SARTORIUS

Simplifying Progress

Batch Process Modeling – Step-By-Step Guide

Lennart Eriksson, Ph.D., Assoc. Prof.
Senior Lecturer and Principal Data Scientist

Born in Data Analytics

- Company founded in 1987 by Professor Svante Wold, in Umeå, Sweden
  - Originator of Chemometrics and the SIMCA® Methodology
- Patented technologies in Design of Experiments and Multivariate Data Analysis

- We help our customers bring high-quality products to market faster
- Part of Sartorius Stedim Biotech since April 2017
- Products like MODDE®, SIMCA® and SIMCA®-online
- Global strength with local presence

Business Growth Through the Entire Product Lifecycle

Active Dashboard
Interactive performance insight

Control Advisor
Avoid problems before they arise

MODDE®
Get it right from the start

SIMCA®
Turn data into growth

SIMCA®-online
Ensured manufacturing success

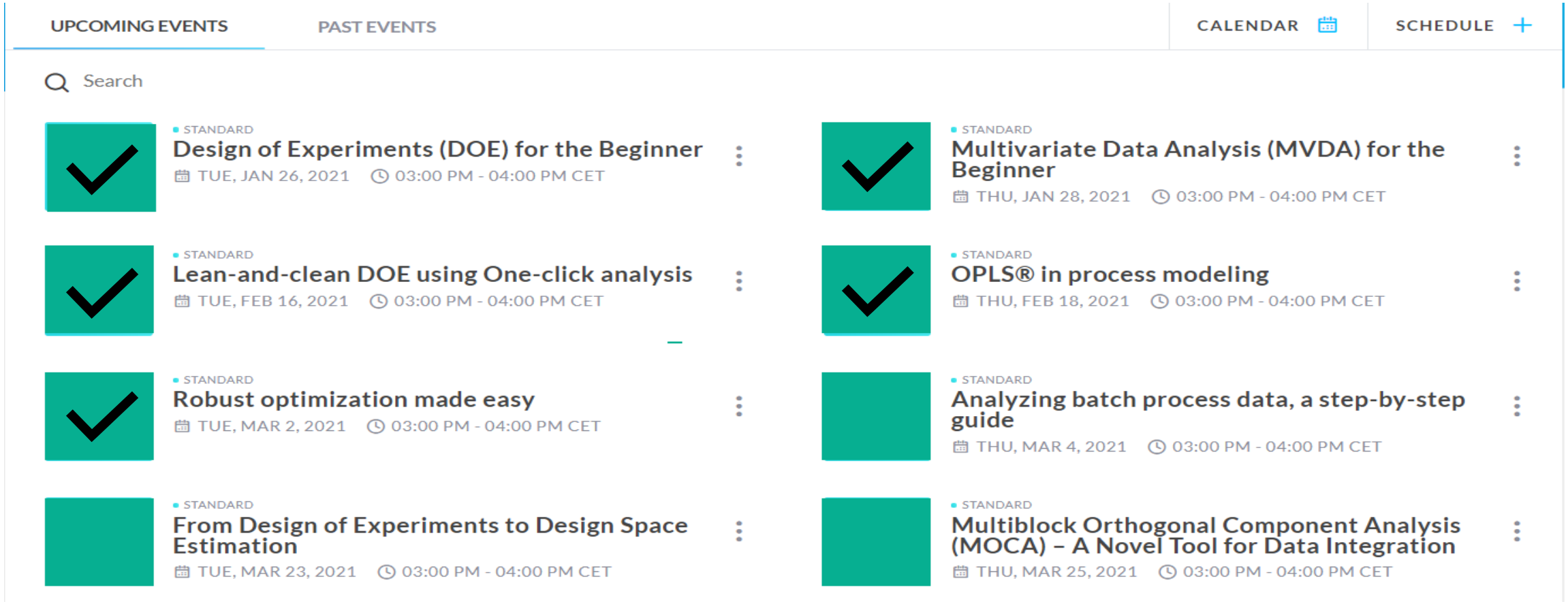
Umetrics® Suite

Education/training and consulting
Transferring expertise to you

Umetrics® Suite of Data Analytics Solution
MODDE® Design of Experiments Solution SIMCA® Multivariate Data Analysis Solution SIMCA®-online Solution

SARTORIUS

Upcoming Webinars (https://www.sartorius.com/en/company/exhibition-conferences)

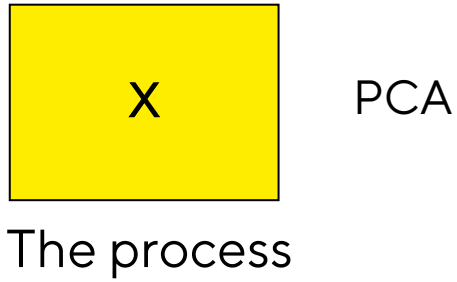


Contents

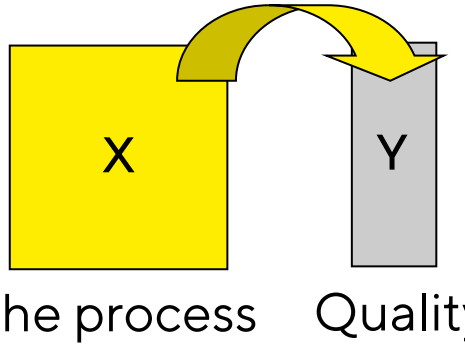
- Multivariate Process data
  - Different blocks of data and methods for their analysis
  - Statistical Process Control (SPC); SPC extensions: MSPC & BSPC
- Data lay-out and data sources
  - Define process and data sources
  - Database import possibilities
  - Process measurements and batch conditions
  - Two model perspectives, BEM & BLM
- Modeling and monitoring a fermentation process producing baker’s yeast
  - Batch control charts
  - Early fault detection and interpretation of deviations
  - Model validation
- Modeling a cell culture process
  - Using the BEM to establish NOC
  - Using the BLM to predict final titer
- Demo & Summary

Multivariate Process Data

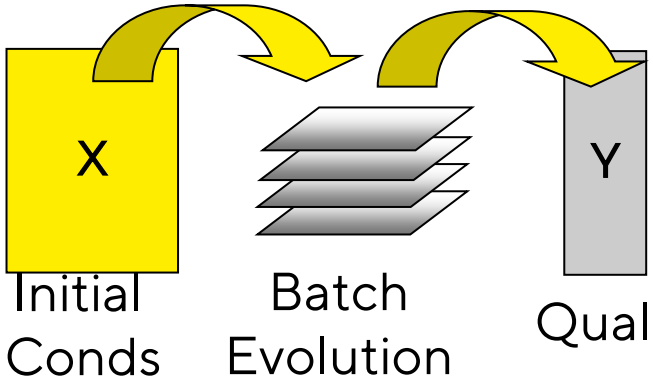
- Monitoring a process
  - Early warning of disturbances
  - Diagnostics - finding "assignable causes"
- Modelling a process output
  - Monitor Quality of final product
- Modelling Batch Processes
  - Majority of industrial processes
  - More complex analysis



PCA



PLS
OPLS



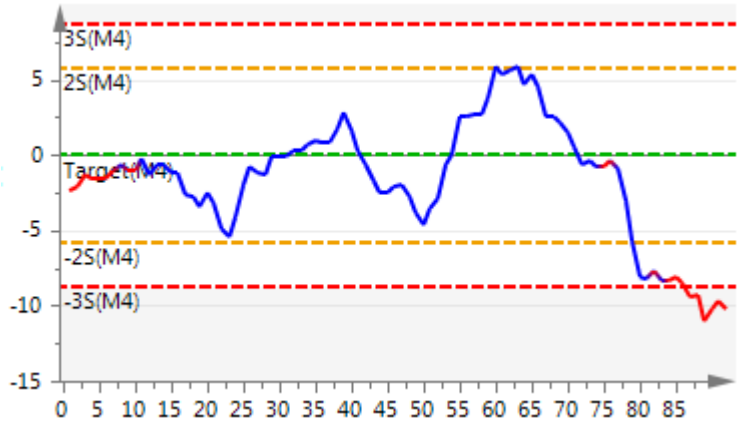
PCA
PLS
OPLS

SARTORIUS

Statistical Process Control, SPC

Walter A. Shewhart ~1930

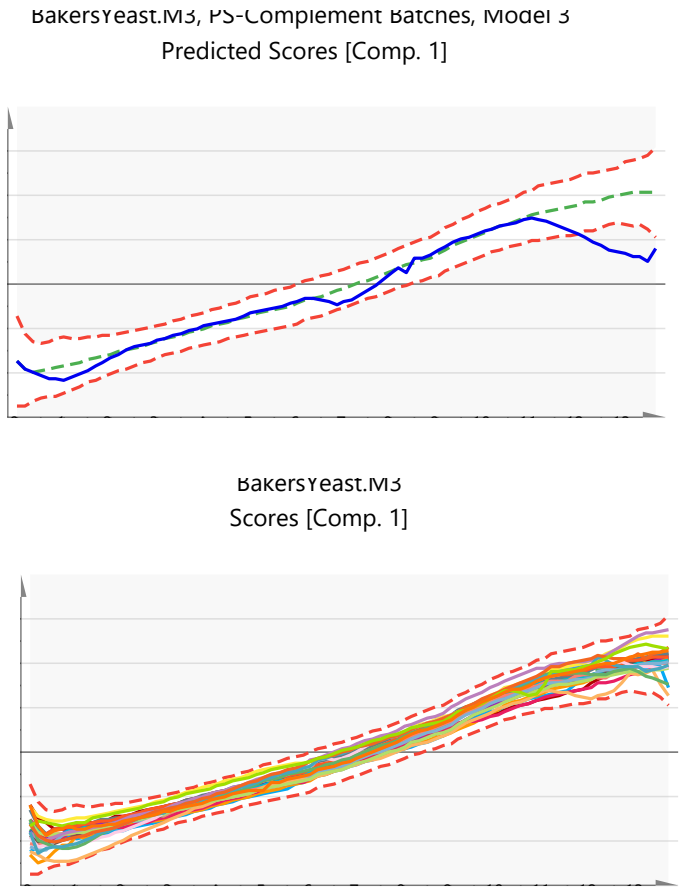
- Assumption
  - A process is in a state of ”statistical control” unless a special event occurs
- Approach
  - Devise a test to detect the occurrence of any special event
- Response
  - Look for ”assignable cause” for the special event. Correct the process back to target
- Result
  - Robust process performance, long term process improvement



Using data analytics, the transition from SPC to multivariate SPC (MSPC) is obvious

Batch statistical process control, BSPC

- The same philosophy is the basis for Batch Multivariate SPC
- For applications in Batch processes the evolution of the batch must be taken into account
- A process “road” for normal evolution is constructed
  - A reliable monitoring model should be able to detect when the process does not evolve in a normal way
- Any deviations are considered as abnormal process events
  - To be analyzed and acted upon

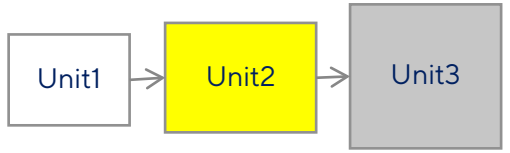


Data Lay-Out and Data Sources

SARTORIUS

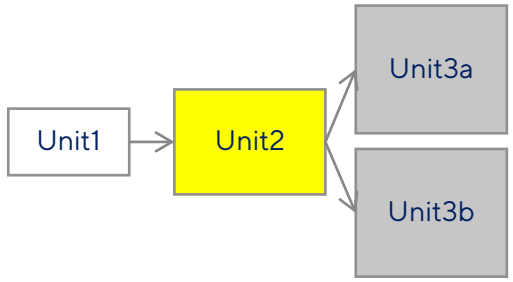
Define Process

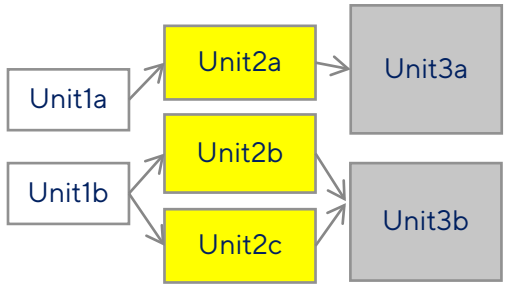
- What does the process look like?
  - How many phases?
    - Process steps/ units
  - Split/ Merge batches



- Which parts of the process should be modeled?

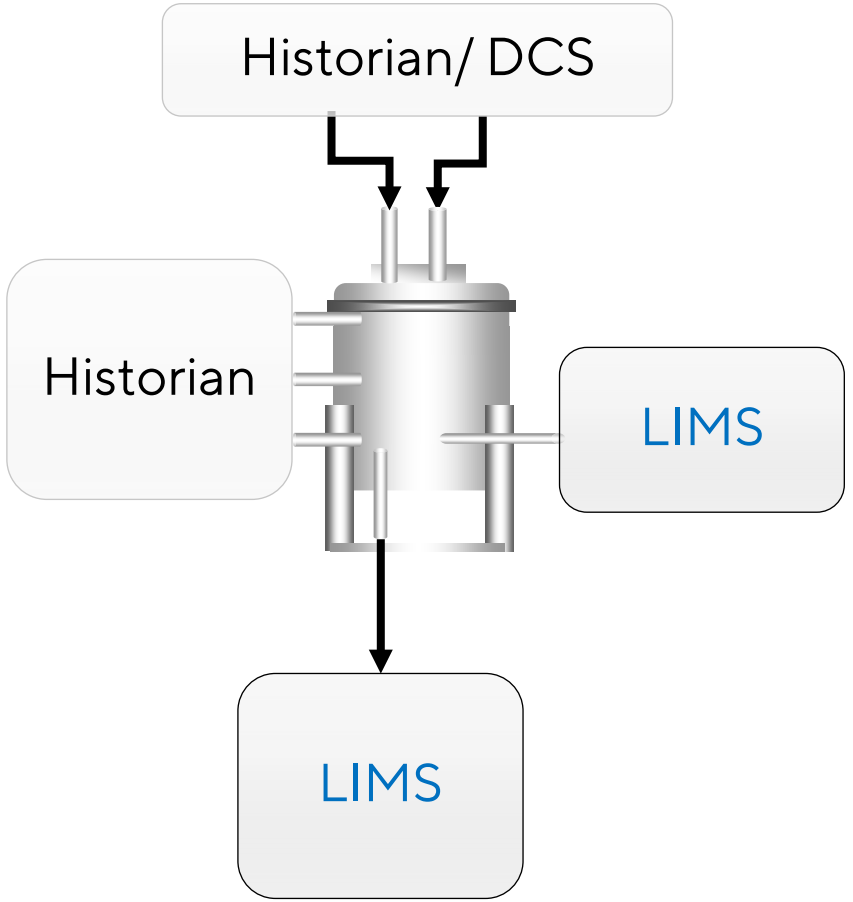
- Which data is relevant?
  - Process tags
  - Quality information





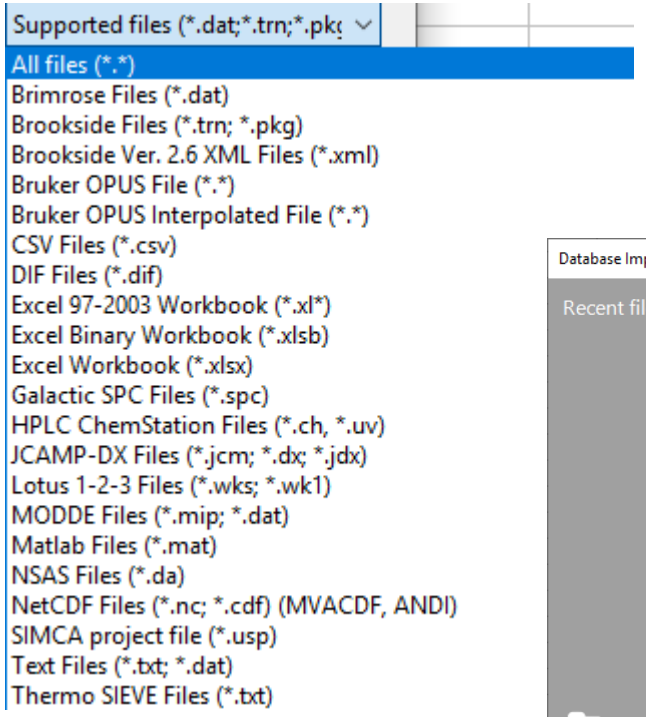
Available Data Sources

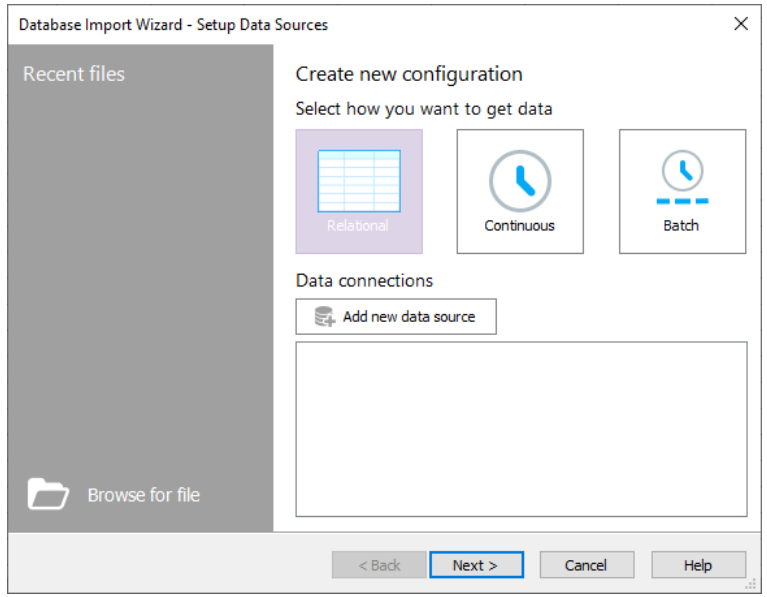
- Different types of data
  - Start-point information
    - Set-points pH, temp
  - Process parameters
    - Controlled parameters
    - Controlling parameters (actuators)
  - IPC / At-line data (ex. daily measurements)
  - Quality attributes
- Data stored in different places
  - LIMS, Historian…
- Depending on objective different data is used
- For analysis they need to be correctly synchronized



Process Database Data Extraction

- Export from DB to SIMCA supported file format
  - Most databases support data extraction to Excel
  - Data must be organized correctly
  - Data may be distributed over several files or excel sheets
- SIMCA can connect directly to a process data base
  - SIMCA import makes sure data is presented as required by SIMCA
  - Requires SimApi compatible with DB





SimApi

	Aspen IP21	CSV	Discovant	June 5	MFCS	ODBC	OPC HDA	OPC UA	OSIsoft PI	SIPAT	SynTQ	Wonderware
Current	O	O	O	O	O	O	O	O	O	O	O	O
Historical	O		O	O	O	O	O	O	O	O	O	O
Discrete			O			O						
Batch	O		O		O	O			O	O		
Write back	O					O	O	O	O			O
Write back discrete						O						
Write back batch						O			O			
Node hierarchy	O						O	O	O			
Array tag expansion	O						O					
Multiple data sources	O			O	O	O		O	O			
Connection resiliency		O		O	O	O		O	O			
3rd party developed				O	O			O		O	O	

https://www.sartorius.com/en/products/process-analytical-technology/data-analytics-software/mvda-software/simapi

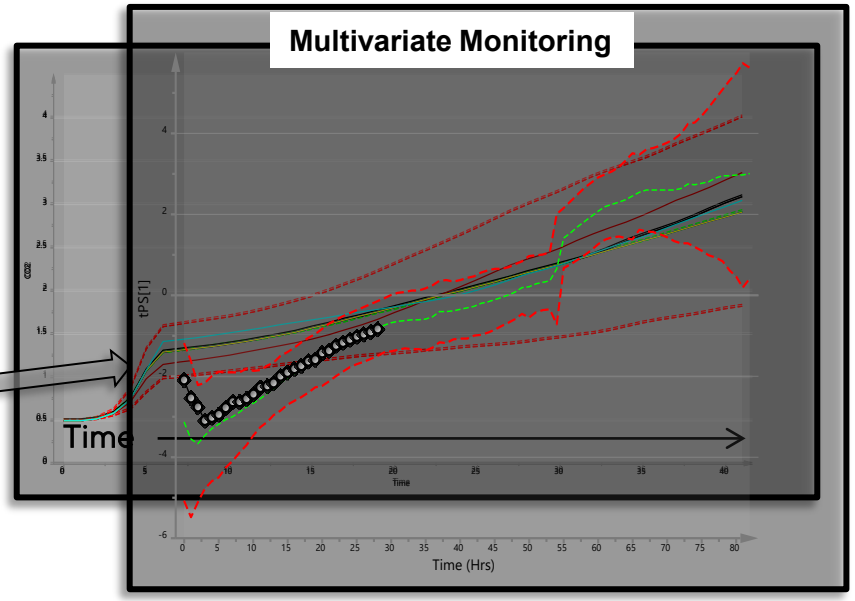
Batch Data

Process Data

Variables →

Time ↓

BatchID	TimeHrs	CL	V	CO2	T	F	Fc	Fb	fg	Ssp	Tsp	pH
1	0	1.16	100	0.5	297	0	10	0	2	0.02	320	5.1
2	0	1.16	100	0.5	297	0	10	0	2	0.02	320	5.1
3	0	1.16	100	0.5	297	0	10	0	2	0.026	330	5.1
4	0	1.16	100	0.5	297	0	10	0	2	0.02	320	5.1
5	0	1.16	100	0.5	297	0	10	0	2	0.02	310	5.1
4	0.083333	1.15695	99.9952	0.50013	296.998	0	10	0	2	0.02	310	5.0963
4	10.0833	1.15172	99.8371	0.52552	296.954	0	2	0.076967	1.86708	0.02	310	4.99307
4	20.0833	1.13769	100.56	0.59000	297.024	0	2	0.201543	1.75939	0.02	310	4.98243
4	30.0833	1.10362	103.303	0.74793	297.381	0	2	0.506962	1.71858	0.02	310	4.95818
4	40.0833	1.03992	110.613	1.08858	298.377	0	2	1.13705	1.81533	0.02	310	4.92515
4	50.0833	1.11777	119.214	1.41535	300.353	0.032947	2	0.002955	1.88646	0.02	310	5.12493
4	60.0833	1.11907	118.991	1.43867	302.636	0.032809	2	0.047028	1.83693	0.02	310	4.99949
4	70.0833	1.11749	118.719	1.4618	304.982	0.033705	2	5.00E-02	1.79198	0.02	310	4.9997
4	80.0833	1.11582	118.314	1.48737	307.398	0.034648	2	0.05304	1.75115	0.02	310	4.99969
4	90.0833	1.11405	117.753	1.51575	309.894	0.035639	4.19453	0.056201	1.71465	0.02	310	4.99968
4	100.083	1.11234	117.039	1.54689	311.286	0.036476	554.565	0.057968	1.68254	0.02	310	4.99996
4	110.083	1.11088	116.311	1.57862	310.745	0.036956	1203.54	0.056805	1.65427	0.02	310	5.00019
4	120.083	1.10958	115.672	1.60909	309.532	0.037288	1213.61	0.054588	1.62934	0.02	310	5.00021
4	130.083	1.10827	115.113	1.63833	309.03	0.037722	689.387	0.053342	1.60766	0.02	310	5.00005
4	140.083	1.10683	114.548	1.66775	309.857	0.038383	347.719	0.053973	1.58946	0.02	310	4.99988
4	150.083	1.10537	113.913	1.69863	310.815	0.039089	646.424	0.054926	1.57495	0.02	310	4.99996
4	160.083	1.10406	113.259	1.7302	310.684	0.03961	1143.03	0.054368	1.56384	0.02	310	5.00011
4	170.083	1.10292	112.648	1.76105	309.889	0.040002	1278.96	0.052841	1.5557	0.02	310	5.00016
4	180.083	1.10182	112.104	1.79088	309.372	0.040425	992.11	0.051669	1.55031	0.02	310	5.00008
4	190.083	1.10069	111.575	1.82053	309.696	0.040982	689.535	0.051634	1.54772	0.02	310	4.99996
4	200.083	1.09953	111.01	1.85108	310.412	0.041612	769.918	0.052185	1.54803	0.02	310	4.99996
4	210.083	1.09954	110.14	1.88053	313.938	0.040985	2	0.039404	1.54806	0.014	325	4.99976
4	220.083	1.09819	108.902	1.91548	317.738	0.041787	2	0.042336	1.55154	0.014	325	4.99972
4	230.083	1.09675	107.235	1.95927	321.692	0.042535	2	0.04464	1.55856	0.014	325	4.99982
4	240.083	1.09532	105.019	2.01075	325.734	0.043048	100.68	0.044996	1.5692	0.014	325	5.0001
4	250.083	1.09397	102.324	2.07725	327.348	0.043261	1293.94	0.04297	1.58325	0.014	325	5.00017
4	260.083	1.09258	99.6942	2.13522	326.148	0.043594	2373.11	0.042058	1.60076	0.014	325	5.00007
4	270.083	1.09142	97.459	2.19425	324.076	0.043885	2308.41	0.040861	1.62139	0.014	325	5.00015
4	280.083	1.09044	95.5296	2.24973	323.26	0.044152	1354.8	0.039537	1.64456	0.014	325	5.0001
4	290.083	1.08944	93.6009	2.30695	324.689	0.044486	725.846	0.038866	1.67018	0.014	325	5.00007
4	300.083	1.08843	91.4383	2.37075	326.462	0.04474	1237.02	0.037657	1.69834	0.014	325	5.00015
4	310.083	1.0874	89.1914	2.43876	326.386	0.044984	2187.11	0.036467	1.72907	0.014	325	5.00009
4	320.083	1.08647	87.1459	2.50508	324.973	0.04525	2541.49	0.035535	1.76223	0.014	325	5.0001
4	330.083	1.0857	85.3565	2.56774	323.883	0.045487	2077.8	0.03442	1.79742	0.014	325	5.00011
4	340.083	1.08498	83.6631	2.63005	324.312	0.045751	1478.67	0.0336	1.83442	0.014	325	5.00007
4	350.083	1.08423	81.8756	2.69693	325.672	0.046001	1551.6	0.032764	1.87326	0.014	325	5.0001
4	360.083	1.08346	79.9893	2.76893	326.198	0.046213	2214.82	0.031715	1.91403	0.014	325	5.0001
4	370.083	1.08272	78.1792	2.84148	325.446	0.046439	2711.22	0.030852	1.95669	0.014	325	5.00008
4	380.083	1.08209	76.5496	2.91124	3.24E+02	0.046653	2598.47	0.029954	2.00101	0.014	325	5.00009
4	390.083	1.08154	75.0396	2.97952	324.351	0.046873	2170.21	0.02915	2.04674	0.014	325	5.00007
4	400.083	1.08097	73.515	3.05056	325.208	0.047093	2062.9	0.028439	2.09389	0.014	325	5.00008

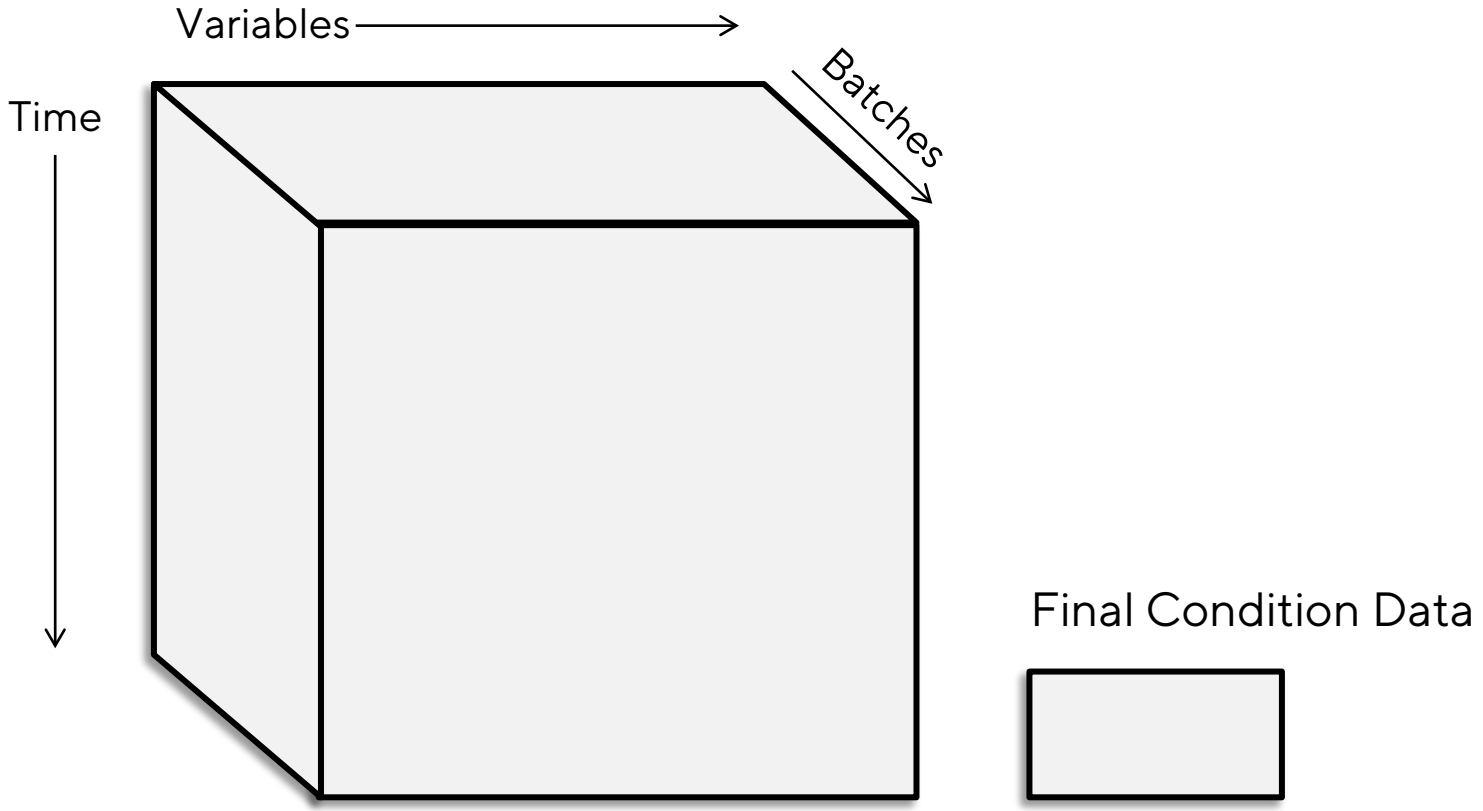


Final Condition Data

Batch ID	[Pen]	Amount	Volume
1	2.305952	164.535	71.3523
2	2.140447	164.238	76.7307
3	2.397035	146.426	61.0863
4	0.467192	76.7307	164.238
5	0.418107	73.515	175.828

Batch Data

- The challenge is to model the relationship between process variation and quality (final conditions)



Process Data and Batch Conditions

- Process data describes the process evolution
  - Timely measurements
  - All parameters are extracted at the same time frequency

- Batch conditions does not have an evolution
  - The same value true for the entire batch
  - Starting conditions, intermediate quality, final batch properties

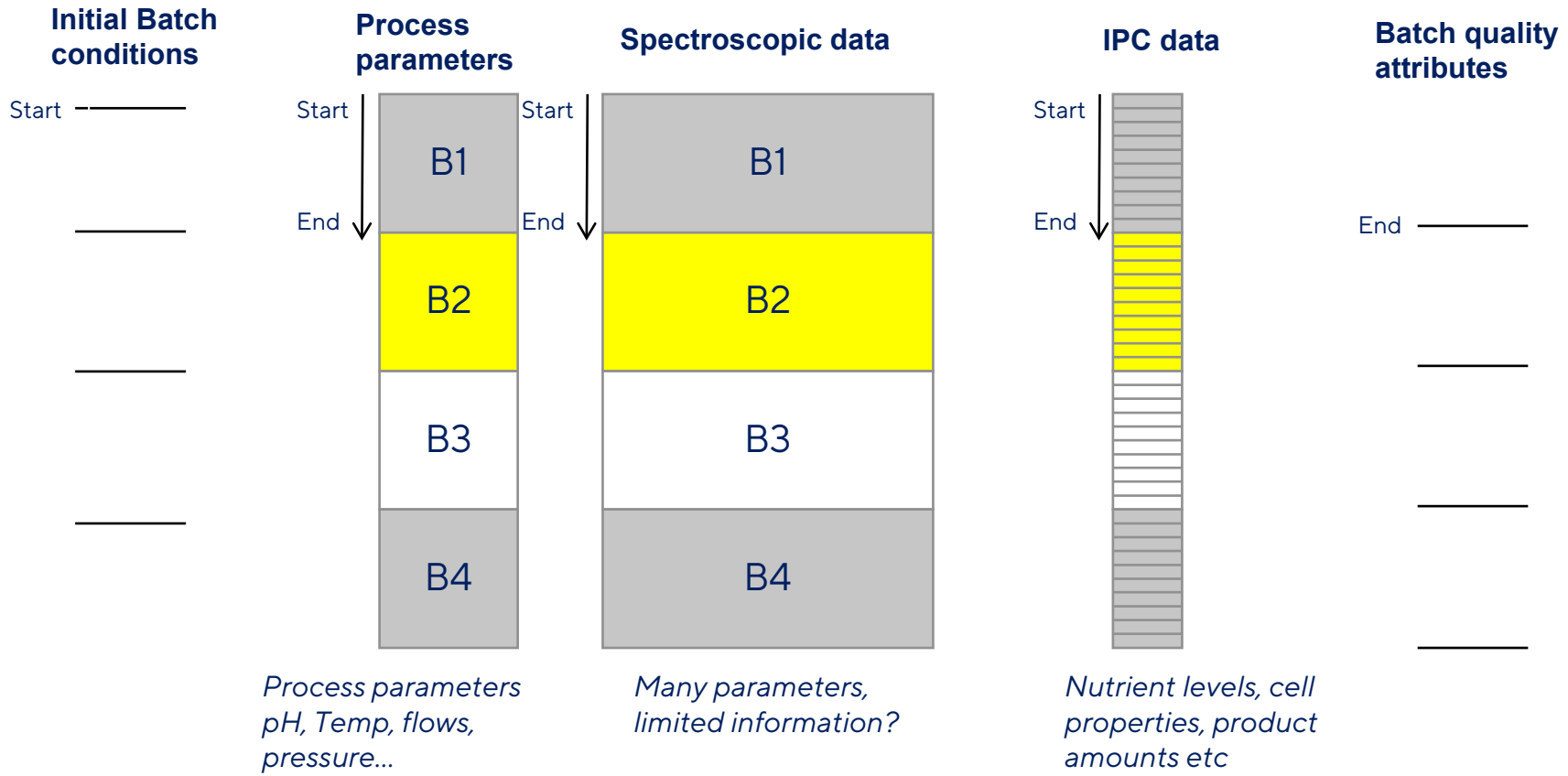
Dataset - BAKERS YEAST

	2	3	4	5	6	7	8	9	10
1	$BatchID	Batch Age	Ethanol	Temp	Molasses	NH3	Air	Level	pH
153	Ba	11,1667	0,0577	30,28	3252,92	98,5	6704,	54,00	5,375
154	Ba	11,3333	0,0528	30,28	3263,22	98,5	6693,	54,24	5,380
155	Ba	11,5	0,0497	30,31	3268,86	98,3	6718,	54,53	5,376
156	Ba	11,6667	0,0469	30,35	3273,44	98,2	6704,	54,89	5,369
157	Ba	11,8333	0,0444	30,38	3279,19	98,0	6689,	55,13	5,363
158	Ba	12	0,0419	30,32	3279,62	10,0	6700,	55,47	5,376
159	Ba	12,1667	0,035	30,29	3270,92	0,10	6700,	55,72	5,445
160	Ba	12,3333	0,0303	30,23	3250,12	0,12	6709,	55,91	5,541
161	Ba	12,5	0,0270	30,21	3223,85	0,10	6695,	56,21	5,650
162	Ba	12,6667	0,0244	30,21	3202,48	0,11	6701,	56,45	5,764
163	Ba	12,8333	0,0223	30,25	3174,19	0,10	6697,	56,72	5,880
164	Ba	13	0,0203	30,28	3122,73	0,12	6702,	57,06	5,994
165	Ba	13,1667	0,0184	30,28	2679,03	0,11	6692,	57,25	6,123
166	Ba	13,3333	0,0165	30,14	2014,34	0,09	6701,	57,41	6,277
167	Ba	13,5	0,0148	30,06	1389,3	0,10	6706,	57,70	6,447
168	Ba	13,6667	0,0134	29,90	859,207	0,12	6718,	57,58	6,591
169	Ca	0	0,0662	29,82	667,187	47,4	2306,	41,23	3,913
170	Ca	0,166667	0,1449	29,69	997,483	73,4	2496,	41,30	4,024
171	Ca	0,333333	0,3917	29,61	1165,18	86,2	2672,	41,39	4,466
172	Ca	0,5	0,7348	29,55	1087,61	80,4	2857,	41,48	4,806
173	Ca	0,666667	1,0436	29,46	995,638	73,8	3033,	41,56	5,001
174	Ca	0,833333	1,2598	29,52	929,693	68,3	3204,	41,63	5,079
175	Ca	1	1,3986	29,53	892,962	66,2	3366,	41,71	4,975
176	Ca	1,16667	1,5052	29,44	871,689	63,8	3533,	41,78	5,078
177	Ca	1,33333	1,5369	29,53	883,85	65,6	3700,	41,85	4,989
178	Ca	1,5	1,5637	29,55	905,478	65,8	3870,	41,93	5,026
179	Ca	1,66667	1,5785	29,42	942,776	72,3	4029,	42,00	5,072
180	Ca	1,83333	1,5812	29,49	974,933	74,9	4195,	42,08	5,080
181	Ca	2	1,5930	29,57	1006,21	78,5	4362,	42,17	5,035
182	Ca	2,16667	1,5873	29,51	1051,41	82,8	4527,	42,24	5,091
183	Ca	2,33333	1,5810	29,50	1088,24	84,3	4690,	42,34	5,066
184	Ca	2,5	1,5668	29,45	1141	88,9	4857,	42,42	5,08
185	Ca	2,66667	1,5518	29,45	1192,38	93,0	5026,	42,53	5,091
186	Ca	2,83333	1,5296	29,50	1263,22	99,2	5195,	42,63	5,060
187	Ca	3	1,5114	29,55	1324,52	103,	5359,	42,73	5,105

Dataset - Bakers Yeast1

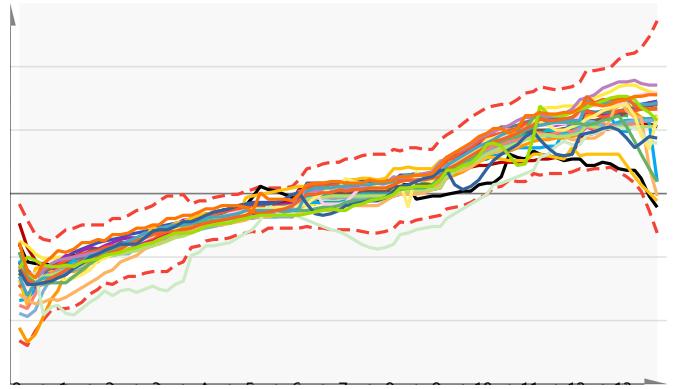
	2	3	4	5	6	7
1	$BatchID	Innoc	QP1	QP2	Amount	Yield
2	bb	914	93	77	5365	0,482205
3	Ga	932	86	77	6089	0,48
4	Ha	940	81	71	5904	0,46
5	hb	943	88	82	5977	0,470986
6	gb	950	89	78	5982	0,47437
7	cb	952	91	83	5875	0,500618
8	db	952	91	82	5835	0,505364
9	eb	952	90	80	5973	0,4816
10	Za	952,32	86	77	5541	0,45427
11	fb	960	81	72	5402	0,46254
12	Ya	964	70	63	4689	0,45
13	ab	967	85	81	5749	0,492269
14	ib	973	91	80	5977	0,497475
15	jb	973	91	79	6147	0,480445
16	kb	980	88	82	6154	0,485785
17	lb	981	93	83	5919	0,481485
18	Xa	993,28	84	68	6597	0,504218
19	Ua	1027	94	91	6427	0,46
20	Aa	1049	92	93	5715	0,42
21	Ta	1049,04			6960	0,505668
22	Ma	1050,56	89	82	6658	0,479405
23	Pa	1050,56	89	82		0,556606
24	La	1051	83	76	6515	0,55
25	Qa	1051	94	89	4442	0,44

Data Overview

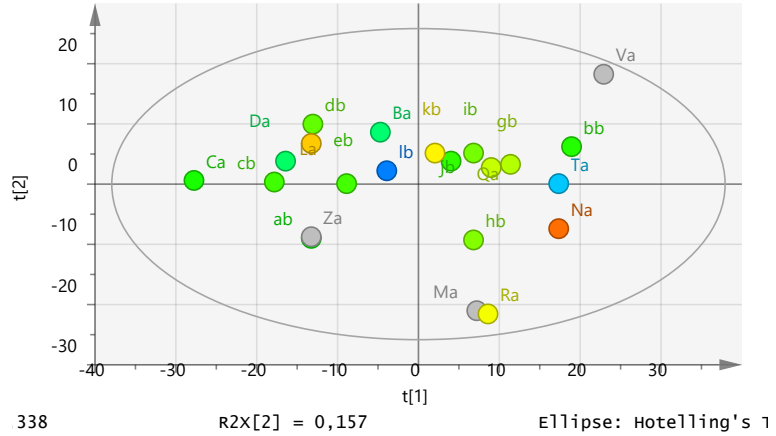


BSPC - From Two Perspectives

- Batch processes evolve over time
  - Many measurements over time
  - Batch Evolution Model (BEM)



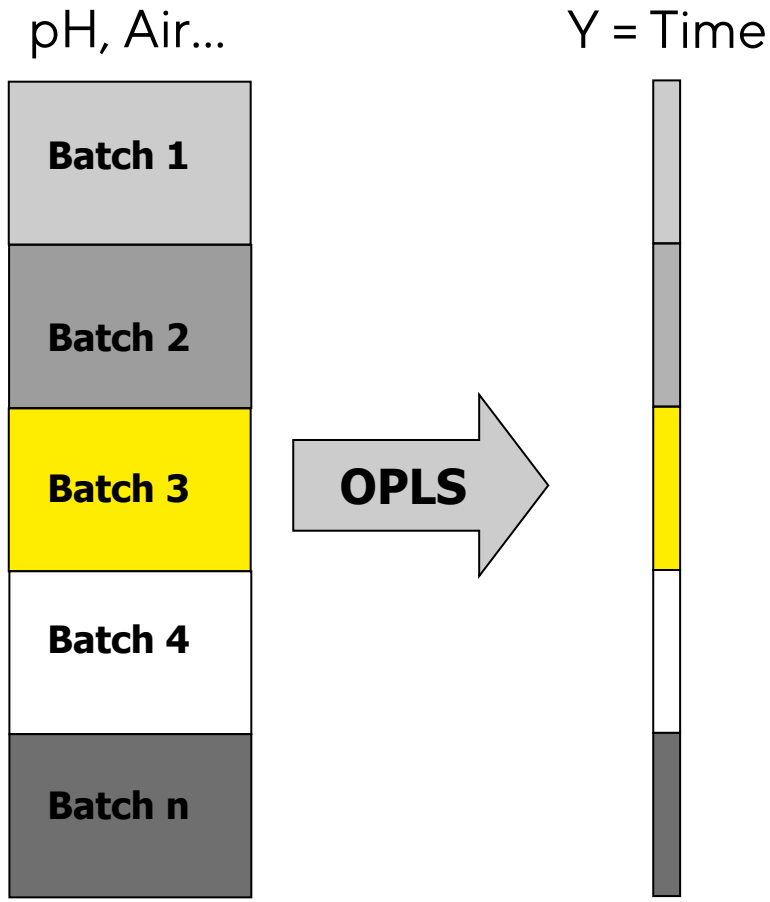
- Quality, yield etc are summarized per batch
  - One measurement per batch
  - Batch Level Model (BLM)



The Batch Evolution Model (BEM)

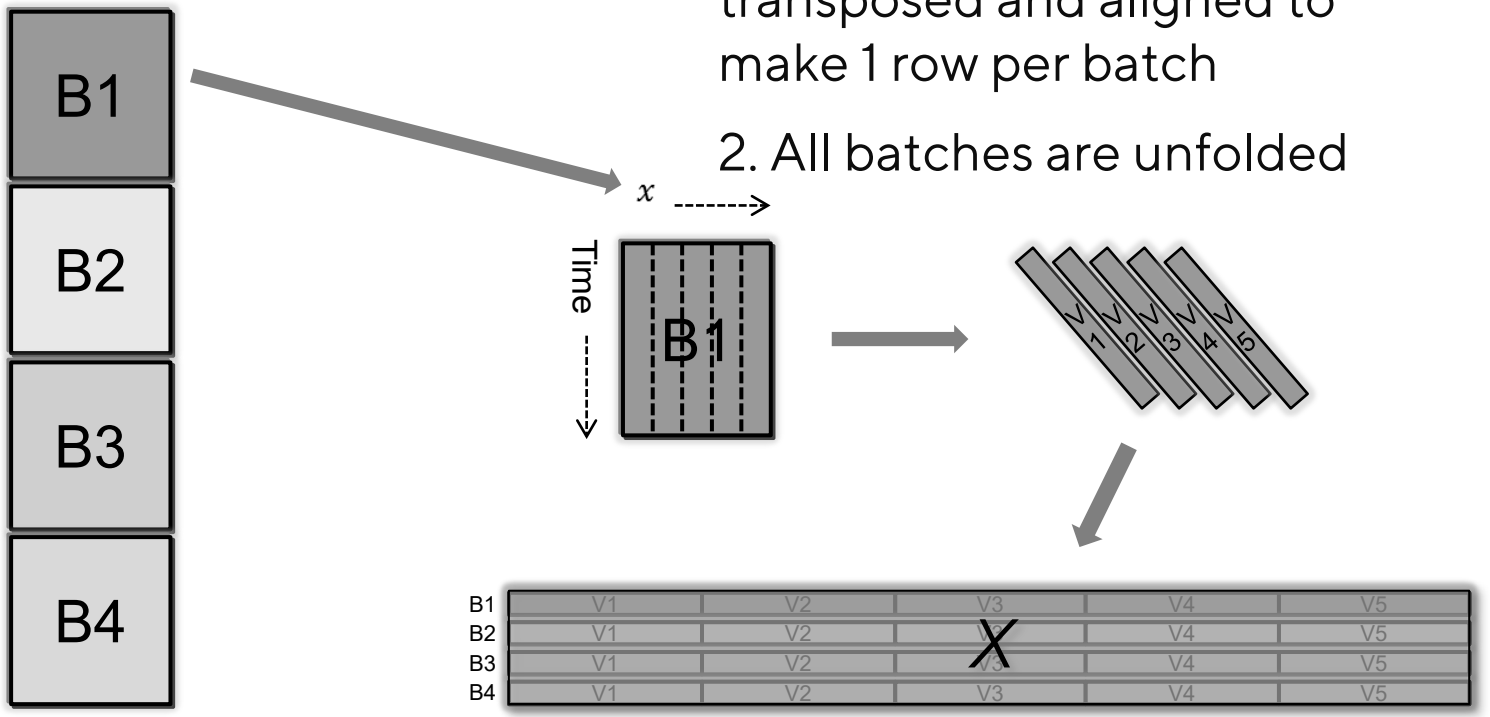
- Time (or maturity) is used as a Y-variable to give the model a direction

- Maturity need not be time. It could be say, for example, be Ethanol in beer brewing



Batch Data; BEM and BLM

- Batch data has time dependency
  - A table of data is generated for each batch
  - Variables measured over time



Batch-wise Unfolding

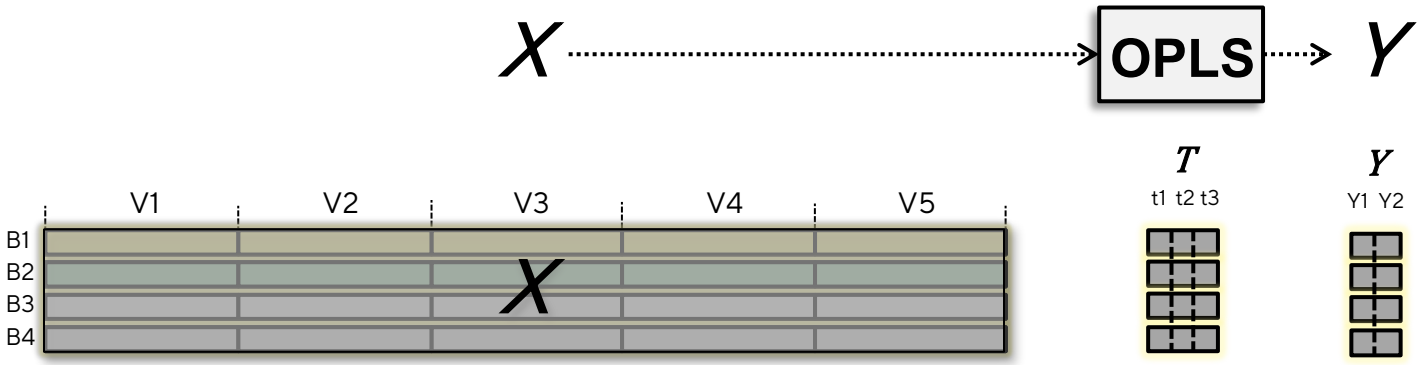
1. Each variable is transposed and aligned to make 1 row per batch

2. All batches are unfolded

SARTORIUS

The Batch Level Model (BLM)

- OPLS can handle the highly correlated batch unfolded dataset



OPLS is used to map the correlation of the process trajectory (X) to final batch conditions (Y)

Example: Baker’s Yeast (BEM)

Modelling of a Batch Process

- Example: Baker’s yeast production
  - Dataset from Jästbolaget AB in Sweden

- 33 batches represented by 7 process variables and 5 batch conditions
  - Batch conditions: X parameter Innoculum and Y parameters QP1, QP2, amount and yield

- Objective: Establish model for normal process evolution
  - Same approach as for a continuous process
  - Model is built on well performing and behaving batches

Data Structure

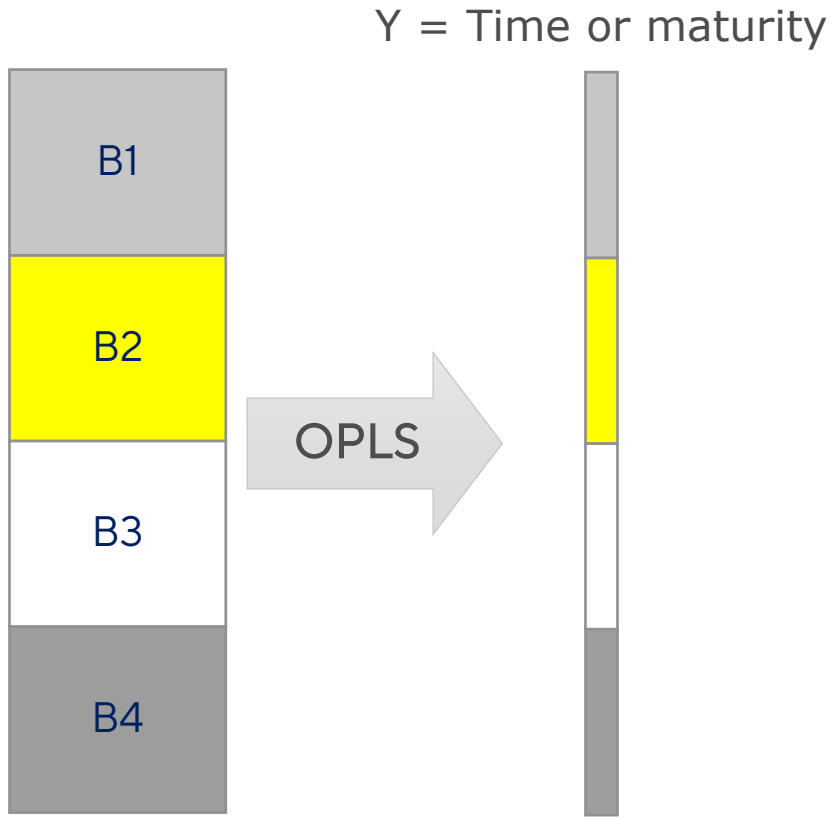
- Data are usually divided into 2 parts
  - Process data (from process DB) and Batch condition data (from LIMS/ lab DB)
- Process/ Evolution data table
  - Each row is one time point
  - One column containing batch ID
  - One column containing time
  - One column for each measured variable
- Batch condition data table
  - One row for each batch
  - Includes both initial (start) conditions and final batch results

BatchID	Innoc	QP1	QP2	Amount	Yield
Ba					
Ca					
Ia	1126,44	82	74	6368	0,493462
Ma	1050,56	89	82	6658	0,479405

BatchID	Batch Age (h)	Ethanol	Temp	Molasses	NH3	Air	Level	pH
	TAGID Age001	TAGID E001	TAGID T001	TAGID FM001	TAGID FN001	TAGID FA001	TAGID L001	TAGID P001
Ba	0	0,03748	35,7145	478,196	23,7236	1842,99	41,738	4,74916
Ba	0,1	0,0583	33,6153	1083,93	73,4786	2154,35	41,8353	4,72612
Ba	0,2	0,22471	31,4477	1110,98	81,7717	2342,81	41,9245	5,03084
Ba	0,3	0,36293	30,3056	1109,31	81,0109	2518,66	42,0162	5,04351
Ba	0,4	0,52197	30,1153	1107,82	80,1329	2704,02	42,0932	5,11628
Ba	0,5	0,59308	30,3526	1125,21	80,9073	2878,89	42,194	5,01217
Ba	0,6	0,62089	30,5014	1172,53	86,2506	3053,62	42,2778	5,05936
Ba	0,7	0,62648	30,492	1223,96	89,2116	3250,84	42,3894	5,11921
Ba	0,8	0,6288	30,3847	1280,65	94,4121	3361,74	42,4982	5,06234
Ba	0,9	0,64116	30,3445	1327,71	98,291	3609,83	42,5993	5,0456
Ba	1	0,63781	30,3887	1385,22	102,406	3668,03	42,7231	5,06487
Ba	1,1	0,6233	30,4464	1454,89	107,695	3894,62	42,8382	5,11761
Ba	1,2	0,63635	30,4196	1505,33	113,947	4061,93	42,9684	5,07876
Ba	1,3	0,63682	30,4062	1561,77	124,486	4228,1	43,1022	5,0702
Ba	1,4	0,66893	30,3954	1597,76	127,286	4336,04	43,2432	5,126
Ba	1,5	0,67144	30,4209	1656,4	132,112	4567	43,386	5,11834
Ba	1,6	0,70631	30,4156	1681,32	134,848	4734,58	43,5261	5,13049
Ba	1,7	0,74643	30,3851	1700,38	136,08	4955,13	43,6768	5,08862
Ba	1,8	0,7561	30,3827	1753,02	139,988	5041,68	43,8232	5,14389
Ba	1,9	0,78523	30,3834	1783,94	143,72	5235,29	43,9784	5,03226
Ca	0	0,06621	29,8236	667,187	47,4298	2306,88	41,2336	3,91324
Ca	0,1	0,14434	29,6393	997,483	73,493	2496,62	41,3048	4,02427
Ca	0,2	0,33917	29,6177	1165,18	86,2716	2672,74	41,3365	4,46631
Ca	0,3	0,73487	29,559	1087,61	80,4525	2857,77	41,4832	4,80647
Ca	0,4	1,04365	29,4681	995,638	73,8315	3033,8	41,5667	5,00187
Ca	0,5	1,2598	29,5215	923,693	68,3734	3204,5	41,6308	5,07965
Ca	0,6	1,33665	29,5349	892,962	66,2413	3366,03	41,7106	4,97938
Ca	0,7	1,50522	29,4454	871,669	63,8538	3533,8	41,7866	5,07938
Ca	0,8	1,53636	29,5349	883,85	65,6356	3700,41	41,8517	4,98369
Ca	0,9	1,56373	29,559	905,478	65,8137	3870,66	41,93	5,02636
Ca	1	1,57653	29,4253	942,716	72,3848	4029,18	42,0022	5,07236
Ca	1,1	1,58124	29,4308	974,933	74,9683	4195,43	42,095	5,08007
Ca	1,2	1,59303	29,5737	1006,21	78,5567	4362,48	42,1714	5,03587
Ca	1,3	1,58734	29,5122	1051,41	82,6364	4527,13	42,2471	5,09182
Ca	1,4	1,58108	29,5082	1088,24	84,301	4690,71	42,3415	5,06636
Ca	1,5	1,56685	29,452	1141	88,9122	4857,34	42,4225	5,08
Ca	1,6	1,55185	29,452	1182,38	93,0529	5026,86	42,5321	5,09136
Ca	1,7	1,52365	29,5095	1263,22	99,2319	5195,34	42,6327	5,06089
Ca	1,8	1,51146	29,555	1324,52	103,298	5353,45	42,7396	5,10565
Ca	1,9	1,48527	29,5803	1394,05	109,377	5532,02	42,8474	5,11778
Ca	2	1,46871	29,6298	1455,8	115,043	5689,43	42,9747	5,09716
Ia	0	0,00755	32,7412	156,488	67,4198	1454,73	41,7391	5,52571
Ia	0,1	0,00327	32,244	646,314	147,272	2060,92	41,7765	5,32144
Ia	0,2	0,02317	31,2055	700,168	147,332	2246,25	41,8504	5,40468
Ia	0,3	0,0268	30,2139	750,155	147,08	2429,64	41,9145	5,60611
Ia	0,4	0,02148	29,8236	774,998	147,196	2610,63	41,984	5,86905
Ia	0,5	0,04886	29,9532	812,629	147,411	2807,66	42,0623	5,67497
Ia	0,6	0,08328	30,0387	883,725	146,967	3088,14	42,198	5,20184
Ia	0,7	0,0883	30,0468	932,47	86,0719	3263,19	42,2805	5,21182
Ia	0,8	0,12259	30,0214	987,423	16,3672	3436,55	42,3538	4,40348
Ia	0,9	0,15003	29,9225	1046,49	54,3661	3602,73	42,4461	4,04615
Ia	1	0,15011	30,004	1092,97	79,3821	3772,63	42,5396	4,23087
Ia	1,1	0,13678	30,008	1143,8	79,7364	3938,46	42,6518	4,33524
Ia	1,2	0,10714	29,9593	1240,44	79,6046	4107,3	42,7587	4,3309
Ia	1,3	0,1039	30,0173	1374,3	79,4123	4276,67	42,8703	4,39864
Ia	1,4	0,1503	30,0254	1436,27	79,1488	4443,29	42,9795	4,43225
Ia	1,5	0,15617	29,9853	1613,38	78,7357	4616,03	43,1051	4,38564
Ia	1,6	0,17027	30,0635	1747,34	78,319	4786,82	43,2636	4,36295
Ia	1,7	0,20631	30,147	1856,78	87,7243	4958,15	43,3863	4,34843
Ia	1,8	0,20635	30,3234	1941,75	31,6167	5141,44	43,5563	4,34115
Ia	1,9	0,22336	30,3742	1993,29	31,2748	5434,43	43,7355	4,34213
Ia	2	0,25643	30,1751	2040,33	120,347	5368,38	43,8976	4,37778
Ma	0	0,17042	31,1183	69,2075	0,11429	1403,82	41,477	4,65089
Ma	0,1	0,27765	31,6729	284,417	15,1214	1961,74	41,4639	3,64269
Ma	0,2	0,37221	31,6595	806,234	65,3619	2250,15	41,5207	3,75869
Ma	0,3	0,66183	31,1309	909,244	73,9214	2450,6	41,634	3,91439
Ma	0,4	0,69476	30,7486	900,713	64,9643	2648,56	41,6714	4,03618
Ma	0,5	1,02756	30,6104	845,771	68,2179	2840,12	41,7223	4,10984
Ma	0,6	1,11893	30,6614	891,461	69,0072	3044,12	41,7939	4,17459
Ma	0,7	1,17465	30,7888	901,657	72,9358	3140,35	41,9016	4,2205
Ma	0,8	1,21139	30,8706	940,624	75,4286	3350,11	41,9143	4,30142
Ma	0,9	1,22247	30,8706	952,2	76,7679	3518,6	42,0396	4,37692
Ma	1	1,20814	30,8358	998,165	80,3529	3682,91	42,0585	4,43667
Ma	1,1	1,19651	30,7768	1117,65	94,8057	3853	42,1959	4,5051
Ma	1,2	1,1726	30,7459	1090,76	83,15	4015,53	42,2967	4,57611
Ma	1,3	1,11846	30,7339	1203,28	102,393	4215,71	42,3106	4,62339
Ma	1,4	1,07127	30,7466	1293,47	111,118	4366,53	42,525	4,68868
Ma	1,5	1,03262	30,7553	1382,33	116,629	4482,46	42,6635	4,76641
Ma	1,6	0,99295	30,754	1425,5	122,025	4631,2	42,6971	4,87714
Ma	1,7	0,94328	30,7767	1553,31	134,543	4836,48	42,8604	4,97206
Ma	1,8	0,90702	30,7382	1538,11	136,654	5022,28	42,9647	5,06176
Ma	1,9	0,85667	30,8278	1674,8	143,8	5192,63	43,113	5,07919
Ma	2	0,81617	30,8438	1758,47	149,468	5336,37	43,2735	5,11083

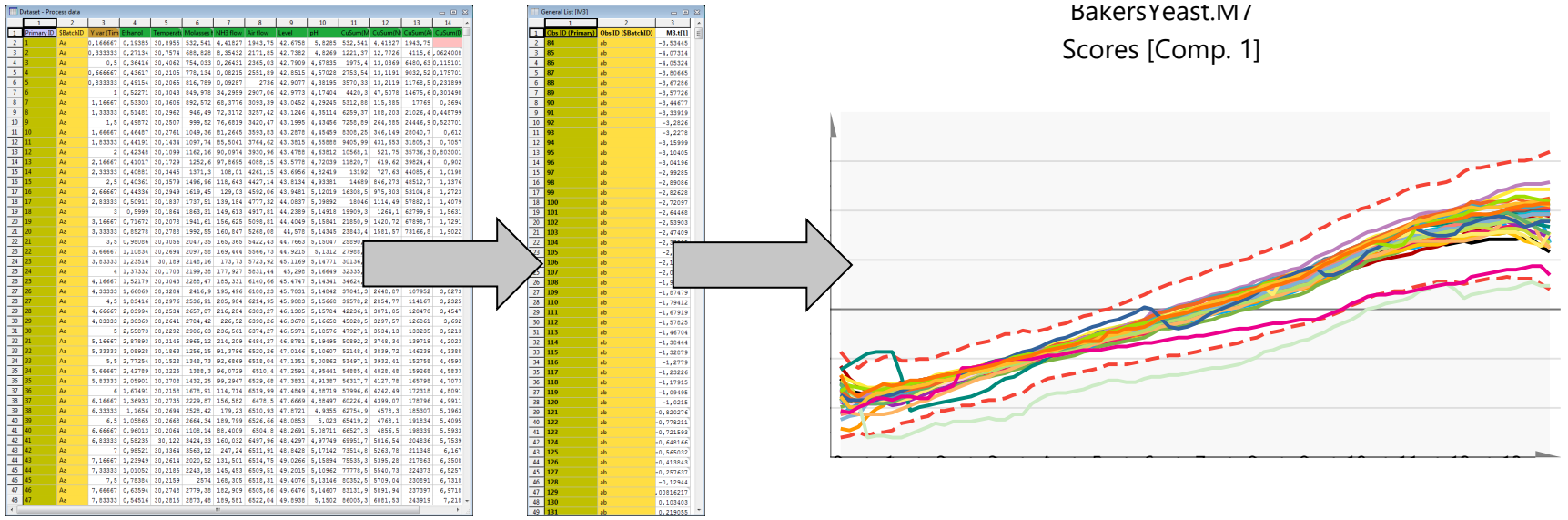
Data Structure In the BEM

- OPLS between process time points and time gives model representing process evolution in data set
- When based on “good” or “normal” batches the model represents normal batch evolution



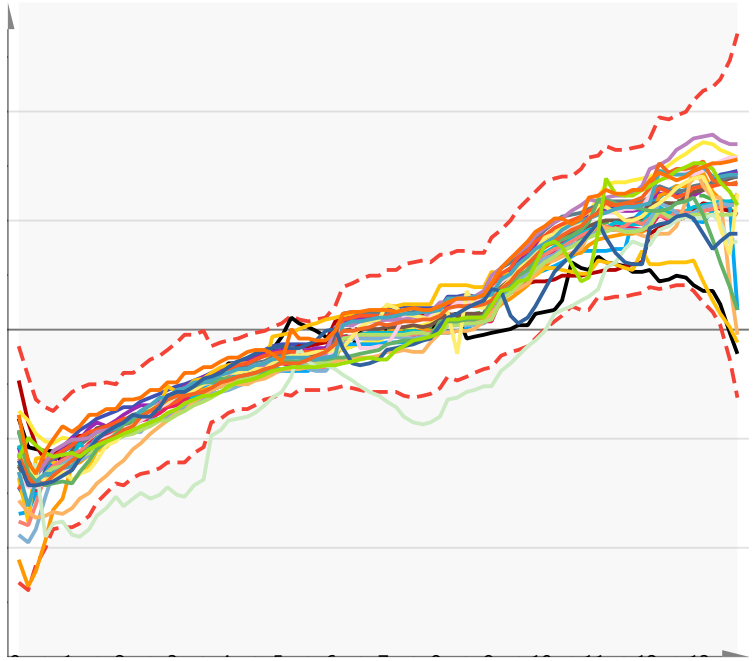
Observations, From Variables to Scores

- Observations are now described by ”scores”
- Scores are linear combinations of original variables
  - Why are observations close to each other?
  - What is the reason between dissimilarities?

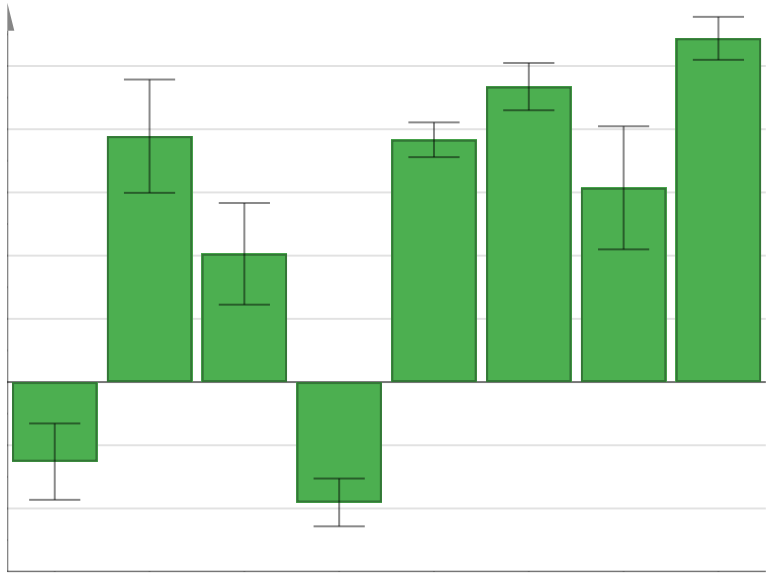


Example – Baker's Yeast

BakersYeast.MT
Scores [Comp. 1]



BakersYeast.MT (OPLS), First model excl low QP2
Normalized to unit length

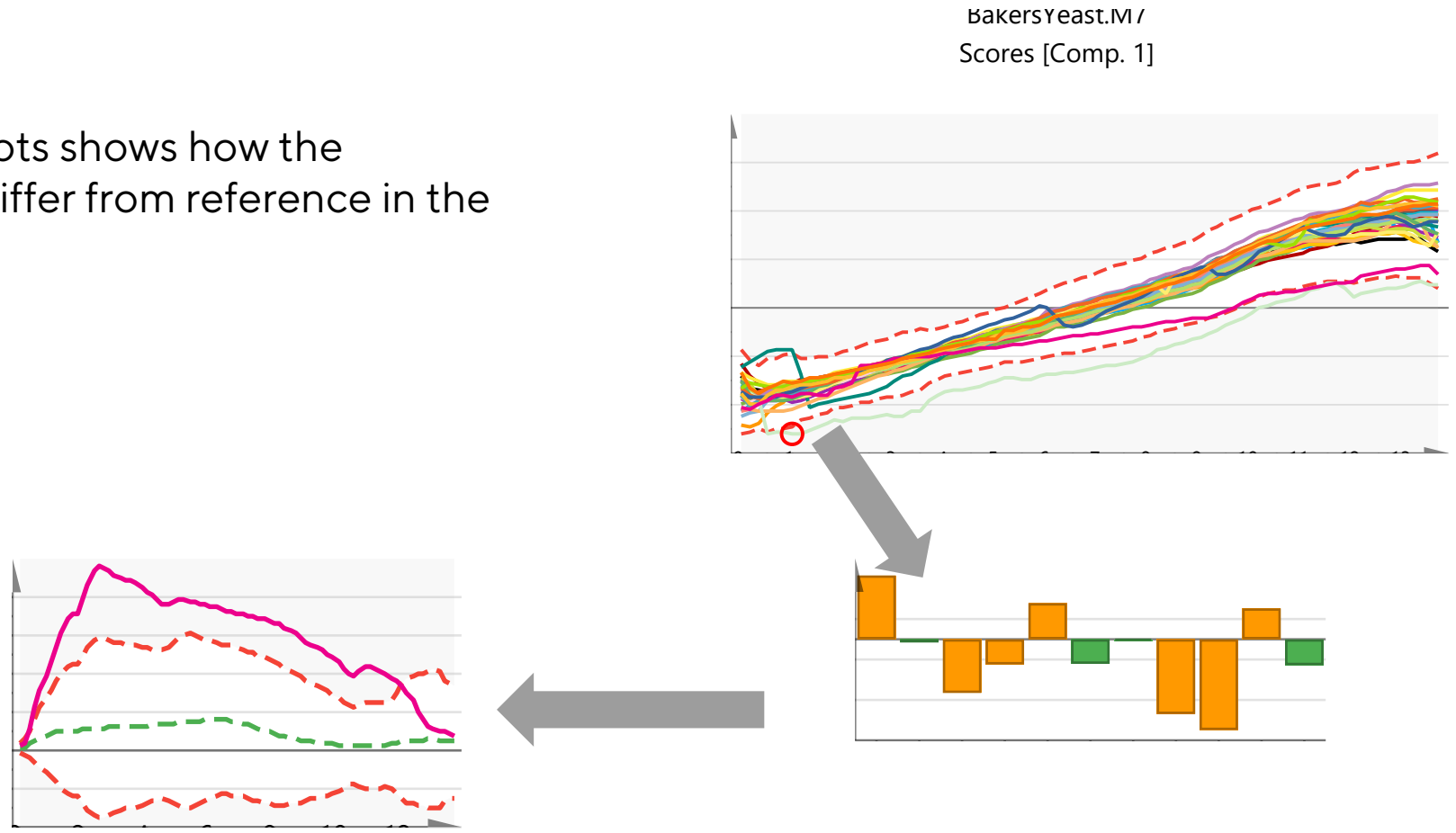


Observations - HOW

Variables - WHY

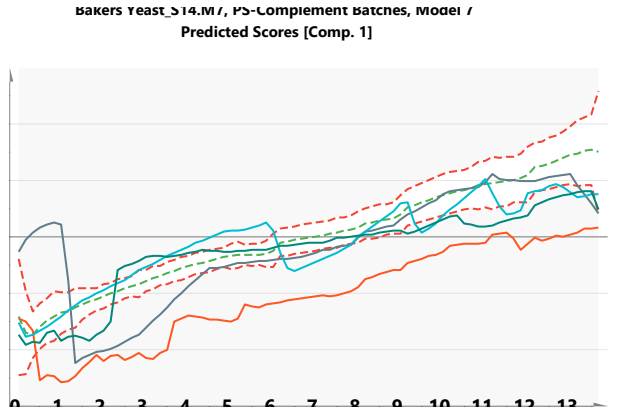
Contribution Plots - Scores

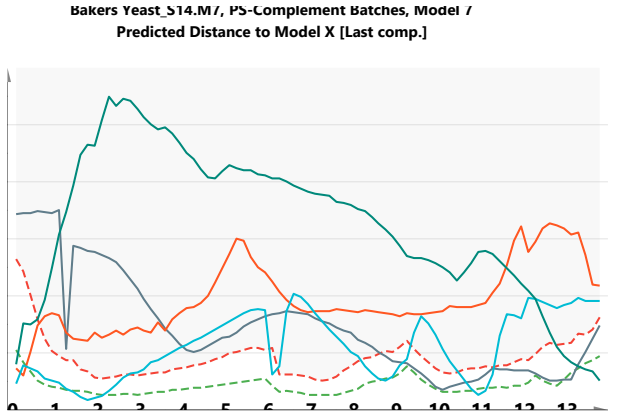
- Score contribution plots shows how the observation / group differ from reference in the model space



Validation of Batch Evolution Model (BEM)

- The poor performing and misbehaving batches are used to validate the model
- Will the model recognize batches with suboptimal behavior?
- Batches outside the +/- 3 StDev limits are considered as definite deviators
- Batches outside the +/- 2 StDev limits are considered as risky





Example: Cell Culture Evolution (BEM/BLM)

Modelling Cell Culture Metabolite Trajectories for Process Monitoring

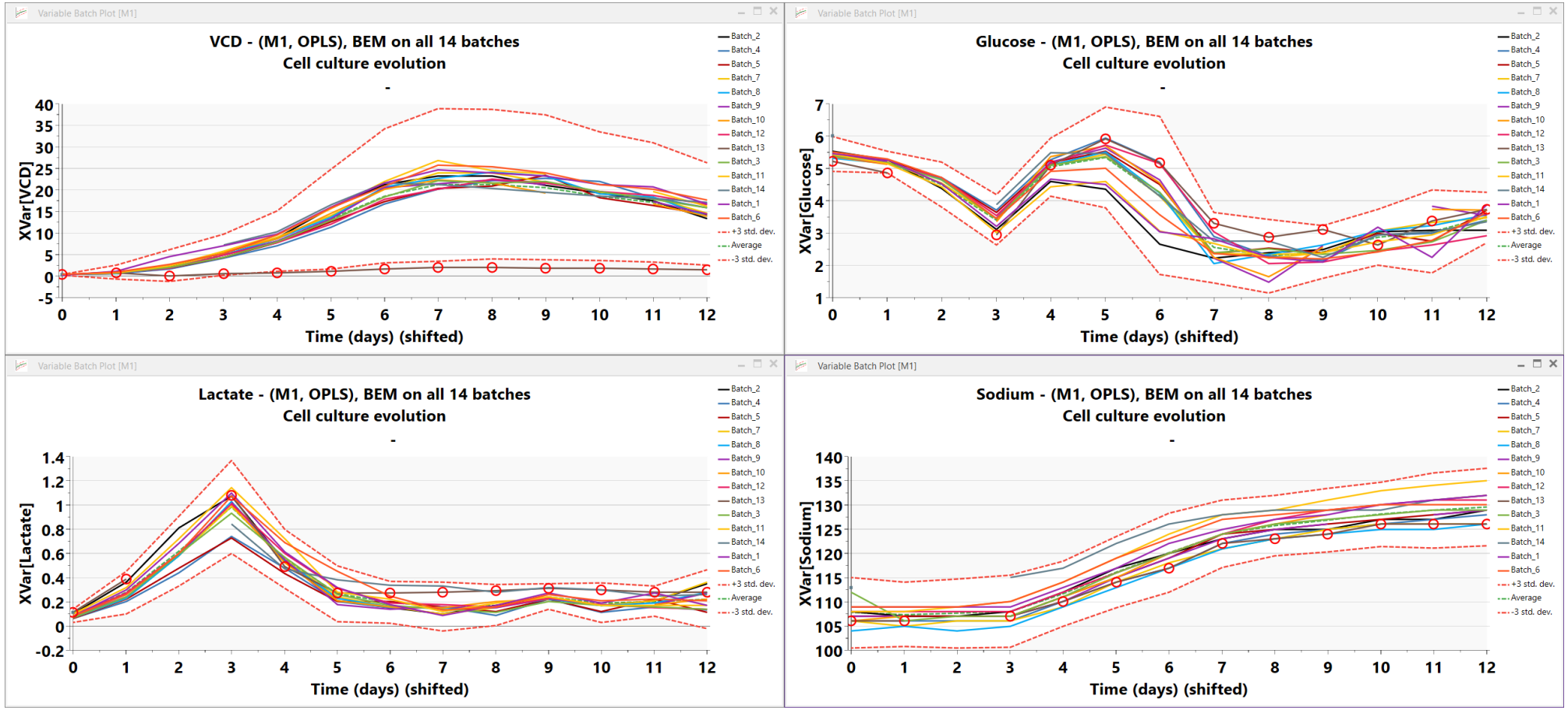
- Several fed-batch cell culture processes were performed for process development objectives.
- The experimental campaign involved processes performed under stable conditions and under unstable conditions.
- As typical parameters of the process performance metabolite measurements, cell characteristics as well as ion and gas concentration measurements were registered once per day.
- There are 14 batches. Process measurements, collected once per day, as well as initial and final batch conditions are available.

Process Evolution Measurements Used to Get the BEM

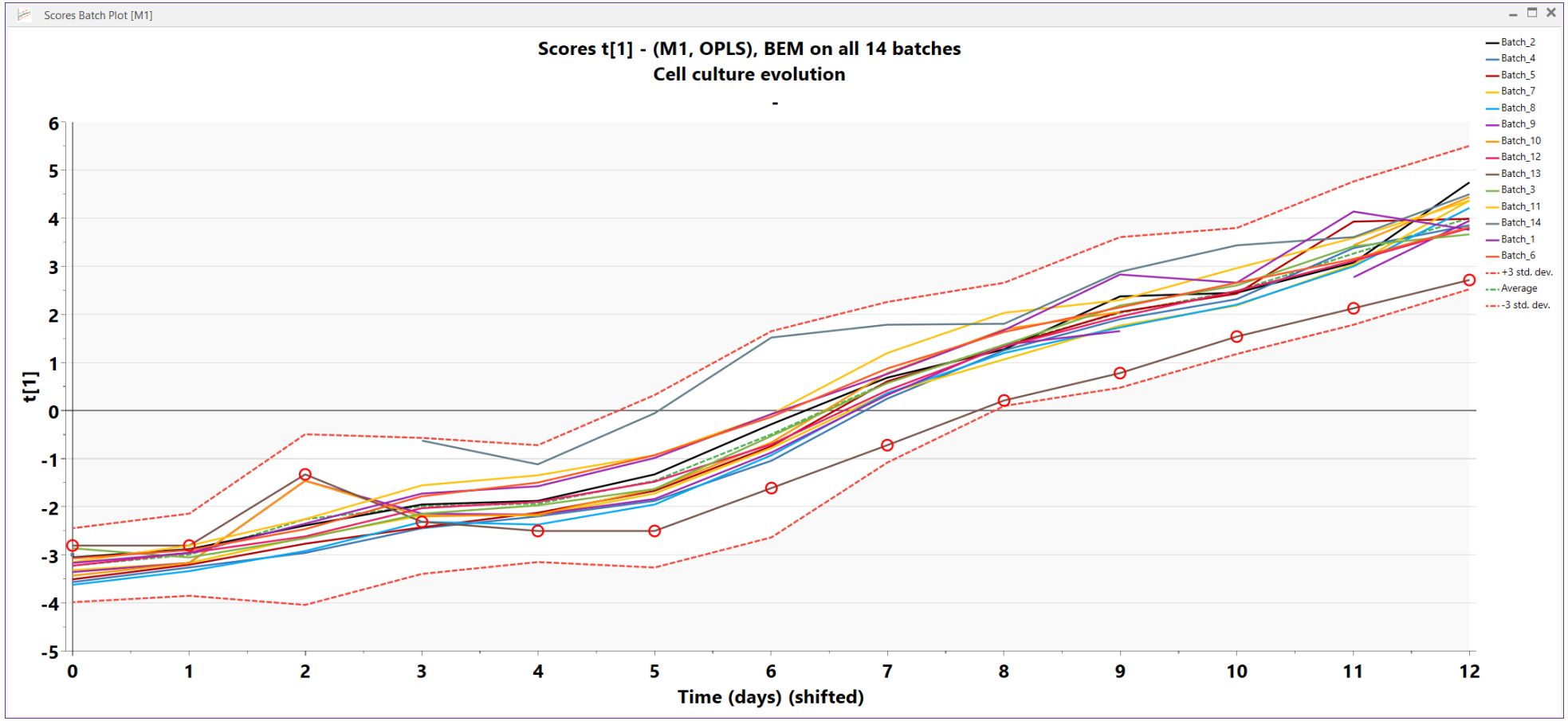
Dataset - Evolution data

	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1	Primary ID	$BatchID	Time (days)	VCD	Viability	Glucose	Lactate	Titer	pCO2	pH Offline	pO2	Osmolality	Potassium	Sodium
41	40	Batch_7	0	0.324	98.2	5.4	0.07	0.0104639	49.5	7.106	91.2	306	9.4	106
42	41	Batch_7	1	0.713	98	5.19	0.23	0.0222557	48.6	7.09	92.5	306	9.3	105
43	42	Batch_7	2	2.036	99	4.56	0.58	0.0647377	36.1	7.102	90.5	304	9.1	106
44	43	Batch_7	3	5.492	99	3.45	0.99	0.165377	24.1	7.07	89.3	290	8.6	106
45	44	Batch_7	4	9.072	99	5.09	0.55	0.339043	38.9	7.094	83.8	309.5	9.4	109
46	45	Batch_7	5	14.1415	99	5.44	0.24	0.620911	58.8	7.085	69.9	319.67	9.7	114
47	46	Batch_7	6	20.1197	99	4.13	0.16	1.07201	70.3	7.09	64.9	315.5	9.4	118
48	47	Batch_7	7	23.984	98.6	2.37	0.1	1.84783	81.9	7.095	61.9	304.5	9.7	121
49	48	Batch_7	8	24.017	98.33	2.35	0.15	2.35966	81.1	7.095	55.6	308	10.9	123
50	49	Batch_7	9	21.848	96.7	2.37	0.24	2.8195	73.5	7.099	52.7	317.5	11.8	125
51	50	Batch_7	10	19.488	95	3.06	0.17	3.17237	77.9	7.103	57.7	331.5	13	126
52	51	Batch_7	11	18.207	89.13	3.32	0.21	3.43495	83.7	7.096	57.9	345	14.1	126
53	52	Batch_7	12	14.51	76.95	3.46	0.36	3.49327	78.1	7.1	77	364	15.5	126
54	53	Batch_8	0	0.31	99	5.32	0.07	0.0100164	54.4	7.064	91.4	303	9.2	104
55	54	Batch_8	1	0.63	99.1	5.15	0.23	0.0215393	46.4	7.1	89.3	300.5	9.1	105
56	55	Batch_8	2	1.829	98.9	4.66	0.58	0.0642541	35.4	7.104	90.3	298	8.9	104
57	56	Batch_8	3	4.872	99.3	3.44	1.03	0.179967	21.5	7.089	89.9	284	8.4	105
58	57	Batch_8	4	8.625	99.3	5.11	0.54	0.344157	41.3	7.07	79.8	307.5	9.2	109
59	58	Batch_8	5	13.635	99.35	5.51	0.23	0.681418	57.3	7.09	77.5	317	9.4	113
60	59	Batch_8	6	20.702	99.3	4.24	0.16	1.16472	68.8	7.09	58.3	309	9.4	117
61	60	Batch_8	7	22.644	98.06	2.05	0.15	1.76629	78.9	7.082	44.1	299.5	9.6	121
62	61	Batch_8	8	24.0935	97	2.34	0.16	2.2537	84.7	7.08	14.7	309	10.7	123
63	62	Batch_8	9	23.251	96	2.64	0.22	2.74823	77.1	7.102	47.3	317	11.8	124
64	63	Batch_8	10	19.175	93.9	3.05	0.17	3.11018	81.4	7.101	46	331.5	12.9	125
65	64	Batch_8	11	18.38	90.05	3.24	0.19	3.32718	85.7	7.095	54.8	349	14.1	125
66	65	Batch_8	12	13.958	79.2	3.56	0.28	3.49662	84	7.1	57.9	367	15.4	126
67	66	Batch_9	0	0.312	99	5.41	0.08	0.00943621	48.2	7.116	89.9	301	9.5	107

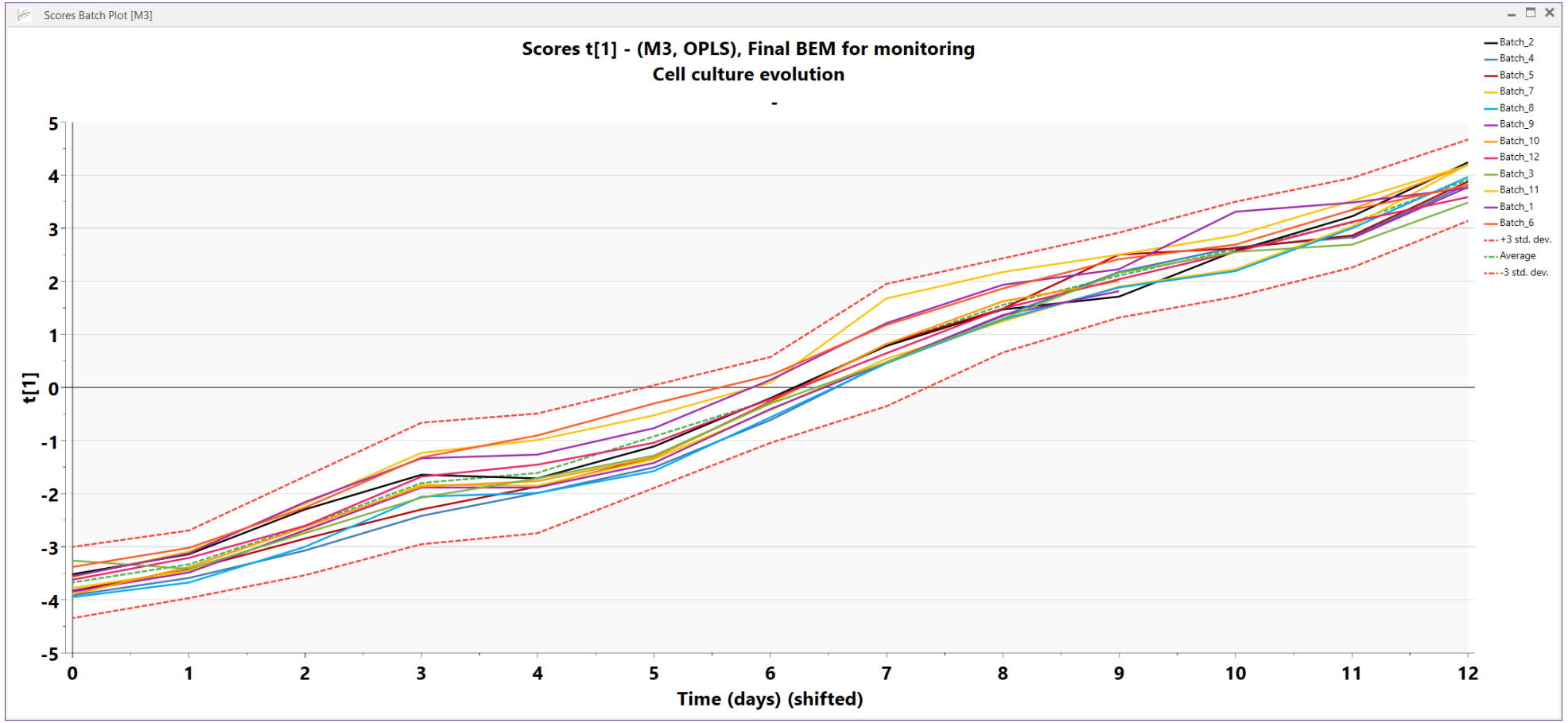
Batch Control Charts (BCC) of Individual Variables (Some Examples)



All Trajectories Summarized by the Score BCC



Result: Two Batches to Be Removed ➔ Final BEM Representing Good Process



Batch Condition Data Used for BLMs

- A BLM is used to overview ALL data of completed batches
- A BLM can be used to predict final conditions
  - E.g. What process evolutions measurements correlate with final titer?
  - And when during cell culture lifetime?

Dataset - BC1_Initial conditions

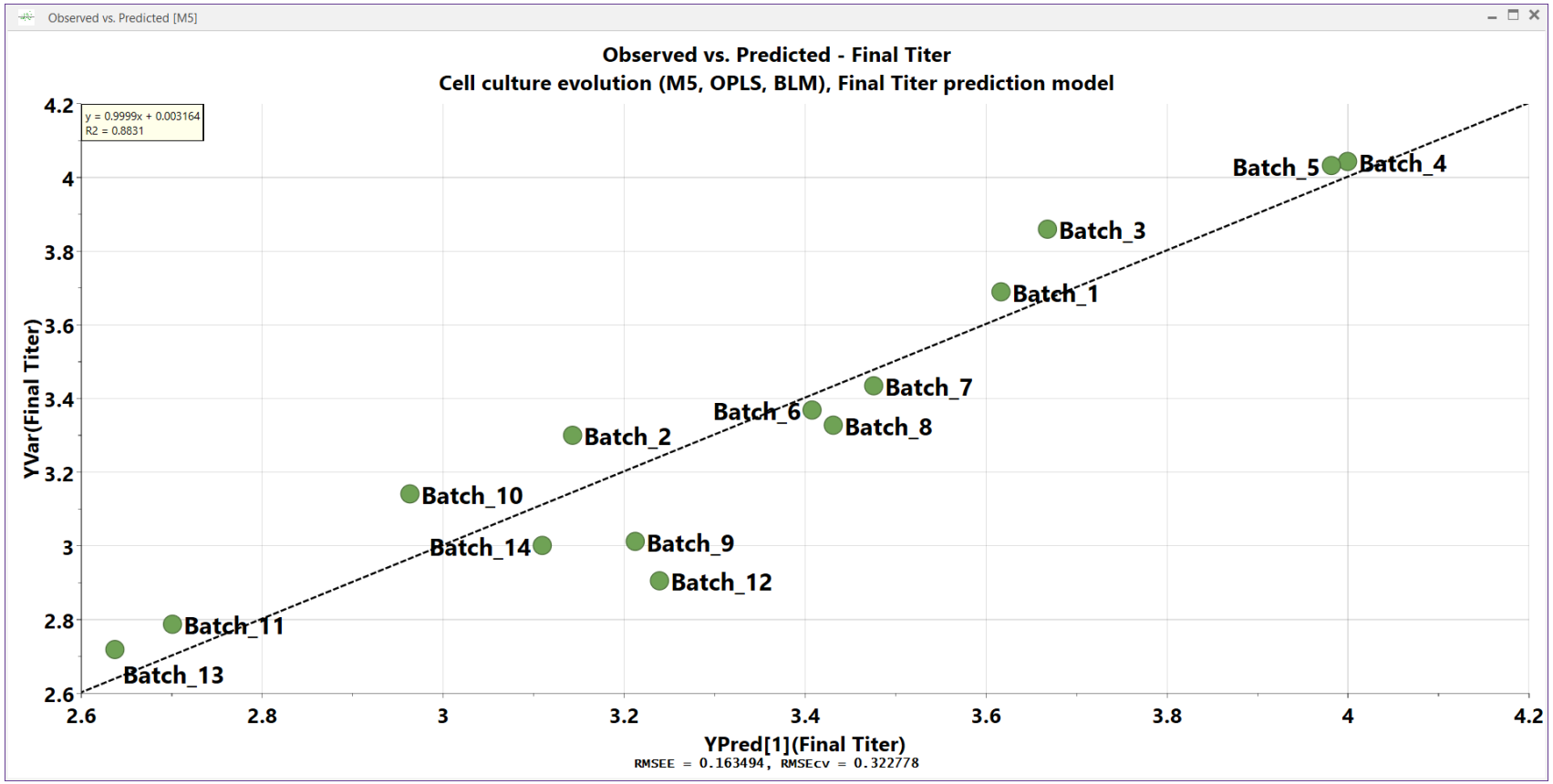
	1	2	3	4	5	6	7	8	9	10	11	12
1	Primary ID	$BatchID	Initial Glucose	Initial Lactate	Initial Osmolality	Initial Potassium	Initial Sodium	Initial VCD	Initial Viability	Initial CO2	Initial pH	Initial O2
2	$Batch_1	Batch_1	5.45	0.08	308.5	9.7	109	0.275	99	50.3	7.1	97.4
3	$Batch_2	Batch_2	5.53	0.1	307	9.6	108	0.3	96.5	47.7	7.109	88.7
4	$Batch_3	Batch_3	5.38	0.07	308	10	112	0.3	97.3	52.5	7.075	83.1
5	$Batch_4	Batch_4	5.44	0.06	307	9.5	106	0.325	99.1	40.9	7.08	88
6	$Batch_5	Batch_5	5.42	0.06	308	9.5	107	0.325	99.1	40.4	7.08	96.5
7	$Batch_6	Batch_6	5.5	0.08	315	9.6	109	0.3	98.2	53.4	7.084	100
8	$Batch_7	Batch_7	5.4	0.07	306	9.4	106	0.325	98.2	49.5	7.106	91.2
9	$Batch_8	Batch_8	5.32	0.07	303	9.2	104	0.3	99	54.4	7.064	91.4
10	$Batch_9	Batch_9	5.41	0.08	301	9.5	107	0.3	99	48.2	7.116	89.9
11	$Batch_10	Batch_10	5.36	0.1	300	9.4	106	0.325	99.4	49.9	7.105	88.6
12	$Batch_11	Batch_11	5.44	0.1								108
13	$Batch_12	Batch_12	5.38	0.08								90.5
14	$Batch_13	Batch_13	5.22	0.11								102
15	$Batch_14	Batch_14	6.01	0.11								94

Dataset - BC1_Process output

	1	2	3	4	5	6	7
1	Primary ID	$BatchID	Peak Titer	Peak VCD	Final Titer	Final VCD	Final Viability
2	$Batch_1	Batch_1	3.6895	24.65	3.6895	20.7	90.8
3	$Batch_2	Batch_2	3.42983	23.3	3.29939	17.5	89.3
4	$Batch_3	Batch_3	3.99882	22.225	3.86063	17.8	92.4
5	$Batch_4	Batch_4	4.26145	22.7	4.04247	18.2	94
6	$Batch_5	Batch_5	4.25435	23.425	4.03154	16.4	93
7	$Batch_6	Batch_6	3.62	25.7	3.37	20.2	92.7
8	$Batch_7	Batch_7	3.49327	23.975	3.43495	18.2	89.13
9	$Batch_8	Batch_8	3.49662	24.075	3.32718	18.4	90.05
10	$Batch_9	Batch_9	3.23808	22.4	3.01178	17.325	91.45
11	$Batch_10	Batch_10	3.40568	22.375	3.14113	16.65	89.8
12	$Batch_11	Batch_11	2.93629	26.825	2.78789	19.6	91.8
13	$Batch_12	Batch_12	3.08995	22.5	2.90643	18.8	94.3
14	$Batch_13	Batch_13	2.92	1.925	2.72	1.6	93.4
15	$Batch_14	Batch_14	3.18	21.625	3	18.1	92.4944

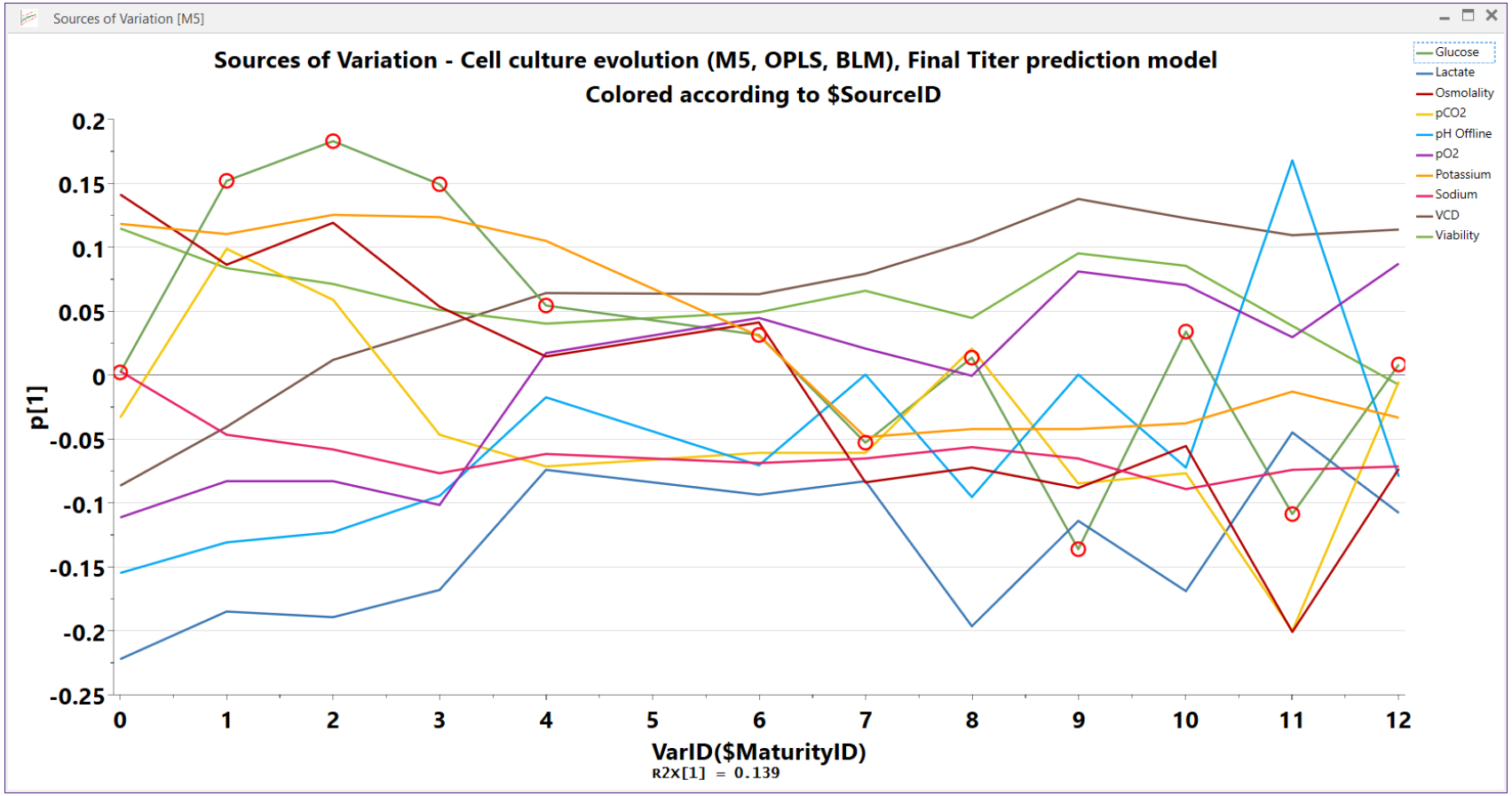
Example of BLM: Final Titer Prediction Model

- Strong agreement between measured final titer and model predictions



Sources of Variation Plot to Reveal Influence of Process Parameters

- The influence of the process parameters across lifetime of cell culture process visualized
- Glucose highlighted as an example
- Possibility to sharpen model by dropping non-influential parameters and/or time segments



Demo (Cell Culture Dataset)

Conclusions From Demo Example

- This example shows a standard strategy to build multivariate models for cell culture processes with the long-term goal of enabling real-time process monitoring.
- The strategy is based on two types of models, the batch evolution model (BEM) and the batch level model (BLM).
- The objective of the BEM is to allow process monitoring of a cell culture process as it is evolving and as early as possible flag any deviation from a normal operating condition.
- The objective of the BLM is to summarize all data (initial conditions, process evolution measurements, final conditions,...) for completed batches to explore possibilities for further process optimization and process enhancement.
- The different models developed in this example suggest high quality data and a strong association between final titer and the way the cell cultures were progressing.

General Functionality of SIMCA Family

Off-line

Historical data

On-line

SimApi

Real-time data

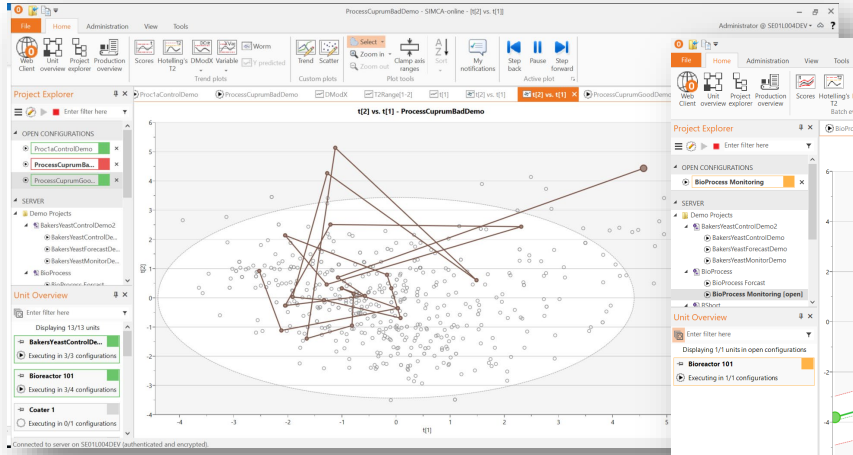
- SIMCA®
  - Creates models from historical data
  - Model objectives include
    - Finding deviating batches
    - Predicting a quality parameter
  - Resulting models are stored in a USP (Umetrics® SIMCA® Project) file

- SIMCA®-online
  - Samples process data at user specified frequency
  - Uses USP files as reference to new data
  - Visualizes the current state of the process

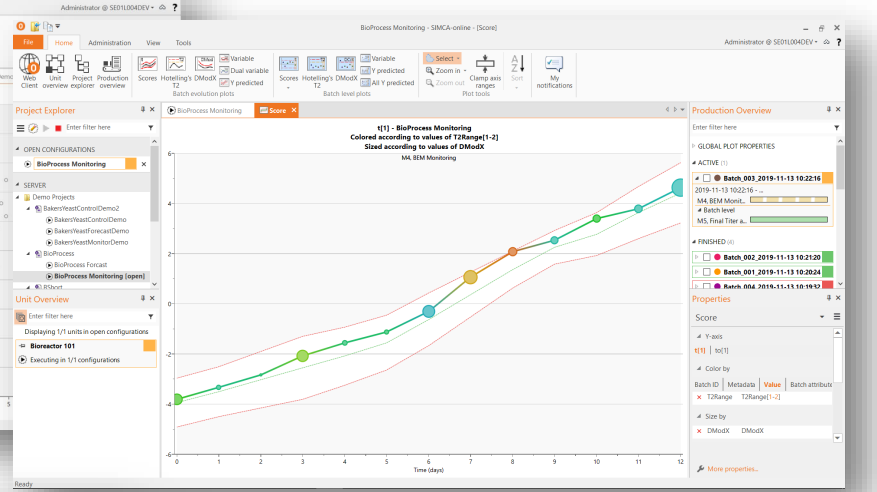
What can SIMCA®-online Do?

- **Monitor** – in real time
  - Create an ideal model of your process
  - Compare your process with the model in real time
- **Predict** – with confidence
  - Final quality prediction
  - Removes expensive measurements
  - All predictions has validity measures
- **Control** – at a glance
  - Forecast quality predictions
  - Learn how to optimize your process

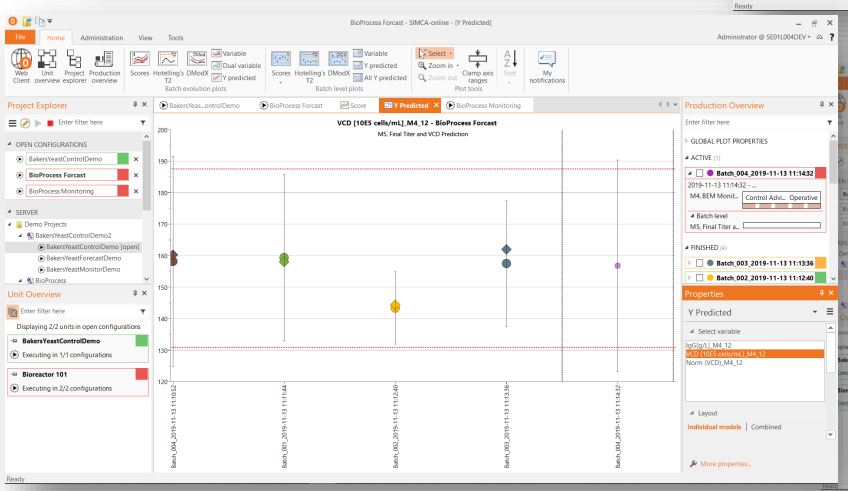
Continuous



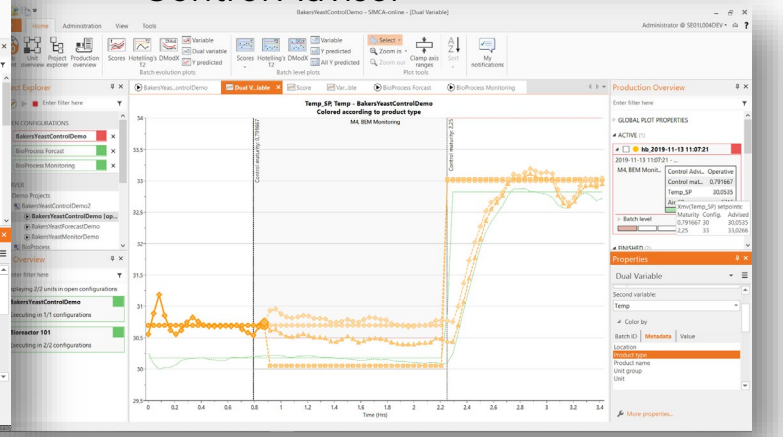
Batch evolution



Batch level



Control Advisor



Upcoming Webinars (https://www.sartorius.com/en/company/exhibition-conferences)