Process Optimization Through Predictive Modelling: Making It Work!

Presentation by :

Vincent Béchard, M.Sc.A.
Dominic St-Onge, P.Eng., MBA

The speakers

Dominic St-Onge, P.Eng., MBA
Consultant in Operational Excellence
Différence GCS inc.
dstonge@difference-gcs.com

Vincent Béchard, B.Eng., MASc.
Associate, Analytical Decision Specialist
Différence GCS inc.
vbechard@difference-gcs.com

- 18 years of experience in pulp and paper industry, in Canada, USA and UK
- Bachelor degree in Chemical Engineering, Master's in Business Administration
- Lean Six Sigma Black Belt since 2015
- Passionate about pulp and paper processes, meeting people, solving problem and analyzing data!

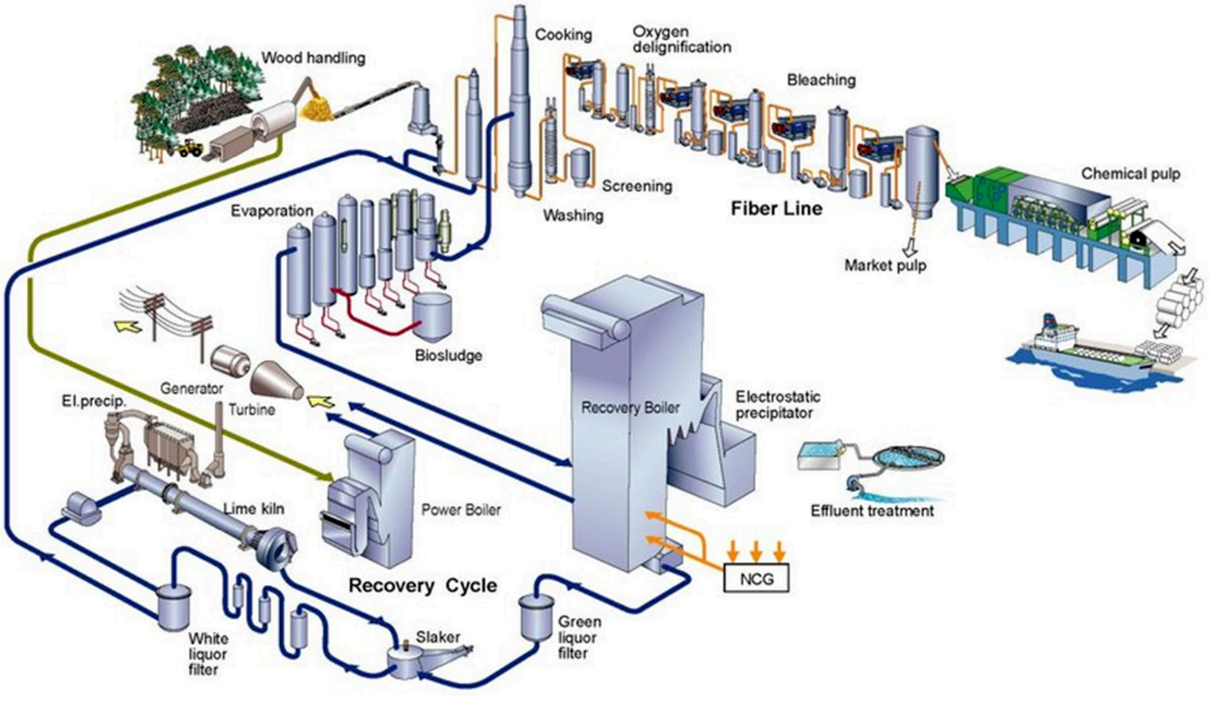
- 19 years of experience in mathematical decision-making, simulation and data science
- Bachelor degree in Chemical Engineering, Master's degree in Applied Mathematics (process optimization)
- Analyzed tons of data and developed many scientific applications, models, simulations…
- Helping clients to unlock value from their databases and to reach the full potential of their assets. If it's data or modelling, I am here!

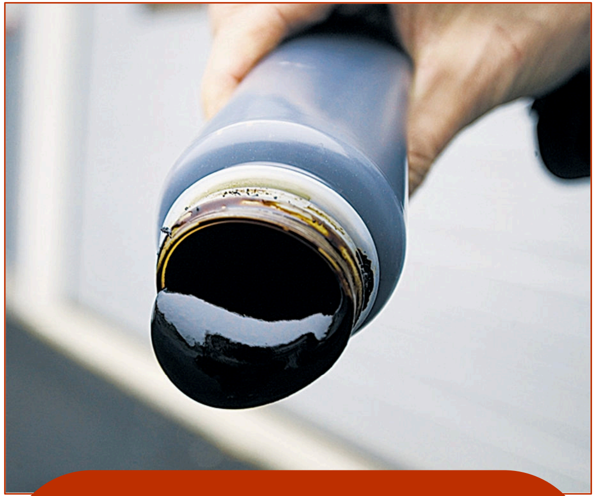
Agenda

- History and technical works
- Towards model-based adjustments
- An happy ending story

History and technical works

Some technical background

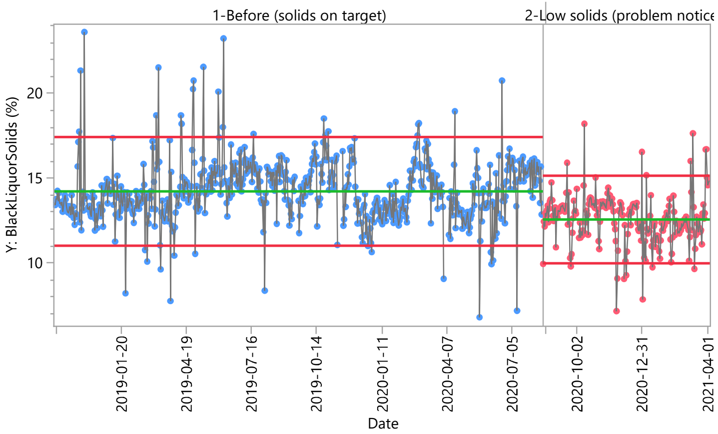


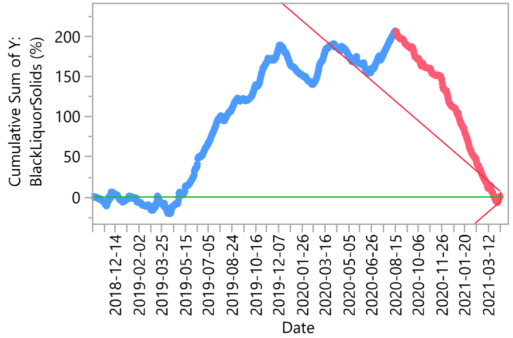


Black liquor is mainly composed of lignin (organic), inorganics and water!

Something went wrong

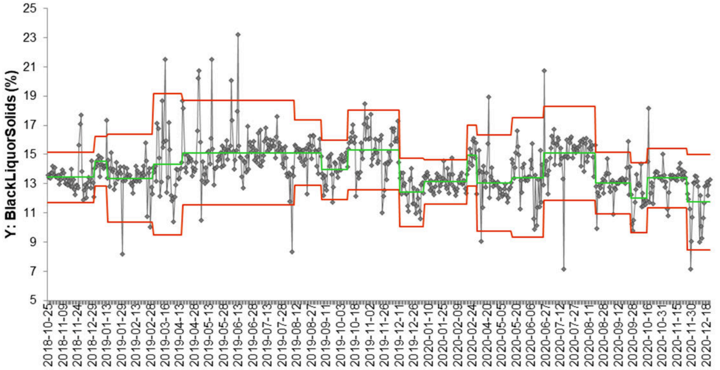
- The problem:
  - A Kraft pulp mill has seen a drop in black liquor (BL) solids and was unable to achieve historical target (15%)
- Rule of thumb
  - For a 1000 t/day mill, an increase of 1% in weak BL solids represents 1M$/year in steam and energy costs





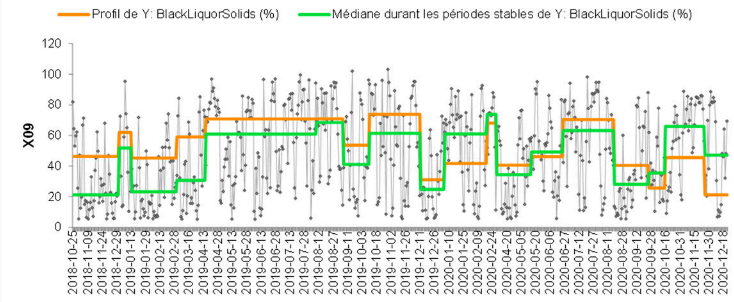
Efficient data analysis

≡ Change Point Analysis (CPA) and Stability Analysis



- ≡ Using the new Difference Excel Add-In (released in 2023):
  - CPA on main output variable (Y)
  - Generates table with statistical change point and dates (easy to relate to an event timeline)

Sommaire des segments			CPA [5000 ¤ 0.7 ¤ 3 ¤ 5000 ¤ 0.9]			
De	À	Valeur p	Moyenne	Nb. obs.	Écart-type	Précision (±)
2018-10-25	2018-11-20		13.4423677	27	0.38336725	0.15165503
2018-11-21	2018-11-25	0.0008	12.650219	5	0.26729246	0.33188743
2018-11-26	2018-12-09	0.0020	15.5228341	13	3.53299023	2.13496463
2018-12-10	2018-12-19	0.0022	13.3203533	10	0.46458322	0.33234281



  - Stability analysis on all Xs at the same time to correlate individuals Xs CPAs with output variable

[Différence Excel Add-in Statistical Data Analysis Simulation (difference-gcs.com)](difference-gcs.com)

Efficient data analysis

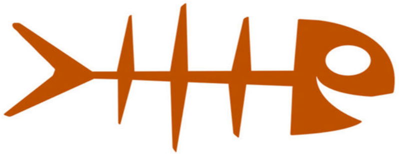
≡ Change Point Analysis (CPA) and Stability Analysis

- ≡ Generates an overview table of all Xs in relation with statistical phases of Y
- ≡ Automatically ranked in order of similarity importance
- ≡ Also gives p-values, correlation slopes and identifies other Xs correlated for each individual X.

Variable (cliquer!)	r (médianes)	Valeur p de r	ΔY / ΔX	Similitude	*Variables corrélées*
X09	0.619	0.0062	0.0437372	88.2%	X10; X12; X16; X21;
X22	-0.513	0.0295	-0.00864659	82.4%	X50;
X18	0.409	0.0923	7.3338E-05	76.5%	X62; X01; X32; X50; X57; X64;
X51	-0.319	0.1973	-0.20629519	70.6%	
X53	0.317	0.1995	0.12249628	70.6%	X57; X64; X85; X101;
X12	0.558	0.0162	9.7905E-06	64.7%	X09; X10; X11; X16; X21;
X21	0.495	0.0367	0.65823415	64.7%	X09; X10; X12; X16;
X16	0.495	0.0367	0.00021923	64.7%	X09; X10; X12; X21;
X17	0.317	0.2007	7.275E-05	64.7%	X13; X23;
X23	0.317	0.2007	0.21843545	64.7%	X13; X17;
X13	0.316	0.2009	1.166E-05	64.7%	X17; X23;
X11	0.518	0.0277	0.527173	58.8%	X12; X32; X57; X64; X85;
X106	0.493	0.0376	0.02403858	58.8%	X18; X32; X43; X45; X47; X50;
X01	0.451	0.0604	0.00010754	58.8%	X18; X62; X35; X64; X68;
X104	0.393	0.1063	0.06350711	58.8%	
X62	0.384	0.1155	0.00045434	58.8%	X18; X01;
X15	-0.326	0.1871	-2.034E-05	58.8%	X20; X25;
X25	-0.323	0.1908	-0.50601941	58.8%	X15; X20;
X20	-0.323	0.1908	-0.00016853	58.8%	X15; X25;
X10	0.317	0.1998	0.03182864	58.8%	X09; X12; X16; X21;
X68	0.285	0.2518	0.00098951	58.8%	X18; X01; X32; X36; X43; X45;
X38	0.259	0.2994	0.00073099	58.8%	
X75	-0.611	0.0071	-0.003058	52.9%	X32; X36; X43; X45; X47; X48;
X50	-0.491	0.0384	-0.0042652	52.9%	X18; X22; X32; X36; X43; X45;
X48	-0.367	0.1345	-0.00034643	52.9%	X32; X36; X42; X43; X45; X47;
X31	0.444	0.0647	0.08562446	47.1%	
X79	-0.444	0.0647	0.06990934	47.1%	X32; X36; X43; X45; X47; X48;
X69	-0.435	0.0711	4.0917915	47.1%	X30; X32; X36; X43; X45; X47;
X35	0.427	0.0775	0.00051477	47.1%	X01;

Typical intervention

- Kaizen:
  - Standard problem solving through DMAIC allowed to identify major causes and resolve technical issues:
    - Mechanical failure at the brown stock washers
      - ie.: water infiltrations through pump water glands
    - Poor loop tuning
      - ie.: conductivity control at decker
    - Process improvement opportunities
      - ie.: more efficient filtrate management
    - Etc.



Experience-based control

≡ Historically, operators adjusted their process manually based on transferred knowledge and long term practices

- Conductivity target of the pulp at the decker outlet
- Dilution ratio at the brown stock
- Extraction ratio at the digester
- Dilutions in washer vats along with defoamer dosage
- Etc.

≡ Following the Kaizen, despite all good will, BL solids were **higher but still variable**!

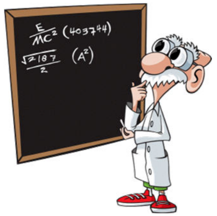
How could we elevate our game to the next level?!

Towards model-based adjustments

Building a 1-step ahead predictive model

≡ Idea: have a statistical model that can predict the next observation of the KPI based on actual process values

- Why? To anticipate the KPI trends to compensate undesirable drifts.
- For what? To ultimately reduce the variability of a KPI; it's a sort of control loop



≡ Danger! This type of prediction must be used carefully… to avoid degrading instead of healing variability!!!

Building a 1-step ahead predictive model

≡ Basically, it's statistical modelling with a twist:

- Organize the data table: the time increment between two rows has to be approximately the same
- Select the Xs to use in the model (already done here!). These Xs must not be colinear, they have to be independently adjustable!
- Build and validate a predictive model where Y (predicted variable) is in fact the next row's Y value…

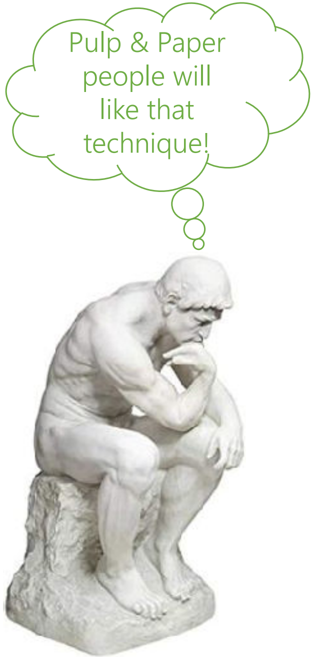
These observations…

predict this value

Date	Y: BlackLiquorSol...	X18	X62	X35	X22	X39
2019-01-24	14.4737558	31710.39...	6495.8998...	5743.312525	533.94260...	0.07802508
2019-01-25	13.6626888	32858.47...	6101.7649...	5359.286698	668.34268...	0.081431...
2019-01-26	13.308412[illegible]	31360.34...	5698.3330...	4687.762147	598.76097...	0.079245...
2019-01-27	13.82677088	34295.29...	7539.1495...	5978.216386	692.00789...	0.069232...
2019-01-28	14.15094186	33466.97...	6813.7698...	5256.662116	598.02002...	0.081431...
2019-01-29	13.69417992	29896.23...	6283.7558...	4651.473281	538.18067...	0.02370575
2019-01-30	8.171630465	13864.32...	2916.11218	[illegible]		
[illegible]	13.15255993	[illegible]				

Building a 1-step ahead predictive model

- We used a machine learning technique called "Random Forests":
  - Based on regression trees: robust handling of missing values and outliers
  - Captures non-linearities and correlations structure
  - Built-in protection against extrapolation
  - Merely impossible to overfit and built-in cross-validation mechanism
  - But: no explicit equation between the Xs and Y; model structure can't be interpreted



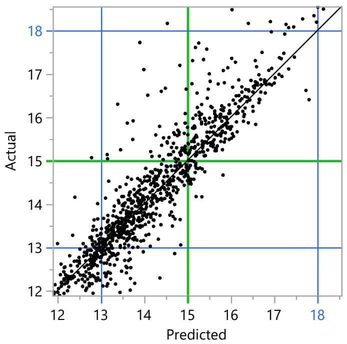
Building a 1-step ahead predictive model

≡ Our 1-step ahead predictive model:

Verification that the Xs are not colinear using the eigenvalues rule (max/min<100):
1.77/0.39 = 4.5 < 100

Number	Eigenvalue
1	1.769600
2	1.122921
3	1.015521
4	0.704031
5	0.387927

RSquare	RASE	N
0.751	0.979969	1032



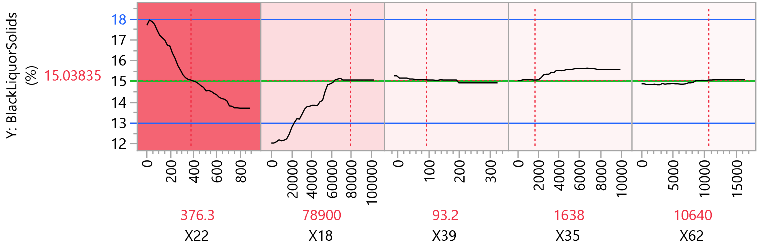
Term	Portion
X22	0.4627
X18	0.2069
X39	0.1131
X35	0.1110
X62	0.1063

Sufficient goodness of fit in the region that matters (inside the specs limits!)

The model shows non-linearities (and the typical saturation profile from random forests).

Most Xs can be set to a robust point (flat profile).

Unfortunately, the most influential X is a unstable one...

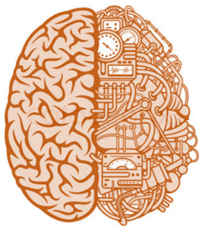


Recommender algorithm

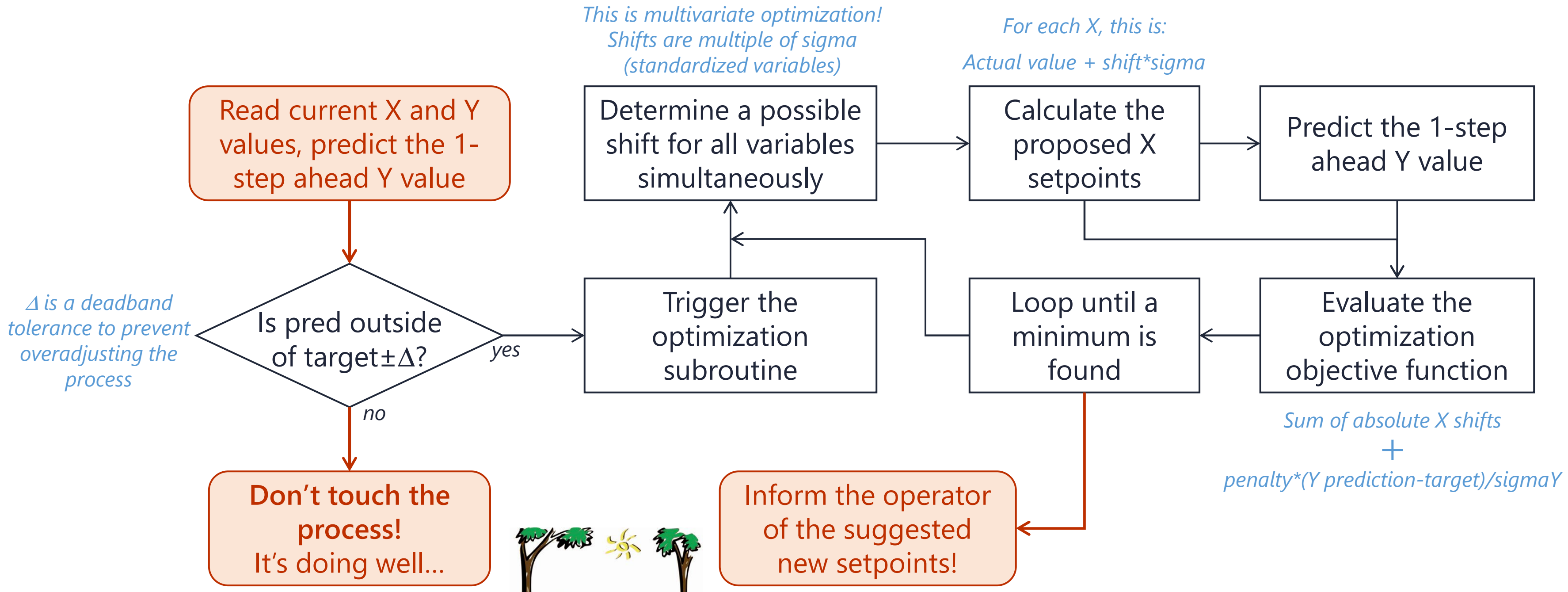
≡ A recommender algorithm is a procedure that suggests (to the user) a multivariate move that should bring the KPI on its target

- At-your-risk approach: automated online process adjustments
  - This is <u>artificial</u> intelligence!

- Safer (wiser!) use: suggest to the operator how to re-adjust the X setpoints
  - Based on her/his experience, the operator can accept or alter the recommendations.
  - This is <u>augmented</u> intelligence!



Recommender algorithm



Recommender algorithm

≡ Complicated? Yeah… a little bit, but it pays a lot!

Random forests (or any machine learning technique!) do not offer an explicit/easy to use equation…

Have to iteratively explore the relationship through optimization!

Why using machine learning then???

Go back to the benefits of random forests! Why would you not use it?

The optimization subroutines really does the magic!!!

Find the overall smallest move from actual Xs setpoint… Call me lazy!

Sum of absolute X shifts

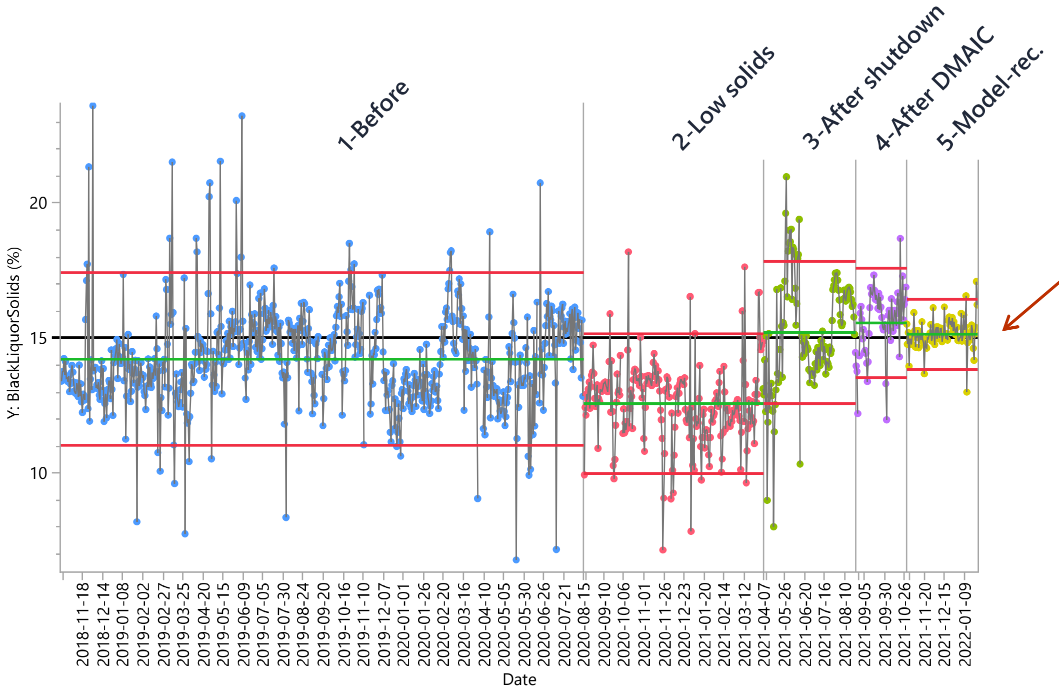
+

penalty(Y prediction-target)/sigmaY*

Control your aggressiveness with the penalty factor! But make sure to move back to the target…

An happy ending story

Can you see the progress?!



On target with minimal variability!

Ppk improved from 0.22 to 1.33!

Off specs reduced from 23.6% to 1.1%!

	Avrg.	Stdev.	Ppk
1-Before (solids on target)	14.23	1.83	0.22
2-Low solids (problem noticed)	12.57	1.47	-0.1
3-After technical shutdown	15.2	2.11	0.35
4-After DMAIC	15.57	1.2	0.68
5-Model-recommended adjust.	15.15	0.54	1.33

différence
consulting in statistics

Différence is a society offering coaching, consulting and training services in statistic, data science, simulation and continuous improvement.

We promote the use of quantitative tools that can be applied at the different steps of an improvement and variability reduction project.

Powerful methods

Adapted approach

Combining hard work with fun

For more information, you can contact:

Dominic St-Onge, P.Eng., MBA
Consultant in Operational Excellence
+1 (819) 230-1189
dstonge@difference-gcs.com
linkedin.com/in/dominic-st-onge

Vincent Béchard, MASc.
Associate, Analytical Decision Specialist
+1 (438) 521-5829
vbechard@difference-gcs.com
linkedin.com/in/vincentbechard