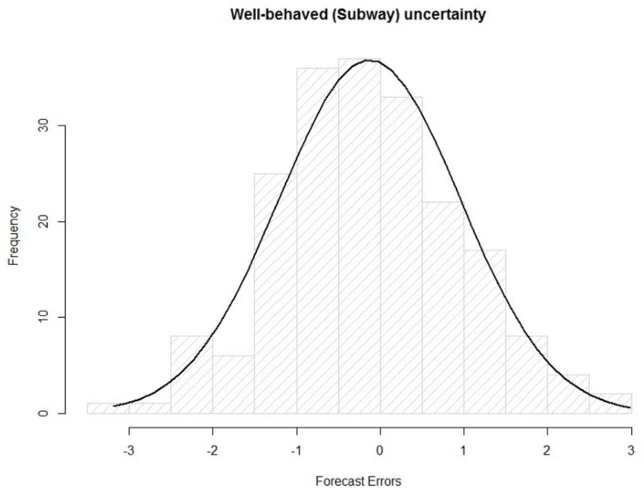
- NN require ***enough data*** to use (train & test) for building a model. But time series data is often very small compared to the data sets used in image processing or NLP.
  - Example: Two years of weekly data are 104 points, 10 years of quarterly data are 40 points.
- With data sets as small as this, we don’t have the luxury of setting aside 20% or 30% of the data for testing purposes.
- Cross validation can’t help, since it is not applicable at all.
  - It doesn’t make sense to try to forecast values for February using training data from January and March. That would amount to leakage.

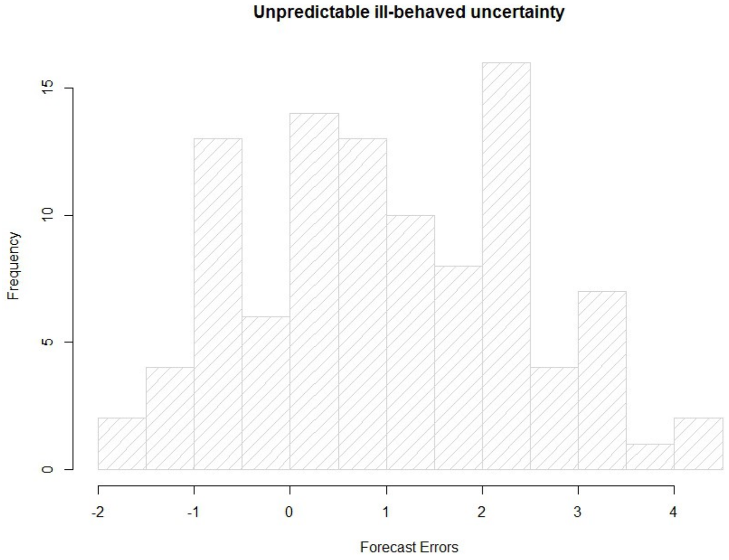
© Robert Bosch GmbH 2018. All rights reserved, also regarding any disposal, exploitation, reproduction, editing, distribution, as well as in the event of applications for industrial property rights.

Source: Towards Data Science

Forecasting Methods
Neural Networks & Timeseries: Uncertainty

- We can reasonably expect that an NN (for image classification or for NLP problem) can eventually classify all new incoming examples ***accurately*** (given enough training data).
- This is usually ***not the case in business forecasting applications***, where forecasts are almost always going to be not accurate.
  - In demand forecasting and inventory applications, ***the uncertainty of forecast is crucial for the applications*** that consume the forecast.
  - The uncertainty of your forecast (represented by forecast intervals or by forecast quantiles) is what we will use to calculate safety stock.





© Robert Bosch GmbH 2018. All rights reserved, also regarding any disposal, exploitation, reproduction, editing, distribution, as well as in the event of applications for industrial property rights.

BOSCH

Source: Towards Data Science

11. JUDGMENTAL FORECASTING METHODS

Forecasting Methods
Judgmental methods:

- Using a group of experts to predict future values.
  - A group will bring to the table a ***variety of different perspectives***, diverse expertise and a wider range of information that an individual.
  - A group also affords the opportunity to ***exchange ideas*** and challenge and ***criticise each other's arguments***.
- Can be used not only in predictive trends, but also in ***decision making*** and ***policy planning***.

© Robert Bosch GmbH 2018. All rights reserved, also regarding any disposal, exploitation, reproduction, editing, distribution, as well as in the event of applications for industrial property rights.

Forecasting Methods
Judgmental methods: When to use & how?

- There are three general settings where judgmental forecasting is used:
  - when there are ***no available data*** so that statistical methods are not applicable and judgmental forecasting is the only feasible approach,
  - when data are available, statistical forecasts are generated and these are then ***adjusted using judgement***,
  - when data are available and statistical and judgmental forecasts are ***independently generated*** and then ***combined***.
- They need a ***systematic and well-structured approach*** in order to reduce limitations impact.
  - A set of 5 basic principles.

© Robert Bosch GmbH 2018. All rights reserved, also regarding any disposal, exploitation, reproduction, editing, distribution, as well as in the event of applications for industrial property rights.
BOSCH

Forecasting Methods
Judgmental methods: Limitations (1)

- Judgmental forecasts can be ***inconsistent***.
  - They depend heavily on human cognition and are vulnerable to its limitations.
    - For example, a ***limited memory*** may render recent events more important than they actually are and may ignore momentous events from the more distant past; or
    - a ***limited attention*** span may result in important information being missed; or
    - a ***misunderstanding of causal relationships*** may lead to erroneous inference.
    - Furthermore, human judgement can vary due to the effect of ***psychological*** factors.
      - A manager, who is in a positive frame of mind one day, can generate forecasts that may tend to be somewhat optimistic, and in a negative frame of mind another day, can generate somewhat less optimistic forecasts.

© Robert Bosch GmbH 2018. All rights reserved, also regarding any disposal, exploitation, reproduction, editing, distribution, as well as in the event of applications for industrial property rights.

Forecasting Methods
Judgmental methods: Limitations (2)

- ***Misunderstanding of causal relationships***:
- Suppose a widely reported study finds that people who drink expensive tee tend to live longer.
  - In reality, this simply reflects the fact that only wealthier people, who on average live longer than poor people anyway, can afford higher-priced tea. But the report doesn't say this.
  - After hearing about the report Mary, a person struggling to make ends meet, decides to sacrifice some of her other food purchases to scrape together enough cash to buy a premium brand of tea regularly.
  - Of course, the prediction that this will lead to a longer life doesn't apply to her. Mary's conditions are different.

  - ***Worse still:*** the switch from nutritious food to expensive tea might even shorten her lifespan.

© Robert Bosch GmbH 2018. All rights reserved, also regarding any disposal, exploitation, reproduction, editing, distribution, as well as in the event of applications for industrial property rights.

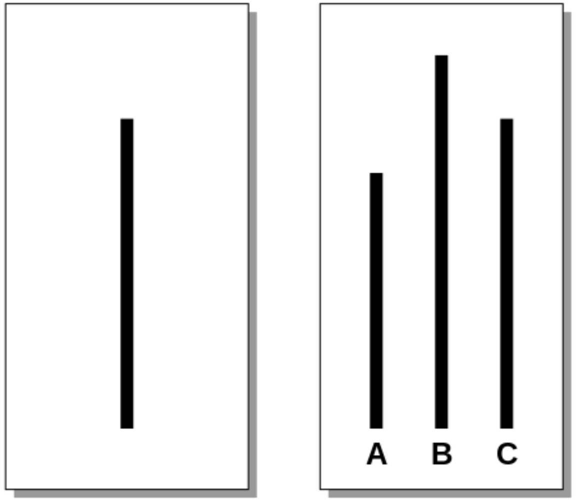
Forecasting Methods
Judgmental methods: Limitations (3)

- ***Misunderstanding of causal relationships***:
- “Nature” scientific journal, 1999: A paper was published showing that children under the age of two were more likely to be short-sighted if they slept with the lights on at night.
  - The best decision seems obvious – switch of the lights in your child’s bedroom.
  - The findings were rejected by other researchers. It was found that ***parents with myopia*** had an increased tendency to employ lights as a night-time aid for their children.
  - The children probably inherited their short-sightness from their parents, not from the lights.

  - **Gain:** Deciding to switch lights off would have at least reduced the electricity bill.

© Robert Bosch GmbH 2018. All rights reserved, also regarding any disposal, exploitation, reproduction, editing, distribution, as well as in the event of applications for industrial property rights.

Forecasting Methods
Judgmental methods: Limitations (4)



- An experiment on “***visual perception***”…
  - 6 people + you are participating.
  - Each person receives 2 cards by the instructor.
  - Everyone has the ***same pair of cards***.
  - So: Which of the lines A, B or C is the same length as the line on the other card?
    - Easy, right? The answer is obviously C.

- First person answers: a confident “A”!
- They all answer: “A”!
- Now it’s your turn. What you will say?
  - Can all the others be wrong?
  - Should you answer “C” and risk making a fool of yourself?
  - Should you ignore your own eyes?

➢ ***32 % agreed that “A” was the correct answer!***

© Robert Bosch GmbH 2018. All rights reserved, also regarding any disposal, exploitation, reproduction, editing, distribution, as well as in the event of applications for industrial property rights.

BOSCH

Source: Asch conformity experiments

Forecasting Methods
Judgmental methods: Limitations (4)

- ***Pressure to conform***:
- Members of the group often feel a pressure to conform to the view of the rest of the group.
  - Even if this view is manifestly wrong!
- Experiments have shown that ***75% of people conformed on at least one occasion*** (out of 12 trials).
  - They did not want to appear “peculiar” in front of the others.
  - They did not want to be ridiculed.

© Robert Bosch GmbH 2018. All rights reserved, also regarding any disposal, exploitation, reproduction, editing, distribution, as well as in the event of applications for industrial property rights.
BOSCH

Forecasting Methods
Judgmental methods: Limitations (5)

- Judgement can be ***clouded by personal or political agendas***, where targets and forecasts are not segregated.
  - if a sales manager knows that the forecasts he generates will be used to ***set the expectation of sales*** (target), he may have the tendency to ***set these low***.
    - in order to show a good performance, thus to exceed the expected targets.
  - Even in cases where targets and forecasts are well segregated***, judgement may be affected by optimism or wishful thinking.***
    - it would be highly unlikely that a team working towards launching a new product would forecast its failure.

© Robert Bosch GmbH 2018. All rights reserved, also regarding any disposal, exploitation, reproduction, editing, distribution, as well as in the event of applications for industrial property rights.
BOSCH

Forecasting Methods
Judgmental methods: Limitations (6)

- The effect of ***anchoring***.
  - Once we have a number in our head it becomes what is known as an anchor.
    - Estimate the percentage of African Countries are members of the United Nations. Is it more that 20?
- In this case, subsequent forecasts tend to converge or be very close to an initial familiar reference point.
  - it is common to take the ***last observed value as a reference point***. The forecaster is unduly influenced by prior information and therefore gives this more weight in the forecasting process.
  - Anchoring may lead to ***conservatism*** and undervaluing new and more current information and thereby create a systematic bias.

© Robert Bosch GmbH 2018. All rights reserved, also regarding any disposal, exploitation, reproduction, editing, distribution, as well as in the event of applications for industrial property rights.

Special Topics
Judgmental methods: Principles (1)

- ***Set the forecasting task clearly and concisely***
  - Attention is needed when setting the challenges and expressing the forecasting tasks.
  - It is important that everyone is ***clear about what the task*** is.
    - All definitions should be clear and comprehensive, avoiding ambiguous and vague expressions.
  - Also it is important to avoid incorporating emotive terms and ***irrelevant information*** that may distract the forecaster.
  - Sometimes, it is useful to conduct a ***preliminary round*** of information gathering before setting the forecasting task.

© Robert Bosch GmbH 2018. All rights reserved, also regarding any disposal, exploitation, reproduction, editing, distribution, as well as in the event of applications for industrial property rights.

Special Topics
Judgmental methods: Principles (2)

- ***Implement a systematic approach***
  - Forecast accuracy and consistency can be improved by using a systematic approach to judgmental forecasting, involving ***checklists*** of categories of information relevant to the forecasting task.
  - It is helpful to identify what information is important and how this information is to be weighted.
    - when forecasting the ***demand of a new product***, what factors should we account for and how should we account for them?
      - Should it be the price,
      - the quality and quantity of the competition,
      - the economic environment at the time,
      - the target group population of the product?
  - It is worth while devoting significant effort and resources in putting together ***decision rules*** that lead to the best possible systematic approach.

© Robert Bosch GmbH 2018. All rights reserved, also regarding any disposal, exploitation, reproduction, editing, distribution, as well as in the event of applications for industrial property rights.

Special Topics
Judgmental methods: Principles (3)

- ***Document and justify***
  - Documenting the decision rules and assumptions implemented in the systematic approach can ***promote consistency*** as the same rules can be implemented repeatedly.
    - requesting a forecaster to document and justify forecasts leads to ***accountability*** which can lead to a ***reduced bias***.
    - formal documentation significantly aids in the systematic evaluation process that is suggested next.
- ***Systematically evaluate forecasts***
  - Systematically monitoring the forecasting process can identify unforeseen irregularities.
    - we must keep records of forecasts and use them to obtain feedback as the forecasted period becomes observed.
    - Although we can do our best as forecasters, the environment is dynamic. Changes occur and we need to monitor these in order to evaluate the decision rules and assumptions.
    - Feedback and evaluation helps forecasters learn and ***improve forecast accuracy***.

© Robert Bosch GmbH 2018. All rights reserved, also regarding any disposal, exploitation, reproduction, editing, distribution, as well as in the event of applications for industrial property rights.
BOSCH

Special Topics
Judgmental methods: Principles (4)

- ***Segregate forecasters and users***
  - Forecast accuracy may be impeded if the forecasting task is carried out by users of the forecasts, such as those responsible for implementing plans of action about which the forecast is concerned.
  - We should clarify that forecasting is about predicting the future as accurately as possible, given all the information available including historical data and knowledge of any future events that might impact the forecasts. Forecasters and users should be ***clearly segregated***.
    - in a case of a new product being launched, the forecast should be a reasonable estimate of the sales volume of a new product. This may be very different to what management expects or hopes the sales will be in order to meet company financial objectives. A forecaster in this case may be delivering a reality check to the user.
  - It is important that forecasters thoroughly ***communicate*** forecasts to potential users.
    - Users may feel distant and disconnected from forecasts and may not have full confidence in them. Explaining and clarifying the process and justifying basic assumptions that led to forecasts will provide some assurance to users.

© Robert Bosch GmbH 2018. All rights reserved, also regarding any disposal, exploitation, reproduction, editing, distribution, as well as in the event of applications for industrial property rights.
BOSCH

Special Topics
Judgmental methods

- ***Conclusions***:
  - Judgmental forecasts are usually ***complementary*** to statistical forecasts. People can take into consideration past events, especially those which are not under a timeseries model, but they are inconsistent and they posses increased bias. On the other hand, statistical methods are rigorous but consistent, and can cope with large data volumes quickly.
  - The effectiveness of combining independent judgmental and statistical forecasts has been examined in several studies. The overall conclusion is that the ***combination improves the forecasts accuracy***, since the components (statistical forecast, judgmental forecast) have the ability to include different aspects of the available information.

© Robert Bosch GmbH 2018. All rights reserved, also regarding any disposal, exploitation, reproduction, editing, distribution, as well as in the event of applications for industrial property rights.

SIMPLE JUDGMENT

Special Topics
Simple (unaided) judgment (1)

- The simplest type of judgmental methods.
- Predictions are being made in an “isolated” way, without the use of a structural methodology and without providing guidance, instructions or other help.
- This method is often been used as a comparison (***benchmark***) with other more sophisticated judgmental methods.
- Commonly used for estimating forecasts in cases ***where contradicting opinions exist***.
- Is the most common approach for predicting a specific point, but that requires the analysis of various factors and conditions in different areas.

© Robert Bosch GmbH 2018. All rights reserved, also regarding any disposal, exploitation, reproduction, editing, distribution, as well as in the event of applications for industrial property rights.
BOSCH

Special Topics
Simple (unaided) judgment (2)

- It has been showed, that the experience often leads the experts who use simple judgment, ***to ignore basic details and to avoid potential aids.***
  - This is leading to a decrease of the accuracy of forecasts.
- It has also been observed that the predictions made by simple judgment tend to be ***over optimistic***.
- Moreover, many times the predictions made by experienced experts are not significant better than the forecasts made by novice or non experts.

© Robert Bosch GmbH 2018. All rights reserved, also regarding any disposal, exploitation, reproduction, editing, distribution, as well as in the event of applications for industrial property rights.
BOSCH

DELPHI METHOD

Special Topics
Delphi method

- Invented by Rand Corporation (1950s), for the purpose of addressing a specific military problem.
  - Based on the principle that ***forecasts from a structured group of individuals are generally more accurate*** that those from unstructured groups or individuals.
    - But: There is often the case of experts and specialists forming a committee in order to discuss for a specific topic and to estimate forecasts. During their meetings it is possible to ***observe strong opposing views***, which will eventually affect the outcome while some members' ideas and views will be devalued.
- Aims to ***construct*** consensus ***forecasts*** from a group of experts in a ***structured iterative manner*** by exploit the advantages of a committee, while eliminating the previously discussed disadvantages.
  - Is a ***structured*** communication technique, developed as a systematic and interactive forecasting method.
  - Main goal is to provide a way of ***exporting a single view*** from a group of experts.
  - A ***facilitator*** is appointed in order to implement and manage the process.

© Robert Bosch GmbH 2018. All rights reserved, also regarding any disposal, exploitation, reproduction, editing, distribution, as well as in the event of applications for industrial property rights.

BOSCH

Special Topics
Delphi method: Implementation Stages (1)

- ***Stage 1: Panel of experts is assembled***
  - identify a group of experts that can contribute to the forecasting task.
  - Suggestion: between 5 and 20 experts with diverse expertise (no physical presence needed).
  - All experts are given an equal say and all are made accountable for their forecasts.
  - Duties of experts: Submit forecasts and provide detailed qualitative justifications for these.
  - ***Anonymity of the participants:*** Usually all participants remain anonymous.
    - Their identity is not revealed, even after the completion of the final report.
      - Minimizes the "*bandwagon effect*" (the probability of any individual adopting a belief is increasing with the proportion that have already done so).
      - Minimizes the "*halo effect*" (an observer's overall impression of a person influences the observer's feelings and thoughts about that person's character).
      - allows free expression of opinions, encourages open critique, and facilitates admission of errors.

© Robert Bosch GmbH 2018. All rights reserved, also regarding any disposal, exploitation, reproduction, editing, distribution, as well as in the event of applications for industrial property rights.
BOSCH

Special Topics
Delphi method: Implementation Stages (2)

- ***Stage 2: Forecasting tasks / challenges are set and distributed***
  - Conduct a preliminary round of information gathering from the experts before setting the forecasting tasks.
  - Alternatively, as experts submit their initial forecasts and justifications, valuable information not shared between all experts can be identified by the facilitator when compiling the feedback.
- ***Stage 3: Collect initial forecasts and justifications***
  - Collect experts forecasts and justifications.
    - The initial contributions from the experts are collected in the form of answers to questionnaires and their comments to these answers.
  - These are compiled and summarised in order to provide feedback.

© Robert Bosch GmbH 2018. All rights reserved, also regarding any disposal, exploitation, reproduction, editing, distribution, as well as in the event of applications for industrial property rights.

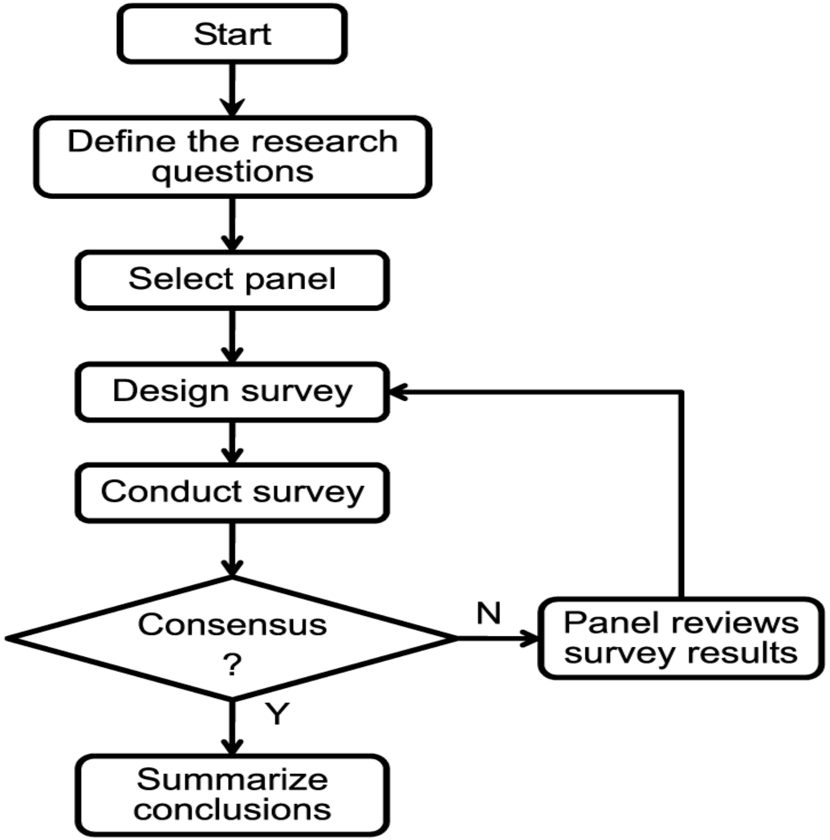
Special Topics
Delphi method: Implementation Stages (3)

- ***Stage 4: Feedback is provided to the experts (an iteration process)***
  - Feedback is provided to the experts who now review their forecasts in light of the feedback.
  - This step may be iterated (2 or more rounds) until a ***satisfactory level of consensus*** is reached.
    - After each round, a facilitator provides an anonymous summary of the experts' forecasts from the previous round, as well as the reasons they provided for their judgments.
    - Thus, experts are encouraged to revise their earlier answers in light of other members of their panel.
    - It is believed that during this process, the range of answers will decrease and the group will converge towards a "correct" single answer.
    - Finally, the process is stopped after a pre-defined criterion (e.g. number of rounds, achievement of consensus, or stability of results) and the mean or median score of the final round determine the results.

© Robert Bosch GmbH 2018. All rights reserved, also regarding any disposal, exploitation, reproduction, editing, distribution, as well as in the event of applications for industrial property rights.
BOSCH

Special Topics
Delphi method: Implementation Stages (4)

- ***Stage 5: Final forecasts***
  - The final forecasts are usually constructed by giving equal weight to all experts' forecasts.
  - The facilitator should keep in mind the ***possibility of extreme values*** which can distort the final forecast.



Internal | CR/AEU2 - Patelis | 2018-03-14
© Robert Bosch GmbH 2018. All rights reserved, also regarding any disposal, exploitation, reproduction, editing, distribution, as well as in the event of applications for industrial property rights.
BOSCH

Special Topics
Delphi method: Facilitator

- ***The role of the facilitator***
  - Selects the group of experts.
  - Send out questionnaires, surveys, etc.
  - Collect and analyse responses. Common and conflicting viewpoints are identified.
- The facilitator needs to be ***experienced*** enough to recognise areas which may need more attention than others, and to direct the attention of the experts to these.
- Also, as there is no face-to-face interaction between the experts, the facilitator is responsible for ***disseminating important information***.
- The efficiency and effectiveness of the facilitator can dramatically improve the probability of a ***successful*** Delphi method in a judgmental forecasting setting.

© Robert Bosch GmbH 2018. All rights reserved, also regarding any disposal, exploitation, reproduction, editing, distribution, as well as in the event of applications for industrial property rights.
BOSCH

ANALOGY & STRUCTURED ANALOGY

Special Topics
Analogy & Structured Analogy (1)

- The method of Analogy refers to the ***recall of similar past events and situations***, in order to provide an interpretation and a prediction to the current situations and conditions.
  - Experts are trying to recall from their memory cases which are similar with the problem under consideration.
  - Often, it is requested from the experts to use simple judgment in order to estimate forecasts, but they ***spontaneously looking for analogies*** in order to support their judgment.
    - Example: if a company wants to make a promotion for a product through an offer, it tries to link the impact of this promotion with other promotions with similar budget, in the same or in comparable products.
    - Example: An appraiser estimates the market value of a house by comparing it to similar properties that have sold in the area. The degree of similarity depends on the attributes considered. With house appraisals, attributes such as land size, dwelling size, number of bedrooms, number of bathrooms, and garage space are usually considered.

© Robert Bosch GmbH 2018. All rights reserved, also regarding any disposal, exploitation, reproduction, editing, distribution, as well as in the event of applications for industrial property rights.

Special Topics
Analogy & Structured Analogy (2)

- A process of estimating judgmental forecasts, which after the definition of the current problem, it involves 3 stages (*Lee et al., 2017)*:
  - The ***recall*** of past analogies
  - The judgmental ***forecasts*** which arise from the analogies, based on their similarity with the current problem.
  - The adjustments (***fine tuning***) of the judgmental forecasts due to the specific conditions and parameters of the current problem.
- Challenge:
  - We should aspire to base forecasts on ***multiple analogies*** rather than a single analogy, which may create biases. However these may be challenging to identify.
  - Identifying or even comparing ***multiple attributes*** may not always be straight forward.

© Robert Bosch GmbH 2018. All rights reserved, also regarding any disposal, exploitation, reproduction, editing, distribution, as well as in the event of applications for industrial property rights.

Special Topics
Analogy & Structured Analogy (3)

- Despite their weaknesses, ***Analogy is useful in complex cases*** and can improve the simple judgment forecasts.
  - However, Analogy is likely to be in favour of a desired outcome. For that reason, the recall of past Analogies should be done in an automated way, ***in order to avoid bias***.
  - The use of Analogies in a structured way (Structured Analogies) will lead to maximum efficiency.
  - The method of Structured Analogies (*created by Green and Armstrong, 2005*) originally applied to produce estimates in cases where ***a contradiction was present***. The method involves 4 steps:
    - Description of the current situation
    - Identification and describe Analogies
    - Evaluation of similarity
    - Estimation of forecasts

© Robert Bosch GmbH 2018. All rights reserved, also regarding any disposal, exploitation, reproduction, editing, distribution, as well as in the event of applications for industrial property rights.

Special Topics
Analogy & Structured Analogy: Implementation steps (1)

- Step 1: ***Describe the target situation***.
  - The administrator prepares an accurate, comprehensive, and brief description. The administrator should:
    - seek advice either from ***unbiased experts*** or from experts ***with opposing biases***.
    - When feasible, include a list of possible outcomes for the target situation to make coding easier.
- Step 2: ***Select experts***.
  - The administrator recruits experts who are likely to know about situations that are similar to the target situation.
  - The administrator should decide how many experts to recruit based on:
    - how much knowledge they have about analogous situations,
    - the variability in responses among experts, and
    - the importance of obtaining accurate forecasts.
  - In practice: At least 5 experts.

© Robert Bosch GmbH 2018. All rights reserved, also regarding any disposal, exploitation, reproduction, editing, distribution, as well as in the event of applications for industrial property rights.

BOSCH

Special Topics
Analogy & Structured Analogy: Implementation steps (2)

- Step 3: ***Identify and describe analogies***.
  - Participants-specialists, who have chosen to estimate judgmental forecasts, are asked to recall and record similar and analogical cases, as many as possible.
- Step 4: ***Rate similarity***.
  - They must list all similarities and differences, in order to rank the recalled cases, by taking into account the proximity and the affinity with the current case.
  - The goal is to provide one or more attributes and results for the current situation, by linking it with past situations with a degree of similarity.
  - A scale against which the experts can rate the similarity is suggested.

© Robert Bosch GmbH 2018. All rights reserved, also regarding any disposal, exploitation, reproduction, editing, distribution, as well as in the event of applications for industrial property rights.

Special Topics
Analogy & Structured Analogy: Implementation steps (3)

- Step 5: ***Derive forecasts***.
  - The administrator should decide on the rules to derive a forecast from experts' analogies
    - is the average of the individual forecasts of each specific participant, or
    - Analogy that the expert rated as most similar to the target and adopt the outcome implied by that analogy as the forecast.

- Ideally, the structured analogy forecast should be performed by specialists with the appropriate experience in the specific business sector, who will be able to think and recall as much relevant past cases as possible.

© Robert Bosch GmbH 2018. All rights reserved, also regarding any disposal, exploitation, reproduction, editing, distribution, as well as in the event of applications for industrial property rights.
BOSCH

Special Topics
Analogy & Structured Analogy (5)

- Support of an Information System:
  - Limited usage (in terms of specialists) of the Analogy method, due to:
    - The limitations of human memory during the recalling similar cases procedure, and
    - The limitations which lies in accessing appropriate information on past special cases of a company.
  - These limitations can be lifted by using an information system which is connected to an appropriate database that stores all past cases relevant data and all related information, such as budget and time information.
  - Studies have shown that when experts have access to such a system, the estimated judgmental forecasts are much more accurate.

© Robert Bosch GmbH 2018. All rights reserved, also regarding any disposal, exploitation, reproduction, editing, distribution, as well as in the event of applications for industrial property rights.

SCENARIO FORECASTING

Special Topics
Scenario Forecasting

- Estimate forecasts based on ***plausible scenarios***.
  - The scenarios are generated by considering all possible factors, their impact, their interaction and the forecasted targets.
  - The method can lead to already early contingency planning
    - Example: Worst, Middle and Best scenario

➢ In contrast to the two previous approaches (Delphi, Analogy) where the resulting forecast is intended to be a likely outcome, here each scenario-based forecast may have a ***low probability of occurrence***.

© Robert Bosch GmbH 2018. All rights reserved, also regarding any disposal, exploitation, reproduction, editing, distribution, as well as in the event of applications for industrial property rights.

Special Topics
Scenario Forecasting: Long-Term

- Forecasting with a horizon of 30-40 years is of a great importance for the design and the strategy within a business.
- But is a very difficult process:
  - The distance future is ***not a simple extrapolation of the past***, mainly due to technological and other changes.
  - It can be affected from many parameters, such as:
    - Economical trends
    - Prices
    - Personal income
    - Cyclical changes
    - Global population

© Robert Bosch GmbH 2018. All rights reserved, also regarding any disposal, exploitation, reproduction, editing, distribution, as well as in the event of applications for industrial property rights.
BOSCH

Special Topics
Scenario Forecasting

- Scenarios are a useful tool for long-term forecasting.
- They are attempts to describe in some detail a ***hypothetical sequence of events*** that could lead plausibly to the ***situation envisaged***.
- They can help for describing future society, given that the existing trends will continue to exist.
  - As a basis for scenario forecasting is to define a simple normal scenario with no surprises.
  - This main scenario will be partially correct, but it will not relate to a specific time but rather to a given long period.
  - The main scenario can be expanded with the development of many extreme scenarios, like the downfall of democracy or capitalism system, the possibility of a third world war and the effects of a nuclear destruction.

© Robert Bosch GmbH 2018. All rights reserved, also regarding any disposal, exploitation, reproduction, editing, distribution, as well as in the event of applications for industrial property rights.
BOSCH

Special Topics
Scenario forecasting: Features (1)

- ***Features***:

1. They serve to ***call attention***, sometimes dramatically and persuasively, to the larger range of possibilities that must be considered in the analysis of the future.
2. They dramatize and illustrate the ***possibilities*** they focus on in a very useful way.
3. They force the analyst to ***deal with details and dynamics*** that he might easily avoid treating if he restricted himself to abstract considerations.
4. They help to ***illuminate*** the interaction of psychological, social, economic, cultural, political, and military factors, including the influence of individual political personalities upon what otherwise might be abstract considerations.

© Robert Bosch GmbH 2018. All rights reserved, also regarding any disposal, exploitation, reproduction, editing, distribution, as well as in the event of applications for industrial property rights.

BOSCH

Special Topics
Scenario forecasting: Features (2)

- ***Features***:

5. They can illustrate forcefully, sometimes in oversimplified fashion, ***certain principles, issues, or questions*** that might be ignored or lost if one insisted on taking examples only from the complex and controversial real world.
6. They may also be used to consider alternative possible outcomes of certain real past and present events, such as Suez, Lebanon, or Laos.
7. They can be used as artificial “case histories” and “historical anecdotes” to make up to some degree for the paucity of actual examples.

© Robert Bosch GmbH 2018. All rights reserved, also regarding any disposal, exploitation, reproduction, editing, distribution, as well as in the event of applications for industrial property rights.

Special Topics
Scenario forecasting: Criticism

- ***Criticism***:
  - only a "paranoid" personality, suspicious, and preoccupied with hostility, could conceive of the kind of crises, provocations, aggressions, and plots that characterize many politico-military scenarios.
    - this seems to have more to do with the kinds of politico-military events the real world provides.
    - any particular scenario may in fact contain paranoid ideas, but this must be judged on the basis of the plausibility of the particular scenario (often a difficult judgment in a world of many surprises).
  - scenarios may be so ***divorced from reality*** as not only to be useless but also misleading and dangerous.
    - we are dealing with the unknown and to some degree unknowable future.
    - ***unrealistic scenarios*** are often ***useful aids to discussion***, if only to point out that the particular possibilities are unrealistic.

© Robert Bosch GmbH 2018. All rights reserved, also regarding any disposal, exploitation, reproduction, editing, distribution, as well as in the event of applications for industrial property rights.
BOSCH

Special Topics
Scenario forecasting

- ***Keep in mind!***
  - History is likely to write scenarios that most observers would find implausible not only prospectively but sometimes, even, in retrospect.
  - Many sequences of events seem plausible now only because they have actually occurred;
    - a man who knew no history might not believe any.
  - Future events may not be drawn from the restricted list of those we have learned are possible;
    - we should expect to go on being surprised.

➢ *"The usage of scenarios for forecasting future situations and events, as a challenge of the conventional and "square" thinking"*, (Makridakis et. al. 1998.)

© Robert Bosch GmbH 2018. All rights reserved, also regarding any disposal, exploitation, reproduction, editing, distribution, as well as in the event of applications for industrial property rights.

BOSCH

Special Topics
Scenario forecasting: Example

- ***Kahn & Wiener, 1967:*** a list with 100 technological innovations that they were likely to occur in the last 20 years of the 20th century.

- ***Successfully*** forecast
  - The ability to produce more reliable long-term meteorological forecasts
  - The widespread use of aerial photography
  - New techniques for reliable birth control
  - The significant increase in average life duration
  - The improvement of healthcare in under-developed areas
  - The automated shops
  - The widely usage of computers

- ***Falsely*** forecast:
  - The expansion of tropical agriculture and forestry
  - The partial control of environmental conditions and climate
  - New reliable educational techniques which will control human behavior
  - The possibility of a human body to hibernate for long periods (months or years)
  - The selection of gender for unborn babies

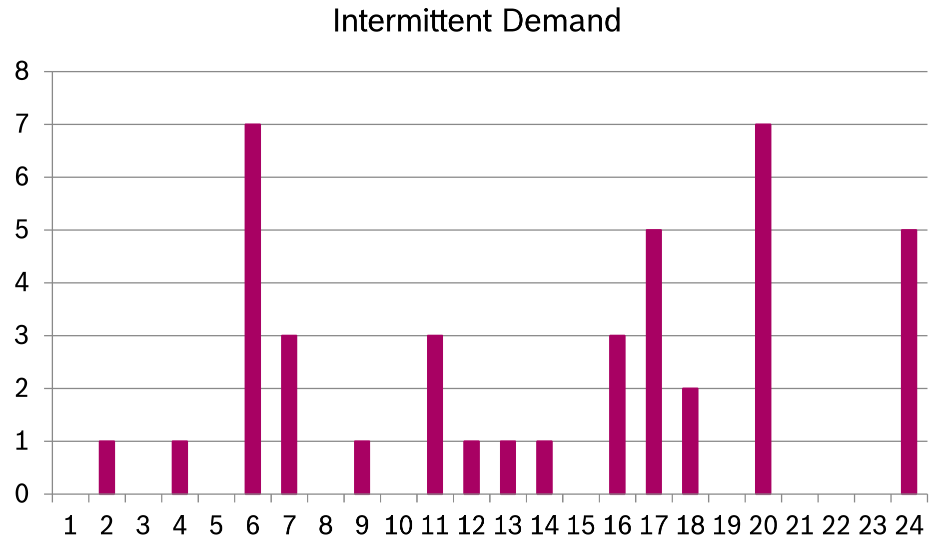
© Robert Bosch GmbH 2018. All rights reserved, also regarding any disposal, exploitation, reproduction, editing, distribution, as well as in the event of applications for industrial property rights.

BOSCH

12. SPECIAL TOPICS INTERMITTED DEMAND

Special Topics
Intermittent Demand (1)

- ***Intermittent (or sporadic) demand***: is a very common problem in business analysis.
  - It includes periods where ***demand is zero***. When demand occurs, the size varies significantly. Items with intermittent, or “slow-moving”, demand have many zero values interspersed with random spikes of non-zero demand.
- This irregular demand pattern is impossible to forecast with traditional, smoothing-based forecasting methods.
- The difficulty of forecasting lies:
  - not only in the ***discontinuity*** of demand, but
  - also in the ***large variation*** between two consecutive non-zero observations.



 Internal | CR/AEU2 - Patelis | 2018-03-14
© Robert Bosch GmbH 2018. All rights reserved, also regarding any disposal, exploitation, reproduction, editing, distribution, as well as in the event of applications for industrial property rights.

Special Topics
Intermittent Demand (2)

- This problem is especially prevalent in companies that ***manage large inventories*** of service/spare parts in industries such as:
  - aerospace, automotive, high tech, and electronics,
  - in MRO (Maintenance, Repair and Overhaul) organizations,
  - companies manufacturing high-priced capital goods,
  - stock keeping units and spare parts.
- In these businesses, as much as 70% of the parts and product items may have intermittent demand.
- The intermittent demand pattern makes it ***difficult*** to accurately ***estimate the safety stock*** and service level inventory requirements needed for successful supply chain planning.

© Robert Bosch GmbH 2018. All rights reserved, also regarding any disposal, exploitation, reproduction, editing, distribution, as well as in the event of applications for industrial property rights.

BOSCH

Special Topics
Intermittent Demand (3)

- Firms make two common mistakes when forecasting intermittent demand:
  - they focus on estimates of per period demand when they should really focus on estimates of ***the inventory stocking requirements*** necessary to meet their desired service levels.
  - they use forecasting methods that are inappropriate for intermittent demand.
- Traditional forecasting methods fail because they try to identify recognizable patterns in the demand data, such as trend and seasonality. However, intermittent demand data don’t exhibit such regular patterns and tend to be characterized by a preponderance of zero values.
- What to use?
  - Croston method
  - Syntetos & Boylan approximation - SBA
  - Aggregate-Disaggregate intermittent demand approach - ADIDA

© Robert Bosch GmbH 2018. All rights reserved, also regarding any disposal, exploitation, reproduction, editing, distribution, as well as in the event of applications for industrial property rights.

CROSTON

Special Topics
Croston (1)

- ***Croston*** (1972): An alternative method, which takes into account both the ***volume of demand*** and the ***interval*** between non-zero observations.
- The logic behind the method, is to separate the initial timeserie into two (2) other timeseries:
  - the first is the ***intervals between the demands***, and
  - the other by the ***number of independent non-zero demands***.
- Croston estimated forecasts by independently applying simple exponential smoothing in:
  - non-zero values of the timeserie, and
  - in the intervals between non-zero values of the timeserie.

© Robert Bosch GmbH 2018. All rights reserved, also regarding any disposal, exploitation, reproduction, editing, distribution, as well as in the event of applications for industrial property rights.

BOSCH

Special Topics
Croston (2)

- The two timeseries are extrapolated independently, using the ***constant level exponential smoothing*** method. In the literature, it is common to use the value of 0.05 as smoothing parameter.
- The final forecast is estimated by using the following formula:

$$DemandForecast = \frac{VolumeForecast}{Interval\ Forecast}$$

  - *Interval Forecast:* is the exponentially smoothed (or moving average) inter-demand interval, updated only if demand occurs in period.
  - *Volume Forecast:* is the exponentially smoothed (or moving average) size of demand, updated only if demand occurs in period.

➢ *Keep in mind: Other forecasting methods can also be used.*

© Robert Bosch GmbH 2018. All rights reserved, also regarding any disposal, exploitation, reproduction, editing, distribution, as well as in the event of applications for industrial property rights.
BOSCH

Special Topics
Croston (3)

- In general, intermittent demand can ***lead to increased levels*** of supply stocks, by making bias in the estimates of average demands.
- Croston however, argues that:
  - the approach of separate estimates for demand and time between demands, can lead to a reduction of prejudice (bias).
  - the additional component of frequency between the demands, allows the stock administrator to define inventory ***orders and costs with greater accuracy***, avoiding the over-stocking.

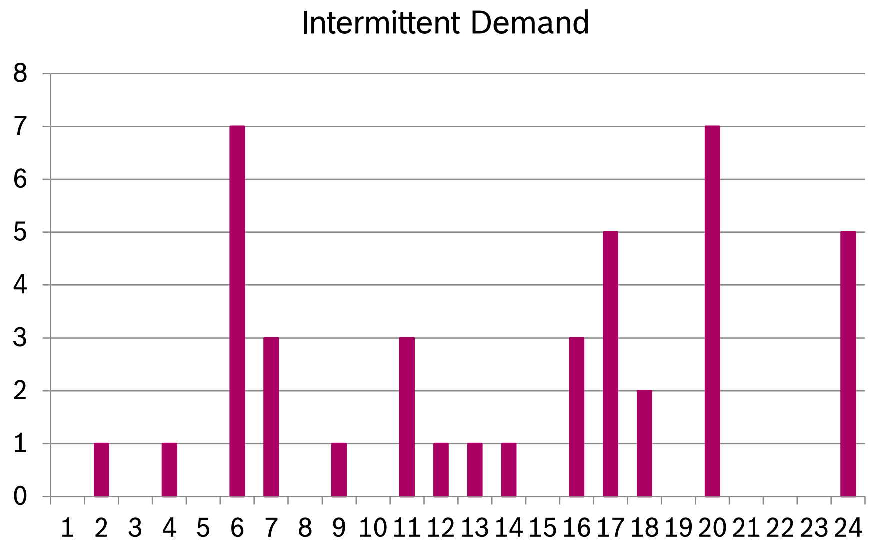
© Robert Bosch GmbH 2018. All rights reserved, also regarding any disposal, exploitation, reproduction, editing, distribution, as well as in the event of applications for industrial property rights.

Special Topics

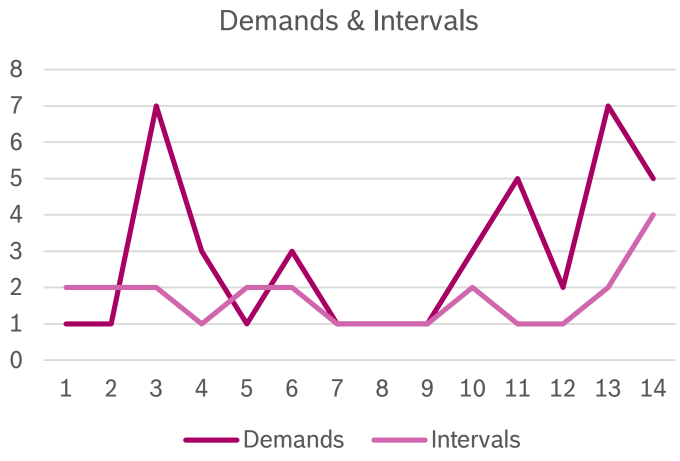
Croston: Example (1)

- Step 1: separate the initial timeserie into two (2) other timeseries

Time (t)	Intermittent Demand
1	0
2	1
3	0
4	1
5	0
6	7
7	3
8	0
9	1
10	0
11	3
12	1
13	1
14	1
15	0
16	3
17	5
18	2
19	0
20	7
21	0
22	0
23	0
24	5



Demands	Intervals
1	2
1	2
7	2
3	1
1	2
3	2
1	1
1	1
1	1
3	2
5	1
2	1
7	2
5	4



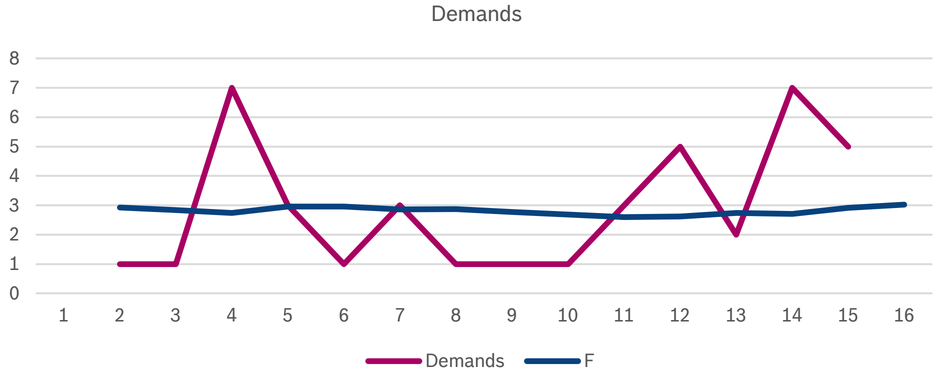
© Robert Bosch GmbH 2018. All rights reserved, also regarding any disposal, exploitation, reproduction, editing, distribution, as well as in the event of applications for industrial property rights.

Special Topics
Croston: Example (2)

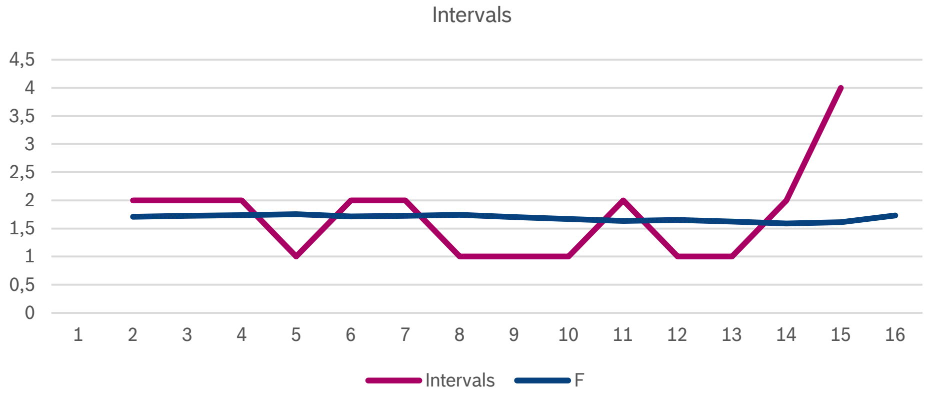
- Step 2: Use SES for extrapolation of each timeserie

Demands	F	e	S
			2,93
1	2,93	-1,93	2,83
1	2,83	-1,83	2,74
7	2,74	4,26	2,95
3	2,95	0,05	2,96
1	2,96	-1,96	2,86
3	2,86	0,14	2,87
1	2,87	-1,87	2,77
1	2,77	-1,77	2,68
1	2,68	-1,68	2,60
3	2,60	0,40	2,62
5	2,62	2,38	2,74
2	2,74	-0,74	2,70
7	2,70	4,30	2,92
5	2,92	2,08	3,02
	3,02		

Intervals	F	e	S
			1,71
2	1,71	0,29	1,72
2	1,72	0,28	1,74
2	1,74	0,26	1,75
1	1,75	-0,75	1,71
2	1,71	0,29	1,73
2	1,73	0,27	1,74
1	1,74	-0,74	1,70
1	1,70	-0,70	1,67
1	1,67	-0,67	1,64
2	1,64	0,36	1,65
1	1,65	-0,65	1,62
1	1,62	-0,62	1,59
2	1,59	0,41	1,61
4	1,61	2,39	1,73
	1,73		



$$S_0 = \overline{Demands}$$



$$S_0 = \overline{Intervals}$$

© Robert Bosch GmbH 2018. All rights reserved, also regarding any disposal, exploitation, reproduction, editing, distribution, as well as in the event of applications for industrial property rights.

Special Topics
Croston: Example (3)

- Step 3: Combine the forecasts

Forecasts		
Demands	Intervals	Croston
2,93	1,71	1,71
2,83	1,72	1,64
2,74	1,74	1,58
2,95	1,75	1,69
2,96	1,71	1,73
2,86	1,73	1,65
2,87	1,74	1,65
2,77	1,70	1,63
2,68	1,67	1,61
2,60	1,64	1,59
2,62	1,65	1,58
2,74	1,62	1,69
2,70	1,59	1,70
2,92	1,61	1,81
3,02	1,73	1,75

Time (t)	Data X	Forecasts
1	0	1,71
2	1	1,71
3	0	1,64
4	1	1,64
5	0	1,58
6	7	1,58
7	3	1,69
8	0	1,73
9	1	1,73
10	0	1,65
11	3	1,65
12	1	1,65
13	1	1,63
14	1	1,61
15	0	1,59
16	3	1,59
17	5	1,58
18	2	1,70
19	0	1,70
20	7	1,81
21	0	1,81
22	0	1,81
23	0	1,81
24	5	1,75

© Robert Bosch GmbH 2018. All rights reserved, also regarding any disposal, exploitation, reproduction, editing, distribution, as well as in the event of applications for industrial property rights.

SYNTETOS & BOYLAN APPROXIMATION – SBA

Special Topics
Syntetos & Boylan (1)

- Empirical evidence has shown that the gains from ***Croston implementation is worse than expected***, compared with simpler forecasting techniques. In some cases, even worse performance is observed.
- Syntetos & Boylan (2001) tried to identify the cause of this unexpected behavior.
  - they found that Croston method is ***positively biased***, thus has an optimistic trend in the forecasting results.
  - they managed to associate the ***level of optimistic trend with the smoothing parameter*** which used for the extrapolation of the two decomposed timeseries.
  - the maximum bias observed when the smoothing parameter has the maximum value of 1. In general, they observed large optimistic bias ***when the smoothing parameter is large***.
  - An empirical result is that Croston method must be used with a smoothing parameter ***not greater that 0.15***.

© Robert Bosch GmbH 2018. All rights reserved, also regarding any disposal, exploitation, reproduction, editing, distribution, as well as in the event of applications for industrial property rights.

BOSCH

Special Topics
Syntetos & Boylan (2)

- The Syntetos and Boylan Approximation (SBA) is in fact a ***modification*** of the Croston method, where the forecasts are estimated using the formula:

$$F_{SBA} = \left(1 - \frac{a}{2}\right) \times \frac{VolumeForecast}{IntervalForecast}$$

© Robert Bosch GmbH 2018. All rights reserved, also regarding any disposal, exploitation, reproduction, editing, distribution, as well as in the event of applications for industrial property rights.
BOSCH

ADIDA

Special Topics
ADIDA (1)

- ***Aggregate-Disaggregate Intermittent Demand Approach***: A process of non-overlapping aggregation of data in less-frequency periods.
  - in case of monthly data, a quarterly aggregation can be applied, setting the level of aggregation to 3 months.
- Advantages:
  - the discontinuity of data (due to the existence of zero values) may be reduced or eliminated.
  - the variation of the new timeserie is expected to be reduced, since a non-overlapping moving average for smoothing has been applied.
  - The proper definition of the level of aggregation will result a ***continuous demand timeserie***, without zero values, in which any forecasting technique can be applied in order to estimate forecasts for the aggregate level.

© Robert Bosch GmbH 2018. All rights reserved, also regarding any disposal, exploitation, reproduction, editing, distribution, as well as in the event of applications for industrial property rights.

Special Topics
ADIDA (2)

- ***Four steps*** for using the ADIDA method:

1. Analysis of the initial timeserie, for selecting the appropriate level of aggregation.
2. Apply the aggregation on the initial data, for creating a new continuous timeserie.
3. Estimate forecasts, by using a predictive model.
4. Disaggregate forecasts, by using one of the following ways:
   - ***Equal weights***: Break the forecasts with similar weights. Suitable for timeseries with great randomness and no seasonal behavior.
   - ***Previous weights***: Break the forecasts by using the observed weights of the last m observations (where m is the aggregation level).
   - ***Average weights***: Break the forecasts by using the average weights (divide all observations into k groups of m observations each, where k*m is equal to the set of available observations). Indicated for timeseries with significant seasonality.

© Robert Bosch GmbH 2018. All rights reserved, also regarding any disposal, exploitation, reproduction, editing, distribution, as well as in the event of applications for industrial property rights.

BOSCH

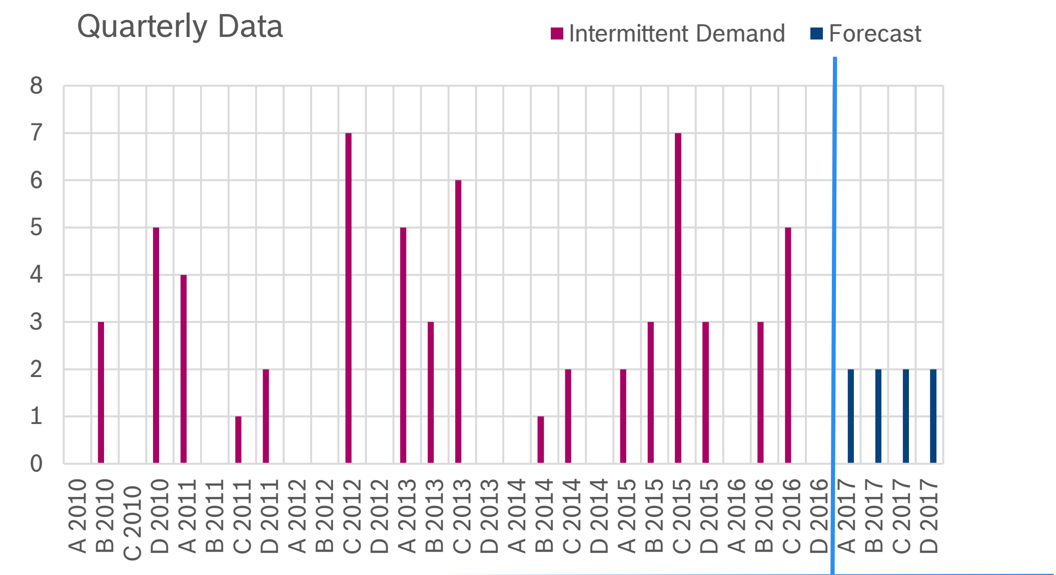
Special Topics
ADIDA (3)

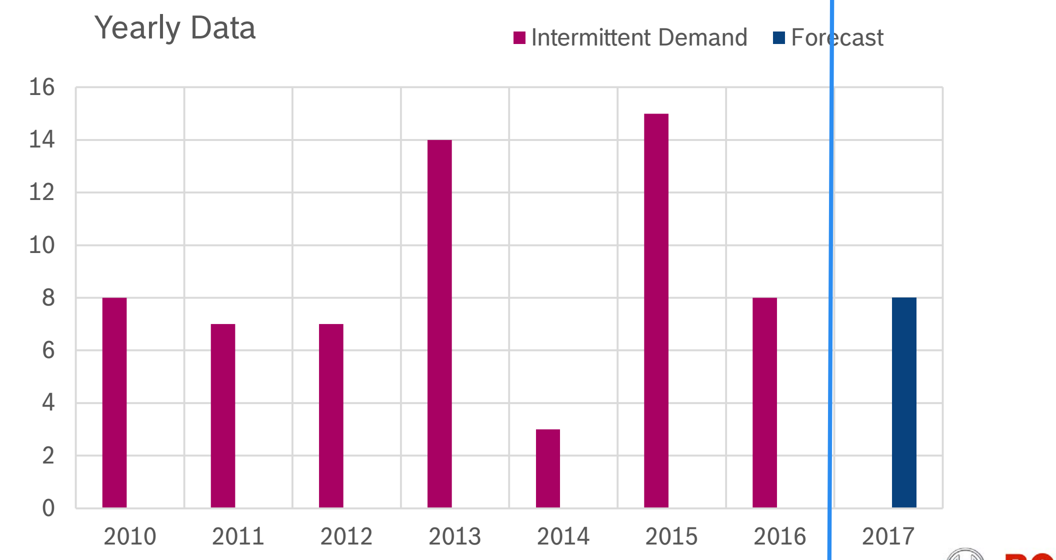
- ***Characteristics***:
  - The non-overlapping data aggregation seems to be a ***promising approach*** for timeseries with interruptible demand, since forecasts in higher levels of aggregation are in general more accurate and have less variation that those in smaller aggregation levels.
  - It is common to use as aggregation level the ***number of forecasting horizon***.
- Empirical results had showed that:
  - there is probably an "optimal" level of aggregation, while the definition of the level of aggregation equal to the lead time plus a period leads to very good results regarding the predictive accuracy.

© Robert Bosch GmbH 2018. All rights reserved, also regarding any disposal, exploitation, reproduction, editing, distribution, as well as in the event of applications for industrial property rights.
BOSCH

Special Topics
ADIDA: Example

- Step 1: Level of aggregation selection.
- Step 2: Apply aggregation.
- Step 3: Estimate forecasts.
  - example: Naïve method
- Step 4: Disaggregate forecasts.
  - example: by equal weights





© Robert Bosch GmbH 2018. All rights reserved, also regarding any disposal, exploitation, reproduction, editing, distribution, as well as in the event of applications for industrial property rights.

13. SPECIAL TOPICS STRUCTURAL CHANGES DETECTION

BOSCH

Special Topics

Structural Changes Detection (1)

- ***Significant changes*** (example: population growth, change of demand of product or service, other products entering a market) can affect a timeserie trend and values.
  - This must be taken into account in all business activities, which are focusing ***in minimizing potential costs*** or ***maximizing profits*** by taking advantage of new opportunities.
- Companies face a emerging need to make decisions that will interrupt the normal production operations and will restrict existing resources and assets.
- ***Tracking of timeseries*** scope is to automatically ***detect significant changes*** in the pattern of a timeserie (for example sudden increases or decreases) as new data become available, ***as soon as possible***, and with the ***least possible false detections***.

© Robert Bosch GmbH 2018. All rights reserved, also regarding any disposal, exploitation, reproduction, editing, distribution, as well as in the event of applications for industrial property rights.
BOSCH

Special Topics
Structural Changes Detection (2)

- Important to use automated tracking applications with ***low cost***.
- These applications must label the timeseries correctly as “no-change” and “significant change”, in order to alert a company to make the necessary business actions.
- Timeseries tracking is directly linked to forecasting:
  - ***change detection is mainly based on the forecasting error***, which resulted from an extrapolation of the timeserie in the future.
    - Several surveys have studied the relationship between the accuracy of timeseries forecasting and the successfully detection of significant pattern changes in a timeserie.
  - tracking timeseries is (or it should be) a ***key component*** in business forecasting systems installed in a company, since business intervention required when simple predictive methods leads to large prediction errors.

© Robert Bosch GmbH 2018. All rights reserved, also regarding any disposal, exploitation, reproduction, editing, distribution, as well as in the event of applications for industrial property rights.

Special Topics
Structural Changes Detection (3)

- The main objective of the forecast is:
  - an estimate of future average value, taking into account the timeserie trend through appropriate methods and parameters and
  - assuming that within the forecast horizon a significant (structural) change in the data pattern will not take place.
- When a relatively ***large forecasting error*** is observed, or when a ***sequence of forecasting errors*** have the same direction (under or over forecasted), then the forecasts are biased.
  - This can be the result of a change in the timeserie pattern, which requires further analysis and control.
- The exponential smoothing methods will adjust their forecasts with a delay of one period, after the completion of the pattern change. But, the estimation errors in the interim periods are enough to detect the structural change on the timeserie data pattern.

© Robert Bosch GmbH 2018. All rights reserved, also regarding any disposal, exploitation, reproduction, editing, distribution, as well as in the event of applications for industrial property rights.

Special Topics
Brown Tracking Method (1)

- The ***Brown method*** of *k-periods* for a given period *n* is calculated as the absolute ration between:
  - the sum of forecasts errors of period *n* and the *k-1* previous periods, and
  - the exponential smoothing average deviation of the absolute errors at period *n*. *(β=smoothing parameter)*

$$Brown_k = \left|\frac{CUSUM_n^k}{MAD_n}\right| = \left|\frac{\sum_{t=n-k+1}^{n} e_t}{(1-\beta)^n * e_0 + \sum_{t=1}^{n} \beta * (1-\beta)^{n-t} * |e_t|}\right|,$$

  - e_0 is the initialization point for MAD and can be calculated by using the following formula:

$$e_0 = \frac{\sum_{t=1}^{6} |e_t|}{6}$$

© Robert Bosch GmbH 2018. All rights reserved, also regarding any disposal, exploitation, reproduction, editing, distribution, as well as in the event of applications for industrial property rights.

BOSCH

Special Topics
Brown Tracking Method (2)

- Originally proposed by Brown (1959, 1963) for estimating CUSUM index for all periods of the timeserie.
- But, Trigg (1964) observed that CUSUM index must be reset immediately after a significant change in the timeserie pattern.
  - If not, the method ***may produce incorrect outputs*** as detections of significant changes, since the cumulative sum will include large errors calculated in the previous periods.
- This observation led to the ***Trigg tracking method***.
  - It incorporates exponential smoothing into the numerator also, in order to significantly reduce the influence of past periods estimation errors.

© Robert Bosch GmbH 2018. All rights reserved, also regarding any disposal, exploitation, reproduction, editing, distribution, as well as in the event of applications for industrial property rights.

Special Topics
Trigg Tracking Method (1)

- Trigg method is calculated by replacing the CUSUM index in the numerator, with a sum of exponentially smoothing forecasting errors E_n.
  - the numerator exponential smoothing allows the index to apply lower weight to forecast errors, as the time lag between the current periods.

$$Trigg_k = \left|\frac{E_n}{MAD_n}\right| = \left|\frac{\sum_{t=1}^{n} \alpha * (1-\alpha)^{n-t} * e_t}{(1-\beta)^n * e_0 + \sum_{t=1}^{n} \beta * (1-\beta)^{n-t} * |e_t|}\right|$$

  - This change causes the pointer to reset automatically, after detecting the pattern change.

- It should be noted that the Trigg method with $\alpha = 1$ is identical to Brown method when $k = 1$.

© Robert Bosch GmbH 2018. All rights reserved, also regarding any disposal, exploitation, reproduction, editing, distribution, as well as in the event of applications for industrial property rights.

Special Topics
Selecting smoothing parameter

- How to select the ***α*** and ***β*** parameters:
  - Gardner (1983) and Golden and Settle (1976) have originally proposed to use the same value, ***equal to the optimal smoothing parameter*** for the exponential smoothing method of constant level (SES).
  - McClain (1988) showed that the selection of different values for α and β has generally better results, suggesting ***lower values for β***.
  - In practice:

$$0.05 \leq \alpha \leq 1$$

$$0.05 \leq \beta \leq 0.5$$

$$\beta \leq \alpha$$

© Robert Bosch GmbH 2018. All rights reserved, also regarding any disposal, exploitation, reproduction, editing, distribution, as well as in the event of applications for industrial property rights.
BOSCH

Special Topics
Tracking the change

- The detection of significant changes, according to Brown and Trigg methods, can be achieved for the period where the calculated index value ***exceeds a certain activation limit*** (threshold).
- The change of the threshold value implies a modification of the method sensitivity.
- In practice, the following activation rules are usually applied, depending on the selected tracking method:

$$if\ Brown_k\ > 0.5\ then\ "Structural\ Change\ Detected"$$

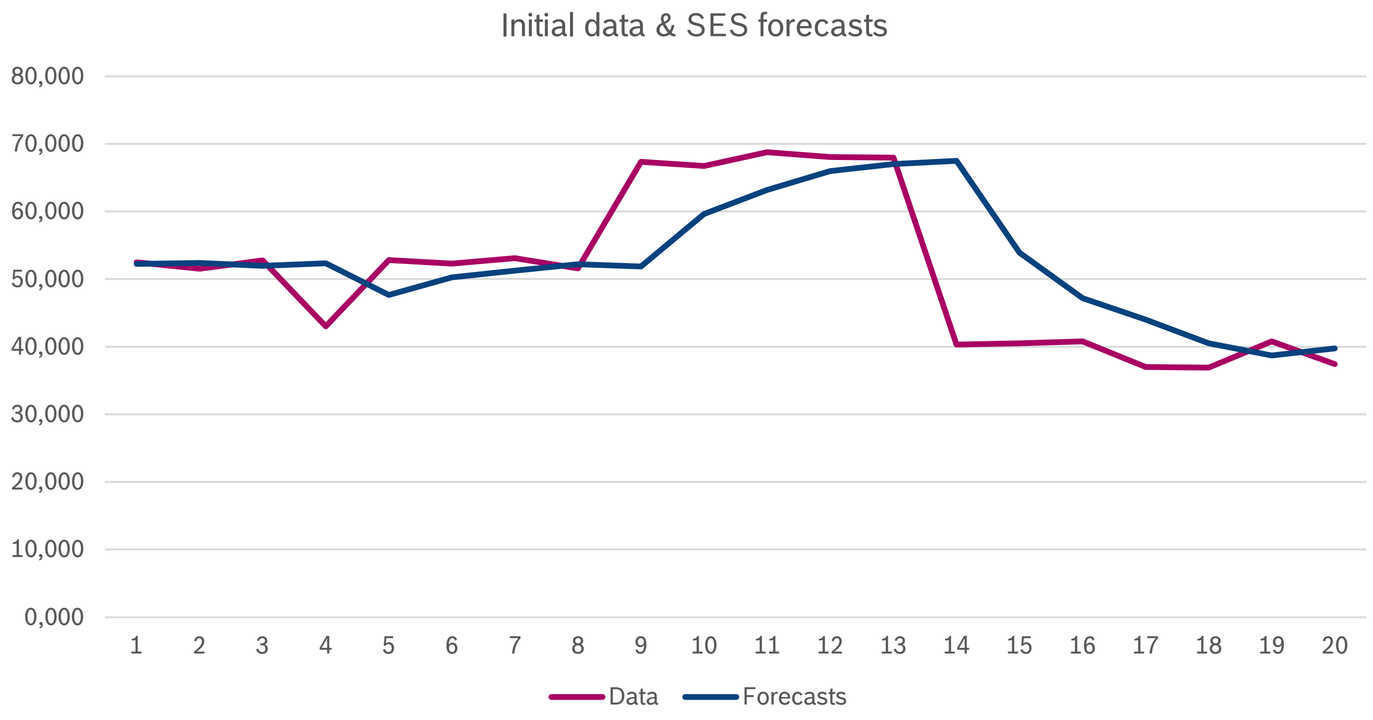
$$if\ Trigg\ > 0.5\ then\ "Structural\ Change\ Detected"$$

Internal | CR/AEU2 - Patelis | 2018-03-14
© Robert Bosch GmbH 2018. All rights reserved, also regarding any disposal, exploitation, reproduction, editing, distribution, as well as in the event of applications for industrial property rights.
BOSCH

Special Topics
Structural Changes Detection: Example (1)

- ***Step 1*:** Apply simple exponential smoothing for estimating forecasts.

S(0) =	52,255	α=0.5		
Time (t)	Data	Forecasts	Error e	Level S
1	52,474	52,255	0,219	52,365
2	51,543	52,365	-0,822	51,954
3	52,749	51,954	0,795	52,351
4	43,000	52,351	-9,351	47,676
5	52,806	47,676	5,130	50,241
6	52,298	50,241	2,057	51,269
7	53,092	51,269	1,823	52,181
8	51,583	52,181	-0,598	51,882
9	67,362	51,882	15,480	59,622
10	66,731	59,622	7,109	63,176
11	68,770	63,176	5,594	65,973
12	68,079	65,973	2,106	67,026
13	67,968	67,026	0,942	67,497
14	40,321	67,497	-27,176	53,909
15	40,491	53,909	-13,418	47,200
16	40,790	47,200	-6,410	43,995
17	36,985	43,995	-7,010	40,490
18	36,897	40,490	-3,593	38,694
19	40,801	38,694	2,107	39,747
20	37,441	39,747	-2,306	38,594



Internal | CR/AEU2 - Patelis | 2018-03-14
© Robert Bosch GmbH 2018. All rights reserved, also regarding any disposal, exploitation, reproduction, editing, distribution, as well as in the event of applications for industrial property rights.

Special Topics
Structural Changes Detection: Example (2)

$$e_0 = \frac{\sum_{t=1}^{6} |e_t|}{6} = 3{,}06$$

- ***Step 2***: Estimate Trigg index.

				Trigg		
Time (t)	Data	Forecasts	Error e	Nominator	Denominator	Index
1	52,474	52,255	0,219	0,044	2,492	0,018
2	51,543	52,365	-0,822	-0,129	2,159	0,060
3	52,749	51,954	0,795	0,056	1,887	0,030
4	43,000	52,351	-9,351	-1,826	3,380	0,540
5	52,806	47,676	5,130	-0,435	3,730	0,117
6	52,298	50,241	2,057	0,064	3,395	0,019
7	53,092	51,269	1,823	0,415	3,081	0,135
8	51,583	52,181	-0,598	0,213	2,584	0,082
9	67,362	51,882	15,480	3,266	5,163	0,633
10	66,731	59,622	7,109	4,035	5,552	0,727
11	68,770	63,176	5,594	4,347	5,561	0,782
12	68,079	65,973	2,106	3,898	4,870	0,800
13	67,968	67,026	0,942	3,307	4,084	0,810
14	40,321	67,497	-27,176	-2,790	8,702	0,321
15	40,491	53,909	-13,418	-4,915	9,645	0,510
16	40,790	47,200	-6,410	-5,214	8,999	0,579
17	36,985	43,995	-7,010	-5,573	8,601	0,648
18	36,897	40,490	-3,593	-5,177	7,599	0,681
19	40,801	38,694	2,107	-3,720	6,501	0,572
20	37,441	39,747	-2,306	-3,438	5,662	0,607

$$E_1 = \sum_{t=1}^{1} a \times (1-a)^{1-t} \times e_t = 0{,}2 \times (1-0{,}2)^{1-1} \times e_1 = 0{,}044$$

$$MAD_1 = (1-\beta)^1 \times e_0 + \sum_{t=1}^{1} \beta \times (1-\beta)^{1-t} \times |e_t| = 0{,}8 \times 3{,}06 + 0{,}2 \times (1-0{,}2)^0 \times |0{,}219| = 2{,}492$$

$$Trigg_1 = \left|\frac{E_1}{MAD_1}\right| = \left|\frac{0{,}044}{2{,}494}\right| = 0{,}018$$

$$E_2 = \sum_{t=1}^{2} a \times (1-a)^{2-t} \times e_t = 0{,}2 \times (1-0{,}2)^{2-1} \times e_1 + 0{,}2 \times (1-0{,}2)^{2-2} \times e_2 = -0{,}129$$

$$MAD_2 = (1-\beta)^2 \times e_0 + \sum_{t=1}^{2} \beta \times (1-\beta)^{2-t} \times |e_t| =$$
$$= 0{,}8^2 \times 3{,}06 + 0{,}2 \times (1-0{,}2)^1 \times |0{,}219| + 0{,}2 \times (1-0{,}2)^0 \times |-0{,}822| = 2.159$$

$$Trigg_2 = \left|\frac{E_2}{MAD_2}\right| = \left|\frac{-0{,}129}{2{,}159}\right| = 0{,}060$$

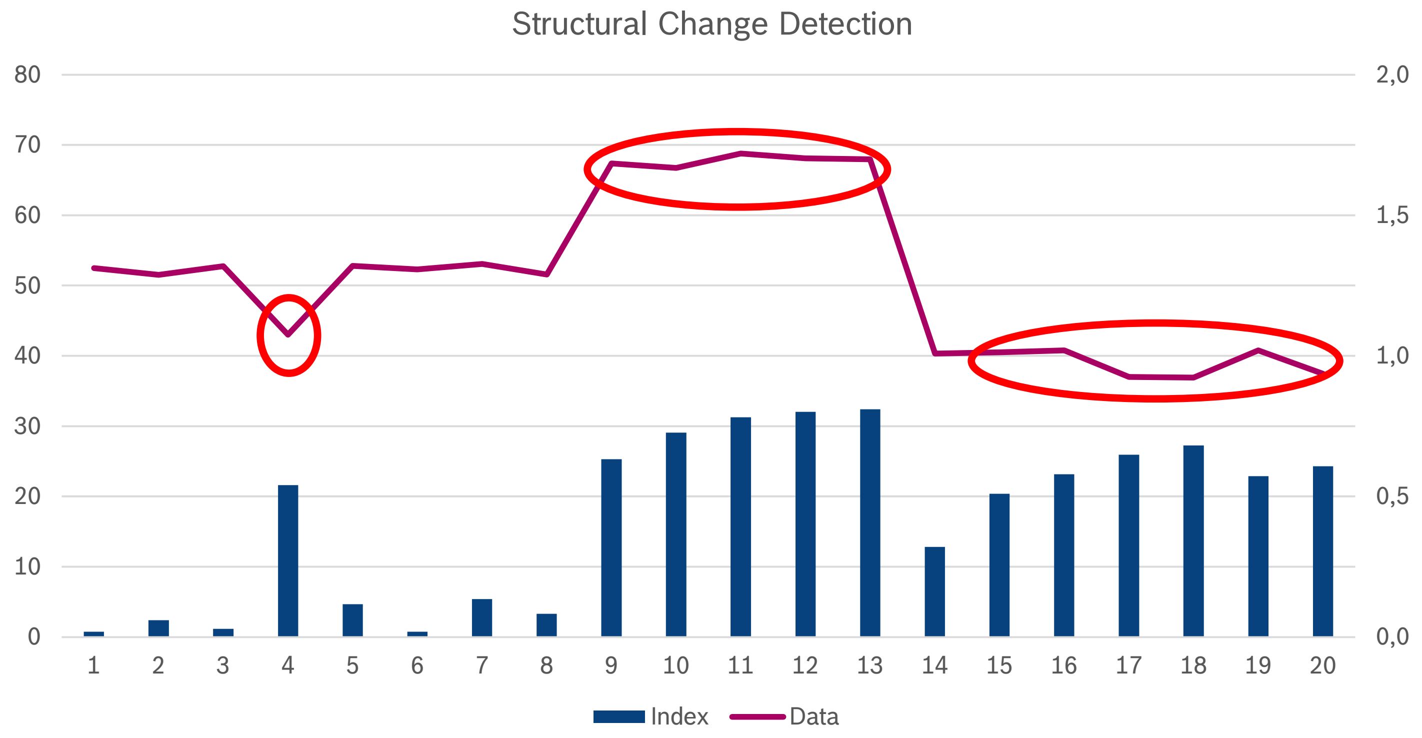
© Robert Bosch GmbH 2018. All rights reserved, also regarding any disposal, exploitation, reproduction, editing, distribution, as well as in the event of applications for industrial property rights.

BOSCH

Special Topics
Structural Changes Detection: Example (3)

- ***Step 3***: Locate structural Changes (Trigg > 0,5)

Time (t)	Data	Index	Str.Change
1	52,474	0,018	
2	51,543	0,060	
3	52,749	0,030	
4	43,000	0,540	yes
5	52,806	0,117	
6	52,298	0,019	
7	53,092	0,135	
8	51,583	0,082	
9	67,362	0,633	yes
10	66,731	0,727	yes
11	68,770	0,782	yes
12	68,079	0,800	yes
13	67,968	0,810	yes
14	40,321	0,321	
15	40,491	0,510	yes
16	40,790	0,579	yes
17	36,985	0,648	yes
18	36,897	0,681	yes
19	40,801	0,572	yes
20	37,441	0,607	yes



Internal | CR/AEU2 - Patelis | 2018-03-14
© Robert Bosch GmbH 2018. All rights reserved, also regarding any disposal, exploitation, reproduction, editing, distribution, as well as in the event of applications for industrial property rights.
BOSCH

14. INTERESTING USE-CASES

INVENTORY DEMAND FORECASTING

Interesting Use-Cases
Inventory demand forecasting (1)

- ***Scope***: A grocery store needs to forecast demand for perishable items
  - Too many → discard valuable product
  - Too few → run out of stock
  - It a challenging task to ***optimise***.

- Two ways of defining a model:
  - By reducing the amount of overstocked items
  - By increasing the amount of understocked items.

© Robert Bosch GmbH 2018. All rights reserved, also regarding any disposal, exploitation, reproduction, editing, distribution, as well as in the event of applications for industrial property rights.

BOSCH

Source: A machine learning approach to Inventory demand forecasting

Interesting Use-Cases
Inventory demand forecasting (2)

- ***Overstocked*** items:
  - A direct loss since expired items must be discarded.
  - A 10% less rate, will result 54.750 € per year.
- Similar issues:
  - Clothing store that overstocks winter coats
    - Need for bigger storage rooms.
    - Reduce limited store space.
    - Must sell coats at discount or loss.

Grocery Chain		
Stores	10	10
Products per Store	10	10
Items per product	100	100
Discard rate	**5,0%**	**4,5%**
Average cost of product	3,00 €	3,00 €
Daily loss	1.500,00 €	1.350,00 €
Annual loss	547.500,00 €	492.750,00 €

© Robert Bosch GmbH 2018. All rights reserved, also regarding any disposal, exploitation, reproduction, editing, distribution, as well as in the event of applications for industrial property rights.

Interesting Use-Cases
Inventory demand forecasting (3)

- ***Understock Items***:
  - A more severe issue.
  - Loss of sales.
  - Lack of product may drive customers to a competitor.

- Similar issue:
  - ***understaffing a call-center*** on a busy day can result in long wait times and poor service.

© Robert Bosch GmbH 2018. All rights reserved, also regarding any disposal, exploitation, reproduction, editing, distribution, as well as in the event of applications for industrial property rights.

Interesting Use-Cases
Inventory demand forecasting (4)

- ***Step 1***: Data
  - Collect, interpret, and analyse historical data.
  - Types of data:
    - Products on stock per day/week.
    - Sales of products per day/week.
    - Additional data (such as: store opening hours, promotion activities, weather data, holidays, etc.)
  - Volume of data:
    - At least 2 to 3 years of data, in order to monitor possible seasonal trends.

© Robert Bosch GmbH 2018. All rights reserved, also regarding any disposal, exploitation, reproduction, editing, distribution, as well as in the event of applications for industrial property rights.
BOSCH

Interesting Use-Cases
Inventory demand forecasting (5)

- ***Step 2***: Define the model
  - What to forecast:
    - Depends on the goal of the business (sales?, minimum overstocked? maximum understocked?).
    - In our case: ***Forecast daily sales for each product in each store***.
  - How:
    - Select a statistical method to use.
    - In our case: Exponential Smoothing.
  - Evaluation metric:
    - There is no best evaluation metric.
    - Again, it depends on the business.
    - In our case: Root Mean Square Error (RMSE).

© Robert Bosch GmbH 2018. All rights reserved, also regarding any disposal, exploitation, reproduction, editing, distribution, as well as in the event of applications for industrial property rights.
BOSCH

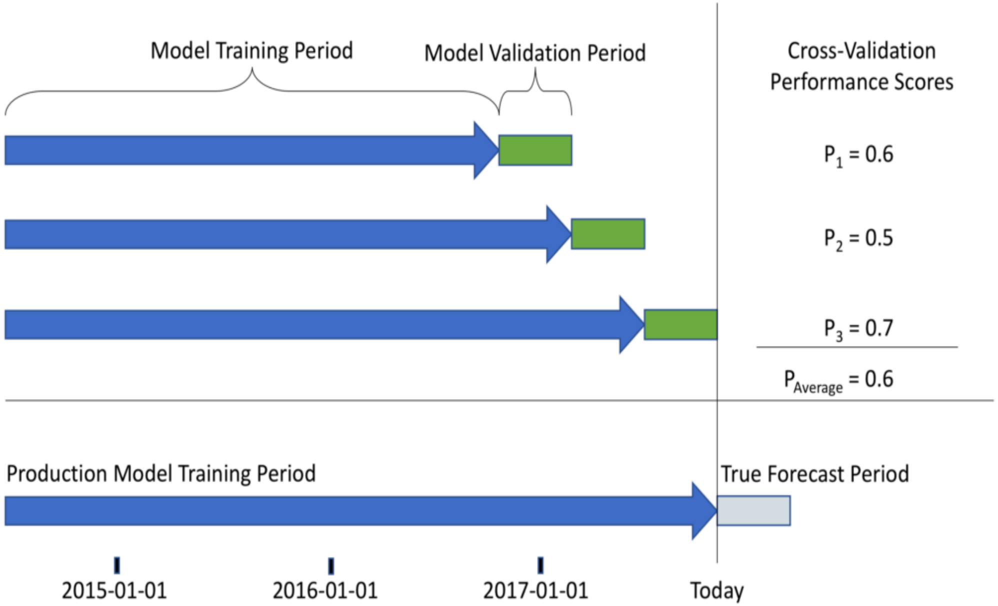
Interesting Use-Cases
Inventory demand forecasting (6)

- ***Step 3***: Define a benchmark
  - We need a baseline score, in order to evaluate our model.
  - Common benchmark:
    - using ***Naive*** (product sales next week will be the same as last week), or
    - using ***Seasonal Naïve*** (product sales next month with be as the same month last year).
    - using ***Judgmental*** (ask each store manager to make predictions).
  - In our case: Benchmark = 0.8

© Robert Bosch GmbH 2018. All rights reserved, also regarding any disposal, exploitation, reproduction, editing, distribution, as well as in the event of applications for industrial property rights.
BOSCH

Interesting Use-Cases
Inventory demand forecasting (7)

- ***Step 4***: Validate the proposed model
  - Let's assume that the historical data are from 01.01.2015 until 31.12.2017.
  - Train the model ***using a data subset***, and predict sales for the next period (week/month). Estimate RMSE(1).
  - Do the same with ***N*** different data subsets.
  - Estimate the average RMSE.
  - Example: ***Model performance = 0.6***
- We can validate more models.



- If the benchmark is better that the model then:
  - Our model is not appropriate, or
  - The store managers are pretty good!

© Robert Bosch GmbH 2018. All rights reserved, also regarding any disposal, exploitation, reproduction, editing, distribution, as well as in the event of applications for industrial property rights.

Interesting Use-Cases
Inventory demand forecasting (8)

- ***Step 5***: Finalize
  - Implement the best performing model into production.
  - Monitor the real-time performance frequently.
    - Use the model for real future forecasts.
    - Check the forecasts for “strange” values (too high, too low, negative, etc.).
    - ***Check if it performs as good as in the evaluation***.

© Robert Bosch GmbH 2018. All rights reserved, also regarding any disposal, exploitation, reproduction, editing, distribution, as well as in the event of applications for industrial property rights.
BOSCH

INTELLIGENT ENERGY MANAGEMENT IN BUILDINGS

Interesting Use-Cases
Intelligent Energy Management in Buildings (1)

- **Scope**: a DSS for the energy managers of buildings, which can assist them in setting indoor temperature set point, based on the feedback received by the occupants.
  - ***optimize energy use***, which is achieved though the efficient management of the heating and cooling systems
    - energy savings and thermal comfort are not always compatible.
    - indoor conditions shaped are not the ideal ones from the occupants' perspective (municipal building).
    - Current results: case-dependent, lead to relative time-independent suggestions (might not always be appropriate).
  - ***Set-point management*** should be dynamically applied in buildings, aiming both at:
    - Creating acceptable comfort levels for occupants by grasping their thermal sensations feeling, and
    - Achieving energy consumption reduction, leading to energy and cost savings.
  - We need a system which effectively captures the "real" sensation of the users and correlates it with the "predicted" one to detect a range of accepted set point temperatures.

© Robert Bosch GmbH 2018. All rights reserved, also regarding any disposal, exploitation, reproduction, editing, distribution, as well as in the event of applications for industrial property rights.

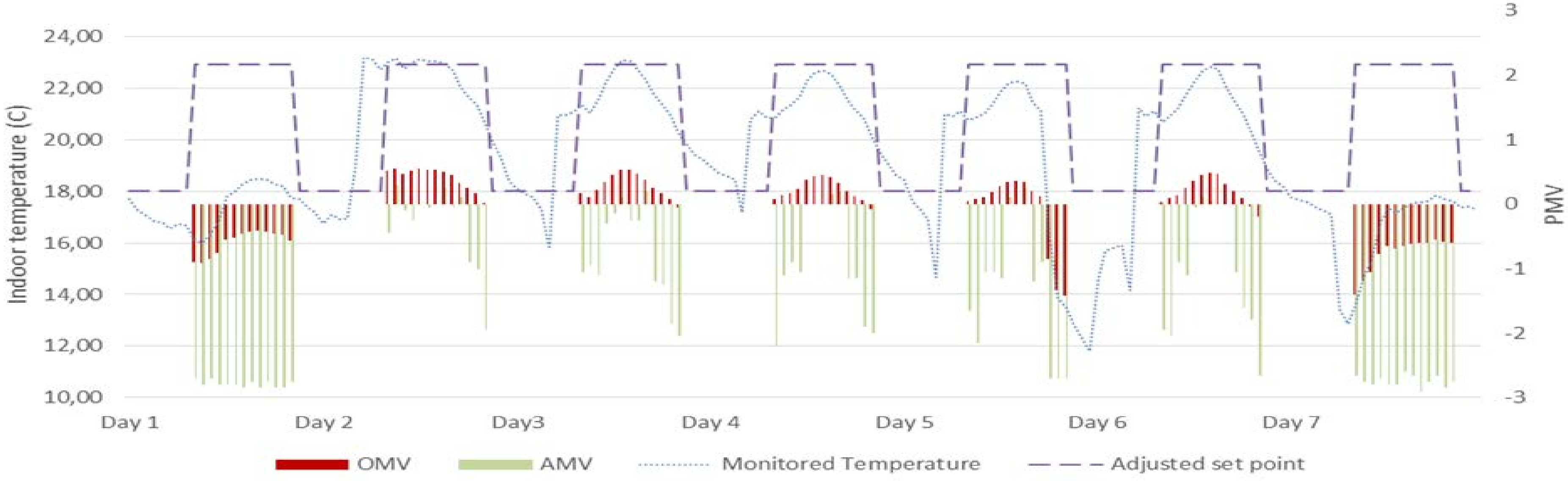
Interesting Use-Cases
Intelligent Energy Management in Buildings (2)

- **Metrics**:
  - Predicted Mean Vote (PMV) index, a 7-point thermal sensation scale.
    - can be calculated (EN ISO 7730:2005) by using predictions for the indoor conditions of a building.
  - Observed Mean Vote (OMV) index.
    - can be calculated using monitored values of temperature and humidity, as measured by building sensors in real-time.
  - Actual Mean Vote (AMV) though Thermal Comfort Validator (TCV).
    - derived directly by the TCV, (http://validator.optimussmartcity.eu), a web application, where occupants are encouraged to submit structured feedback on their thermal sensation.

© Robert Bosch GmbH 2018. All rights reserved, also regarding any disposal, exploitation, reproduction, editing, distribution, as well as in the event of applications for industrial property rights.

Interesting Use-Cases
Intelligent Energy Management in Buildings (3)

- Visualization of the ***real indoor temperature*** monitored, the ***set point temperature suggested***, as well as the OMV/AMV values estimated for the operating hours of the Sant-Cugat Town Hall (Spain)



__Source__: *"Decision Support System for Intelligent Energy Management in Buildings using the Thermal Comfort Model", Marinakis et. al., International Journal of Computational Intelligence Systems, Vol. 10 (2017) 882–893*

© Robert Bosch GmbH 2018. All rights reserved, also regarding any disposal, exploitation, reproduction, editing, distribution, as well as in the event of applications for industrial property rights.

THE FORECASTING POWER OF SOCIAL MEDIA

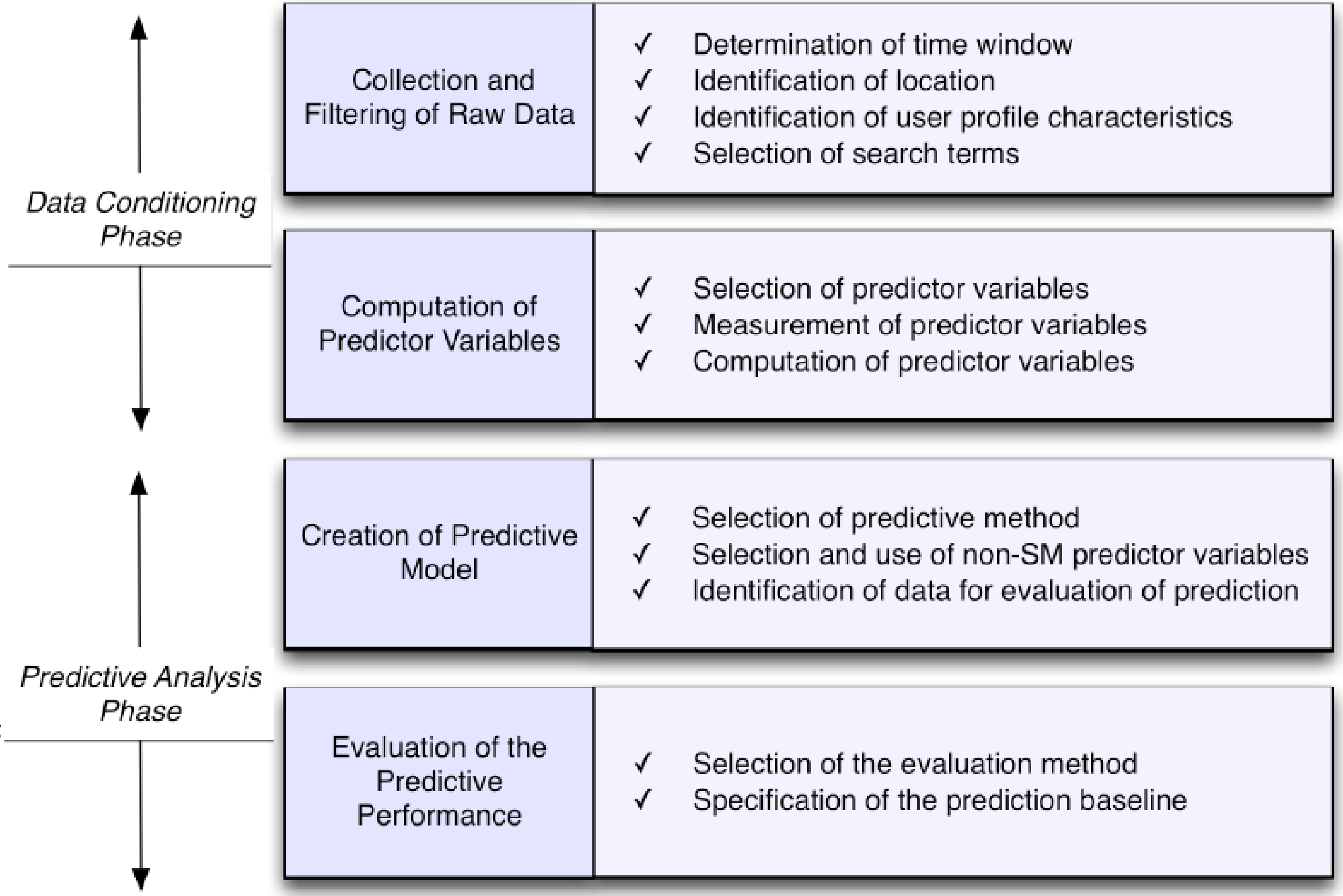
Interesting Use-Cases
The forecasting power of social media (1)

- Use of ***Social Media (SM) data*** for estimating forecasts in various areas, such as:
  - Product sales,
  - Stock market volatility,
  - Disease outbreaks,
  - Elections outcome,
  - etc.
- SM data includes: textual content, rating scores, like or dislike indicators, web search queries, tags and profile information.
- SM data incorporates personal opinions, thoughts and behaviors making it a vital component of the Web and a fertile ground for a variety of business and research endeavours.

© Robert Bosch GmbH 2018. All rights reserved, also regarding any disposal, exploitation, reproduction, editing, distribution, as well as in the event of applications for industrial property rights.
BOSCH

Interesting Use-Cases
The forecasting power of social media (2)

- **Data Conditioning phase**:
  - Transformation of noisy raw data into ***high quality data*** that are structured based on some predictor variables.
    - When? Where? Who? What?
- **Predictive Analysis phase**:
  - Creation and evaluation of a ***forecasting model*** that enables estimating outcome from a new set of observations



© Robert Bosch GmbH 2018. All rights reserved, also regarding any disposal, exploitation, reproduction, editing, distribution, as well as in the event of applications for industrial property rights.
BOSCH

Interesting Use-Cases
The forecasting power of social media (3)

- ***SM-Predictor variables***:
  - ***Volume-related*** variables: they measure the amount of SM data
    - Examples: number of tweets, number of reviews, number of queries etc.
  - ***Sentiment-related*** variables: they measure the sentiment expressed through the data.
    - Examples: bullishness index, review valence, review rating, etc.
  - ***Profile characteristics*** of online users:
    - Examples: Facebook friends, number of followers of users that posted a tweet, total posts, the location of the reviewer, in-degree.
- ***Non SM-Predictor variables***:
  - Past values of phenomenon, demographics, budget, etc.

© Robert Bosch GmbH 2018. All rights reserved, also regarding any disposal, exploitation, reproduction, editing, distribution, as well as in the event of applications for industrial property rights.
BOSCH

Interesting Use-Cases
The forecasting power of social media (4)

- ***Forecasting method***:
  - Most common in literature: Linear Regression.
  - Other methods: Markov models, neural networks, support vector machine, Causality models

- ***Don’t confuse forecasting and explanation!***
  - Forecasting power refers to the ability of predicting new observations accurately, while
  - Explanatory power to the strength of association indicated by a statistical model.
  - *“A statistically significant effect or relationship* ***does not guarantee*** *high predictive power, because the precision or magnitude of the causal effect might not be sufficient for obtaining levels of predictive accuracy that are practically meaningful”.*

Source: *“Understanding the Predictive power of Social Media”, Kalampokis et. al., Internet Research, Vol. 23 (2013)*

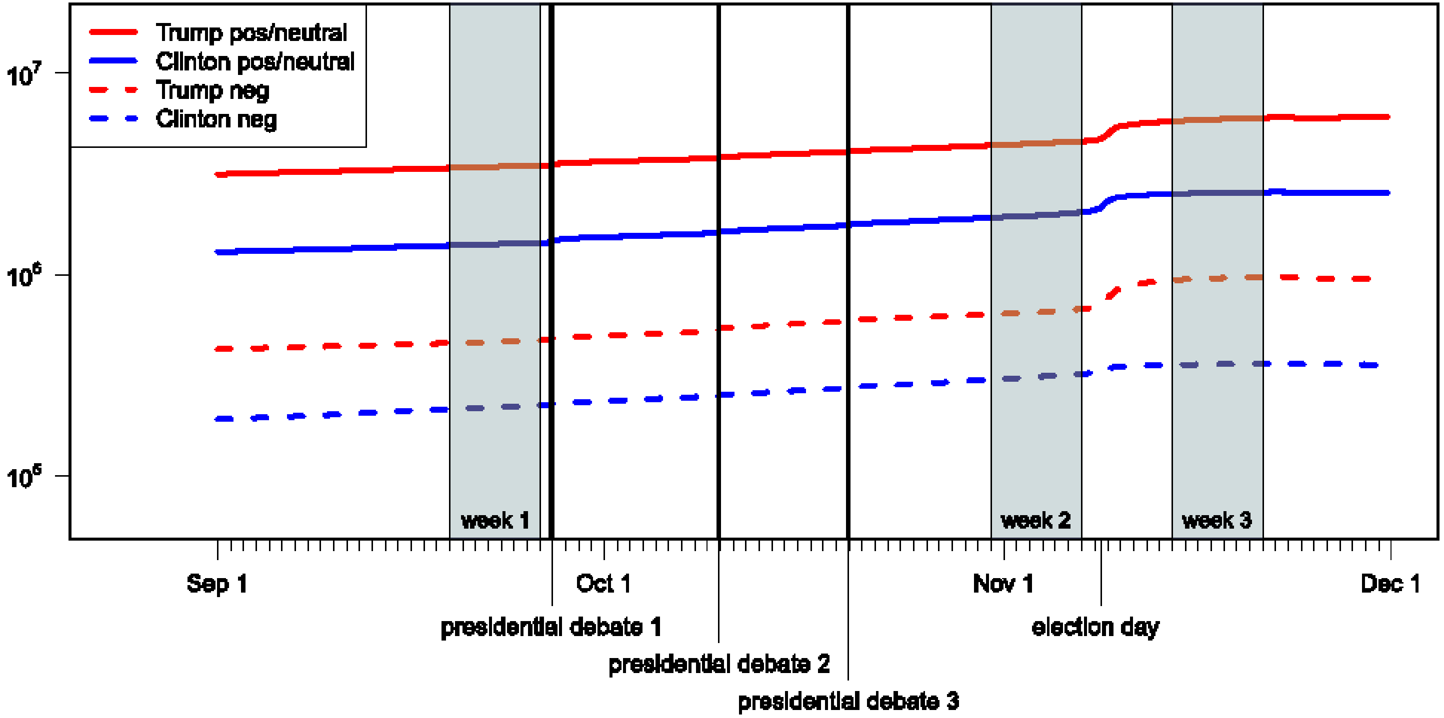
© Robert Bosch GmbH 2018. All rights reserved, also regarding any disposal, exploitation, reproduction, editing, distribution, as well as in the event of applications for industrial property rights.

BOSCH

Interesting Use-Cases
The forecasting power of social media: Example (1)

- The 2016 US presidential election:
  - **Analyze the time of media posting related to Clinton and Trump**.
  - 4 timeseries:
    - Clinton vs. Trump,
    - supporters vs. opponents.
  - Instagram posts:
    - Users are 60% between 18-35
    - 16 hashtags included

Number of media, pos/neutral and negative hashtags (logarithmic scale).



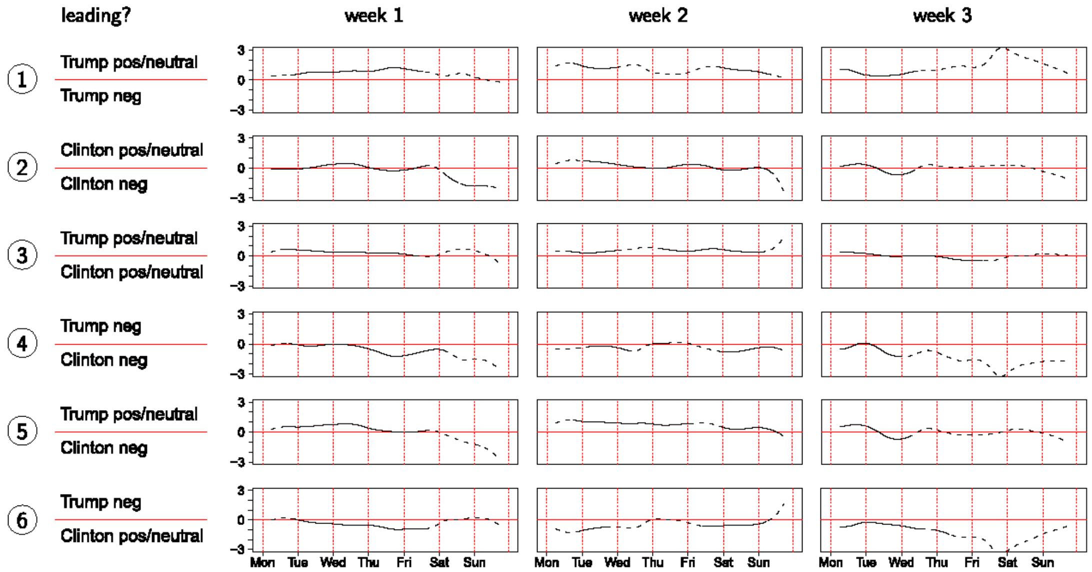
Source: *"The 2016 US presidential election and media on Instagram: Who was in the lead?", Schmidbauer et. al., Computers in Human Behavior, Vol. 81 (2018), pp 148-160*

© Robert Bosch GmbH 2018. All rights reserved, also regarding any disposal, exploitation, reproduction, editing, distribution, as well as in the event of applications for industrial property rights.

Interesting Use-Cases
The forecasting power of social media: Example (2)

- Instagram media postings in favor of, or neutral towards, Trump was ***massively higher*** than any other category.
- Clinton-related postings rose significantly in number, but were still behind on the election day.

Time series of phase differences (in hours) at period 12 h.



➢ ***Lesson learned*: Monitoring media uploads on Instagram and the analysis of the upload behavior could provide a real-time barometer of public opinion and sentiments for policymakers.**

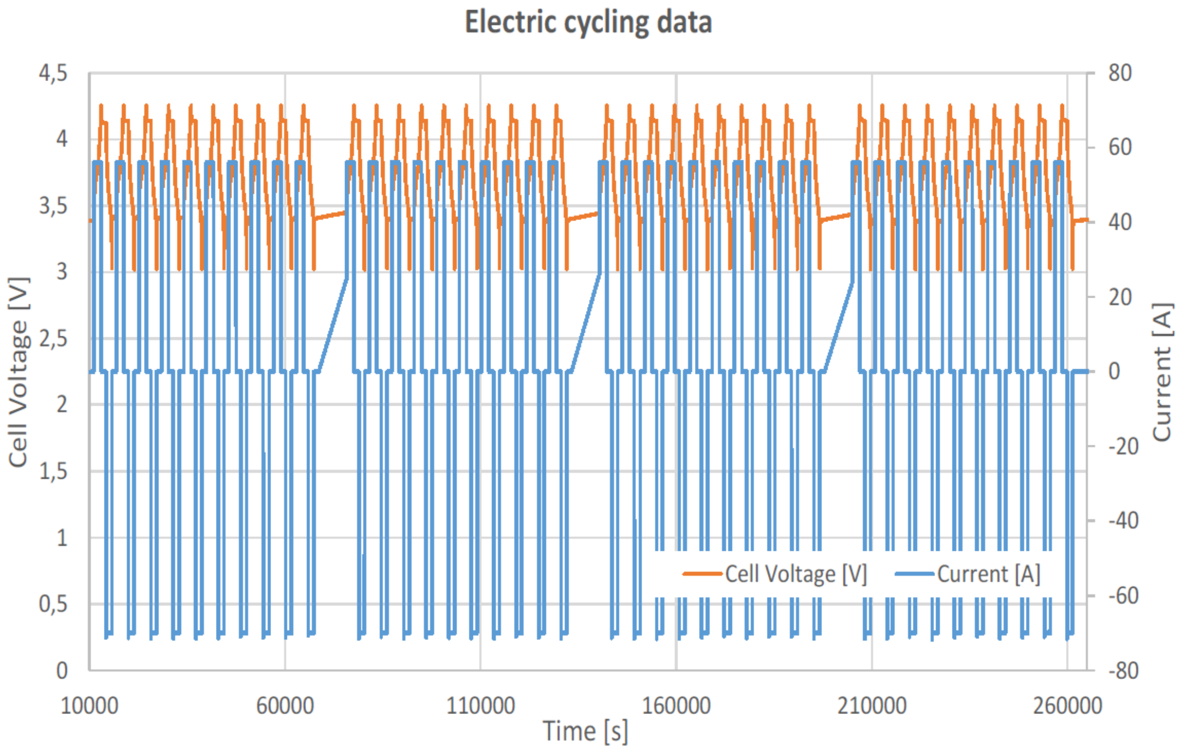
<u>**Source**</u>: *"The 2016 US presidential election and media on Instagram: Who was in the lead?", Schmidbauer et. al., Computers in Human Behavior, Vol. 81 (2018), pp 148-160*

© Robert Bosch GmbH 2018. All rights reserved, also regarding any disposal, exploitation, reproduction, editing, distribution, as well as in the event of applications for industrial property rights.

BATTERY DATA ANALYSIS

Interesting Use-Cases
Battery Data Analysis (1)

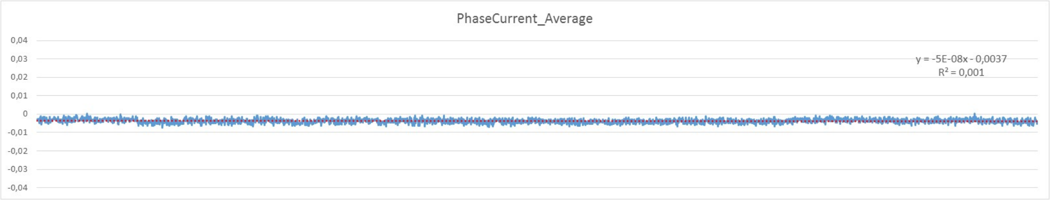
- ***Test data of 6 battery cells***, provided by CR/ARM:
  - Around 1.000 charging/discharging cycles
  - 4.6M data points
  - Data include:
    - Voltage, Current, and
    - Temperature.
  - Data intervals in msec.

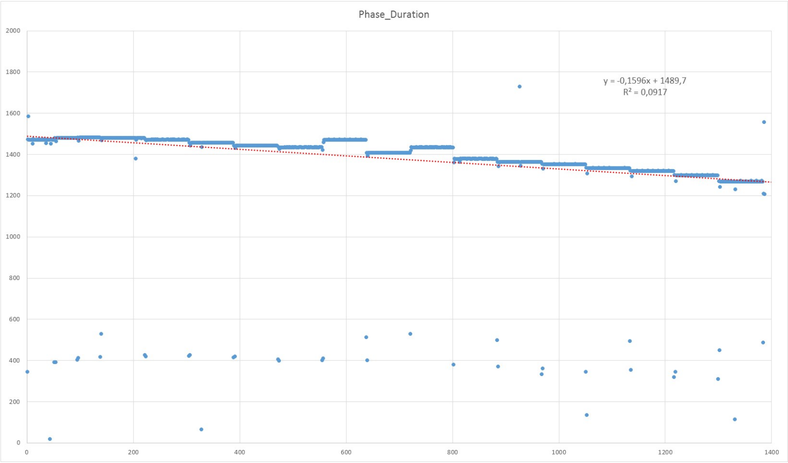


© Robert Bosch GmbH 2018. All rights reserved, also regarding any disposal, exploitation, reproduction, editing, distribution, as well as in the event of applications for industrial property rights.
BOSCH

Interesting Use-Cases
Battery Data Analysis (2)

- Data were:
  - Checked for quality.
  - Outlies were identified (source: measurement faults) and corrected.
  - A custom classification algorithm was created to differentiate between the various operating modes (phases) in the test data:
    - Discharging, Idle, Charging with continues current, Charging with Continues voltage.
  - An electrons-flow timeserie was estimated.



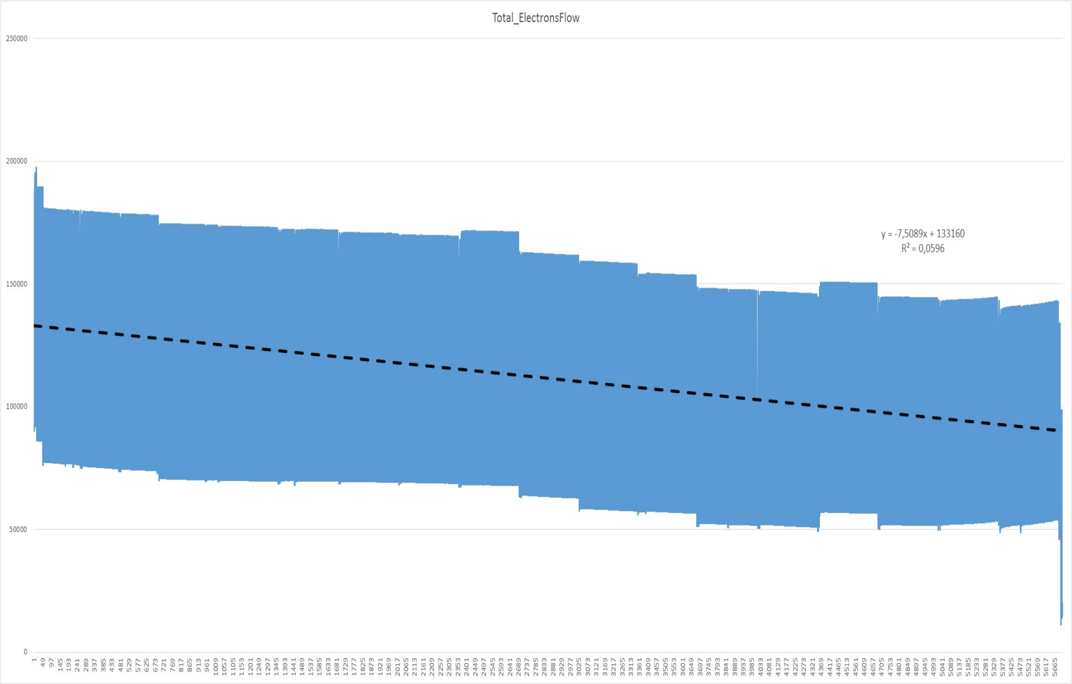


 Internal | CR/AEU2 - Patelis | 2018-03-14
© Robert Bosch GmbH 2018. All rights reserved, also regarding any disposal, exploitation, reproduction, editing, distribution, as well as in the event of applications for industrial property rights.

BOSCH

Interesting Use-Cases
Battery Data Analysis (3)

- ***Goal***: Evaluation and analysis of test data
- Conclusions:
  - Classification algorithms works quite well.
  - Further methods would need to be applied to the classified data in order to eventual identify abnormalities or other non-trivial properties in the data.
  - Electrons-flow timeserie can be used:
    - to detect the behavior of the battery cell, and
    - to forecast when ***total energy will fall below*** a predefined benchmark.



 Internal | CR/AEU2 - Patelis | 2018-03-14
© Robert Bosch GmbH 2018. All rights reserved, also regarding any disposal, exploitation, reproduction, editing, distribution, as well as in the event of applications for industrial property rights.

BOSCH

OTHER USE-CASES

Interesting Use-Cases
Other use-cases (1)

- A large company ***manufacturing*** disposable tableware such as napkins and paper plates.
  - They needed forecasts of each of hundreds of items every month. They know that their products sales have a range of patterns, some with trends, some seasonal, and some with neither.
  - They need to create a software for providing accurate forecasts.

- A large ***car fleet company*** needs to forecast vehicle re-sale values.
  - They purchase new vehicles, lease them out for three years, and then sell them. Better forecasts of vehicle sales values would mean better control of profits.
  - Understanding what affects resale values may allow ***leasing and sales policies*** to be developed in order to maximize profits.

© Robert Bosch GmbH 2018. All rights reserved, also regarding any disposal, exploitation, reproduction, editing, distribution, as well as in the event of applications for industrial property rights.

BOSCH

Interesting Use-Cases
Other use-cases (2)

- A ***European leading airline*** company has weekly traffic data from the previous six years.
  - They require forecasts of passenger numbers for each major route and for each class of passenger (economy class, business class and first class).
  - They need to develop a model for forecasting weekly air passenger traffic on major routes.
  - ***Need to take into account***:
    - Air passenger numbers are affected by school holidays, major sporting events, advertisement campaigns, competition, etc.
    - A major pilots' strike took place one year before, thus there are no traffic data for several months.
    - Also, a new cut-price airline has launched and folded.

© Robert Bosch GmbH 2018. All rights reserved, also regarding any disposal, exploitation, reproduction, editing, distribution, as well as in the event of applications for industrial property rights.
BOSCH

Interesting Use-Cases
Other use-cases (3)

- A ***power company*** wants to remotely control the maintenance of power plants
  - Air pollution measurement data and other sensor data are remotely collected from power plants.
  - They need a system that can analyze the measurements, detect and analyze behavioral patterns and outliers in the timeseries.
  - The cause of each outlier must be identified, in order to create useful information (rule) for the future.
  - The set of identified rules must be transformed into automated data controllers.
  - ***Automated alerts*** must be sent when data are not following the normal pattern and creating abnormal outliers.

© Robert Bosch GmbH 2018. All rights reserved, also regarding any disposal, exploitation, reproduction, editing, distribution, as well as in the event of applications for industrial property rights.

BOSCH

Interesting Use-Cases
Other use-cases (4)

- A major ***advertisement*** company wants to forecast television audience:
  - Needs to create a campaign for a new product on the market. The campaign will run mainly in television.
  - Wants to show their commercial in advertisement slots, which have the biggest television audience of men 25-34 years old.
  - They need to forecast the television viewership (targeting men age 25-34) for each commercial break of each television channel for the next month.
- A ***real estate*** company wants to estimate apartment prices.
  - The company has a big dataset for every apartment, like: geographic location, area, city, address, construction year, surface volume, floor, number of bedrooms, view, orientation, etc.
  - Their scope is to create a software that can automatically estimate the price of all apartments.

© Robert Bosch GmbH 2018. All rights reserved, also regarding any disposal, exploitation, reproduction, editing, distribution, as well as in the event of applications for industrial property rights.

15. PROPOSED LINKS & LITERATURE

Timeseries Forecasting – Data Analytics & Quantitative Methods
Proposed Links & Literature

- R. Hyndman, G. Athanasopoulos (2013): "*Forecasting: Principles and Practice*"
- J. Ord, R. Fildes (2012): "*Principles of Business Forecasting*"
- T. Ruud, S. Babangida (2009): *"On the bias of Croston's forecasting method*", European Journal of Operational Research, Vol.194, pp.177-183.
- J. Hanke, A. Reitsch (2008): "*Business Forecasting*"
- K. Nikolopoulos, P. Goodwin, A. Patelis, V. Assimakopoulos (2007): *"Forecasting with cue information: A comparison of multiple regression with alternative forecasting approaches"*, European Journal of Operational Research, Vol.180, pp.354-368.
- G. Rowe (2007): *"A guide to Delphi",* International Journal of Applied Forecasting 8, 11-16.
- P. Goodwin, R. Fildes, M. Lawrence, K. Nikolopoulos (2007): *"The process of using a forecasting support system"*, International Journal of Forecasting (IJF), Vol.23, pp.391-404.
- K. C. Green, J. S. Armstrong (2007): *"Structured analogies for forecasting"*, International Journal of Forecasting 23(3), 365–376.
- R. Hyndman, A. Koehler (2006): *"Another look at measures of forecast accuracy"*, IJF, Vol.22, pp.679-688.
- E. Gardner (2006): *"Exponential smoothing: The state of the art – Part II"*, IJF, Vol.22, pp.637-666.

© Robert Bosch GmbH 2018. All rights reserved, also regarding any disposal, exploitation, reproduction, editing, distribution, as well as in the event of applications for industrial property rights.
BOSCH

Timeseries Forecasting – Data Analytics & Quantitative Methods
Proposed Links & Literature

- J. De Gooijer, R. Hyndman (2006): *“25 years of time series forecasting”*, IJF, Vol.22, pp.443-473.
- M. Lawrence, P. Goodwin, M. O’Connor, D. Önkal (2006): “*Judgmental forecasting: A review of progress over the last 25 years”,* International Journal of Forecasting 22(3), 493–518.
- Synetos, J. Boylan (2005): *“The accuracy of intermittent demand estimates”*, IJF, Vol.21, pp.303-314.
- Synetos, J. Boylan, J. Croston (2005): *“On the categorization of demand patterns”*, Journal of Operational Research Society, Vol.56, pp.495-503.
- R. Buehler, D. Messervey, D. Griffin (2005), *“Collaborative planning and prediction: Does group discussion affect optimistic biases in time estimation?”*, Organizational Behavior and Human Decision Processes 97(1), 47–63.
- R.G. Brown (2004), *“Smoothing, Forecasting and Prediction of Discrete Time Series”*, Courier Corporation ,Technology & Engineering.
- Hyndman & Billah (2003): *“Theta: SES with drift?”*
- R. Hyndman, A. Koehler, R. Snyder, S. Grose (2002): *“A state space framework for automatic forecasting using exponential smoothing methods”*, IJF, Vol.18, pp.439-454.
- V. Assimakopoulos, K. Nikolopoulos (2000): *“The Theta model: A decomposition approach to forecasting”*, IJF, Vol.16, pp.521-530.
- S. Makridakis, M. Hibon (2000): *“The M3-competition: Results, conclusions and implications”*, IJF, Vol.16 (special issue), pp.451-476.

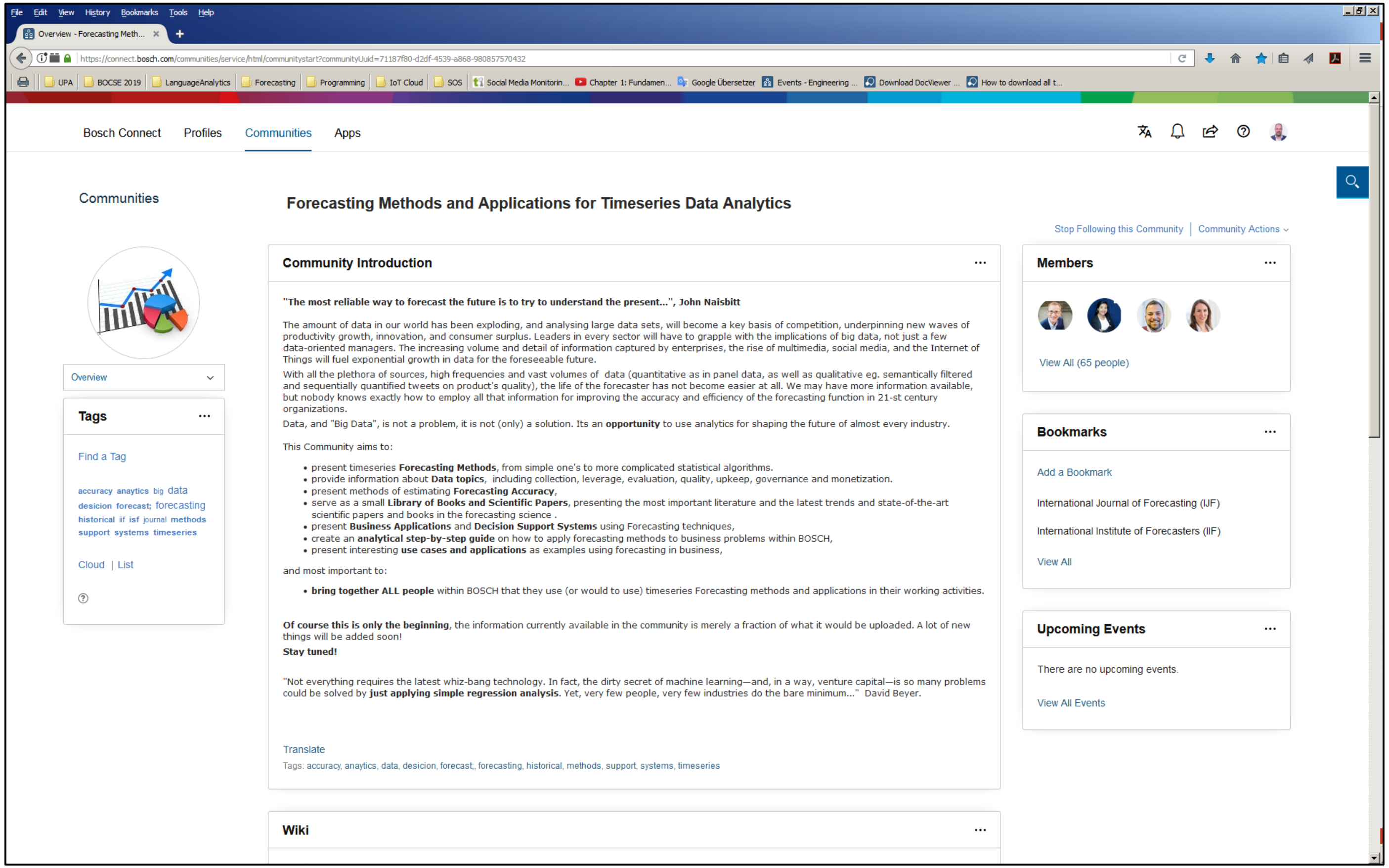
© Robert Bosch GmbH 2018. All rights reserved, also regarding any disposal, exploitation, reproduction, editing, distribution, as well as in the event of applications for industrial property rights.

Timeseries Forecasting – Data Analytics & Quantitative Methods
Proposed Links & Literature

- G. Rowe. G. Wright (1999): *“The Delphi technique as a forecasting tool: issues and analysis”*, IJF, Vol.15, pp.353-375.
- S. Makridakis, S. Wheelwright, R. Hyndman (1998): “*Forecasting: Methods and Applications*”
- Assimakopoulos V. (1995): *“A successive filtering technique for identifying long-term trends”*, IJF Vol14, 35-43.
- J. D. Hamilton (1994): “*Timeseries Analysis*”
- J. Armstrong, F. Collopy (1992): *“Error Measures for generalizing about forecasting methods: Empirical comparisons”*, IJF, Vol.8, pp.69-80.
- S. Makridakis (1990): *“Sliding Simulation: A new approach to time series forecasting”*, Management Science, Vol.36, pp.505-512.
- S. Wheelwright, S. Makridakis (1985): “*Forecasting methods for Management*”
- S. Makridakis, A. Andersen, R. Carbone, R. Fildes, M. Hibon, R. Lewandowski, J. Newton, E. Parzen, R. Winkler (1982): *“The accuracy of extrapolation (time series) methods: Results of a forecasting competition”*, IJF, Vol.1, pp 111-153.
- H. Kahn, A.J. Wiener (1967): “*The use of scenarios*”, Hudson Institute
- http://www.forecasters.org
- http://www.sciencedirect.com
- http://www.elsevier.com

© Robert Bosch GmbH 2018. All rights reserved, also regarding any disposal, exploitation, reproduction, editing, distribution, as well as in the event of applications for industrial property rights.

Timeseries Forecasting – Data Analytics & Quantitative Methods
Bosch Community



© Robert Bosch GmbH 2018. All rights reserved, also regarding any disposal, exploitation, reproduction, editing, distribution, as well as in the event of applications for industrial property rights.
BOSCH

THANK YOU

BOSCH