NEW TRENDS IN PROCESS SIMULATION: PAVING THE WAY TO DIGITAL TWIN. APPLICATION TO HYDROCRACKING

B. CELSE

Benoit.celse@ifpen.fr

3rd European Forum on New Technologies - CHEMICAL ENGINEERING in the PLANT OF THE FUTURE

PRESENTER

RESPONSIBLE
OIL AND GAS

- **Benoit Celse (benoit.celse@ifpen.fr)**
- **IFPEN**
- **Background:**
  - BS in Applied Mathematics and Computer Science Engineering from Grenoble University, PHD on Signal Processing from Lille University (France) in 1999.
  - 15 years as a data scientist and data engineer: Data Science, Fault Diagnosis, Signal and Image processing, Information System.
  - 10 years as a process engineer
  - Now Project manager in Hydroprocessing and a digitalization expert and **Data Science Evangelist** ☺

AGENDA

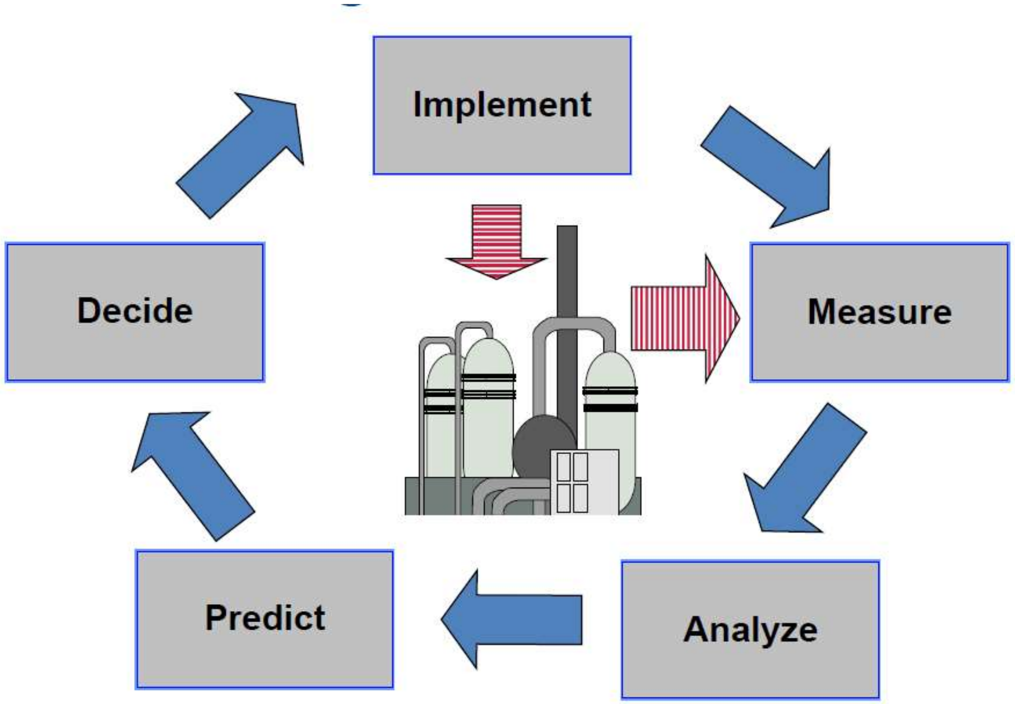
RESPONSIBLE
OIL AND GAS

- What Do We Mean by Digital Transformation ?
- Needs for process simulation ?
- Some new trends
  - Diagnosis
  - Performance models (Yields)
  - Product properties models
  - Deactivation models

"DIGITAL TRANSFORMATION" FOR REFINING AND PETROCHEMICAL PLANT OPERATIONS

RESPONSIBLE
OIL AND GAS

- Using modern digital/advanced technologies to provide improved plant safety, productivity, efficiency, and environmental performance
- Improve the elements of the plant decision cycles through increased use of
  - Measurements
  - And other data sources
    - Analytics
    - Models
    - …



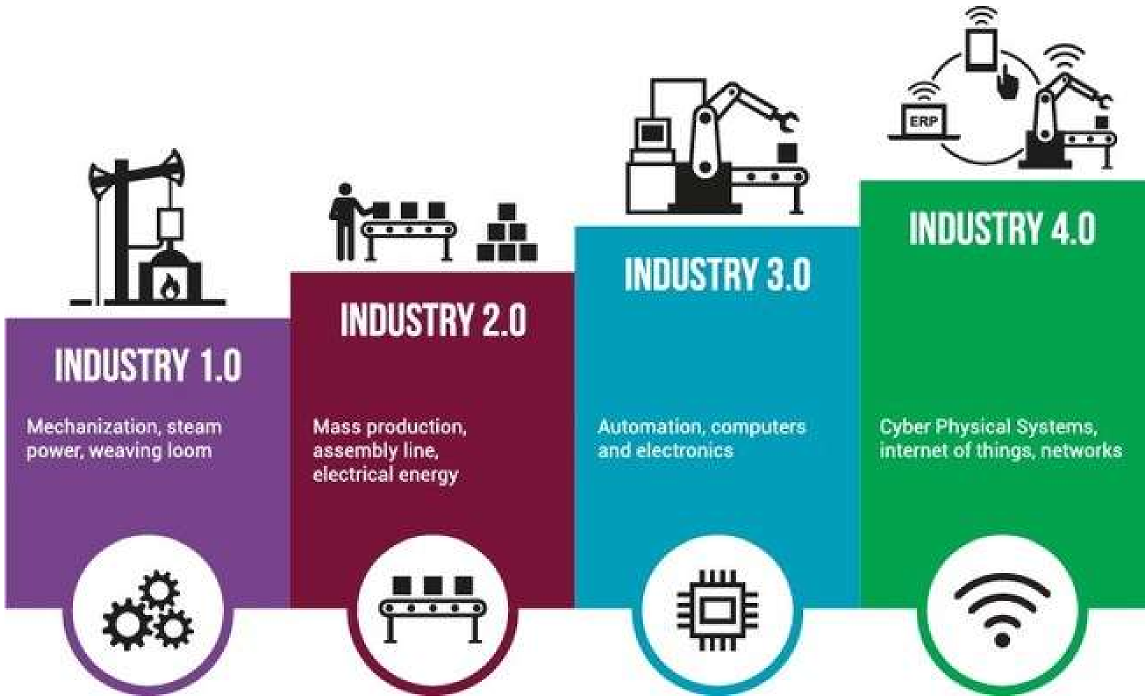
Typical Plant Decision Cycle

Doug White, Aiche Conference, Paper 224c: Digital Transformation In The Process Industries: Looking Back, Looking Forward

INDUSTRY 4.0 (FORTH INDUSTRIAL REVOLUTION)

- Industry 4.0 is a collective term particularly used in manufacturing to emphasize technologies and concepts of value chain organizations
- Related terms
  - Cyber-Physical Systems
  - The Internet of Things
  - Cloud computing
  - Digital Twin
- Although the term originates from the manufacturing industry, the elements of Industry 4.0 are relevant for most businesses (Maintenance 4.0, Safety 4.0, Ship 4.0,…)

*Focus on the **elements** rather than the term Industry 4.0 as such !*



IOT - INTERNET OF THINGS

RESPONSIBLE OIL AND GAS

- The Internet of Things (IoT) is the network of items embedded with ***electronics***, ***software***, ***sensors***, ***actuators***, and ***network connectivity***
- which enable these objects to ***connect and exchange data***

IoT is what we need to ***connect*** *(sensors, analyzers, …)*

CLOUD COMPUTING

RESPONSIBLE
OIL AND GAS

- Cloud computing is an information technology paradigm that enables ***access*** to shared pools of configurable system ***resources***
- In some presentations the term Internet of Services (IoS) rather than cloud computing

With cloud computing we do not need to think about platforms, how to connect etc

DIGITAL TWIN

RESPONSIBLE OIL AND GAS

- The digital twin refers to a digital ***replica*** of physical assets, processes and systems that can be used in real-time for control and decision purposes
  - Computerized mathematical model (what we have done over years)
  - Real-time, thanks to IoT
- In contrast to a physical asset, the digital twin can immediately respond to ***what-if*** inquiries

- -> Impact on developed models

AGENDA

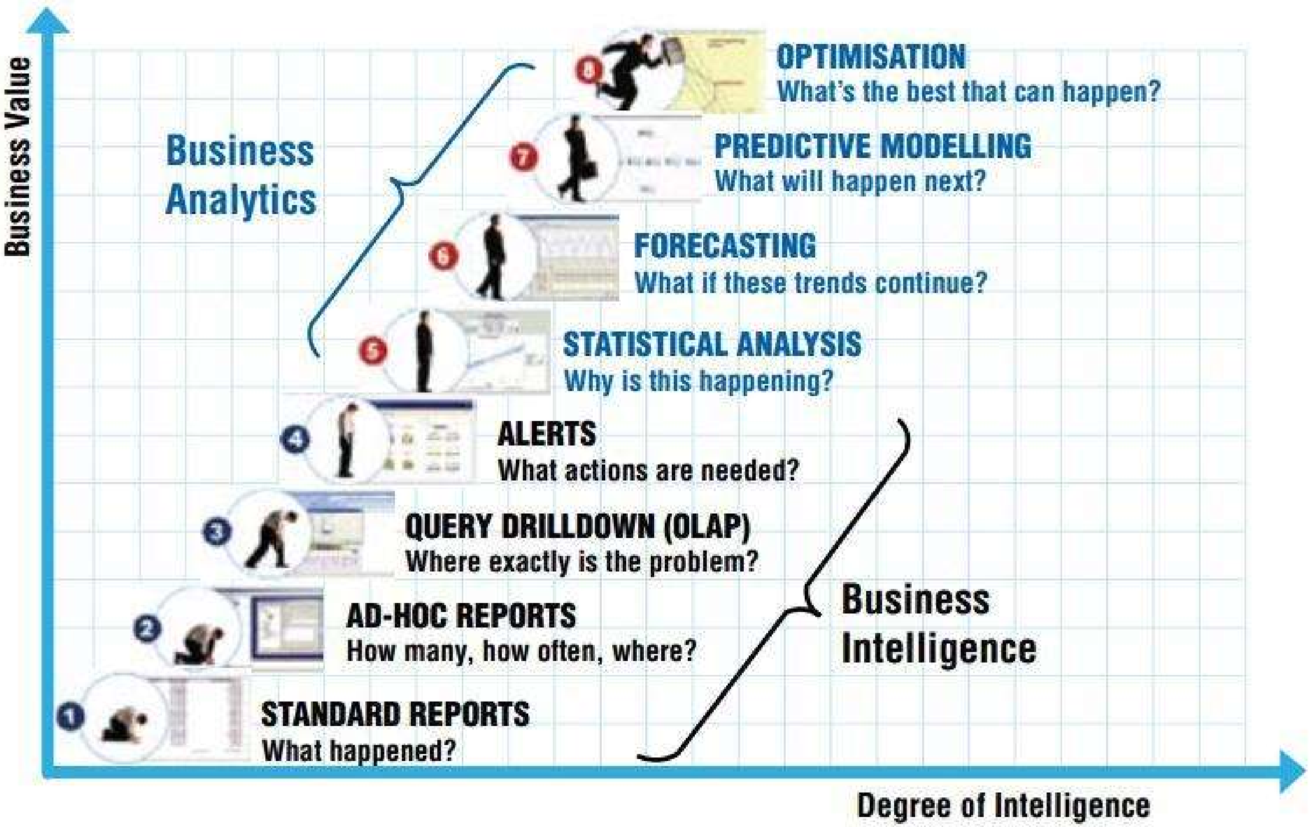
RESPONSIBLE
OIL AND GAS

- What Do We Mean by Digital Transformation ?
- **Needs for process simulation ?**
- Some trends
  - Diagnosis
  - Performance models
    - Yields
    - Product properties

REFINING

RESPONSIBLE OIL AND GAS

- **Where we are, where we want to GO ?**
- **Maintenance** Fault Detection and Isolation: Detect any problem and locate the root cause
- **Deactivation** When my cycle will finish ?
- **What If analysis** What happen if my feedstock changes, what happen if the flowrate increases ….
- **Real time optimization**



Key Point : Accurate data + Accurate models

NEEDS FOR PROCESS SIMULATION

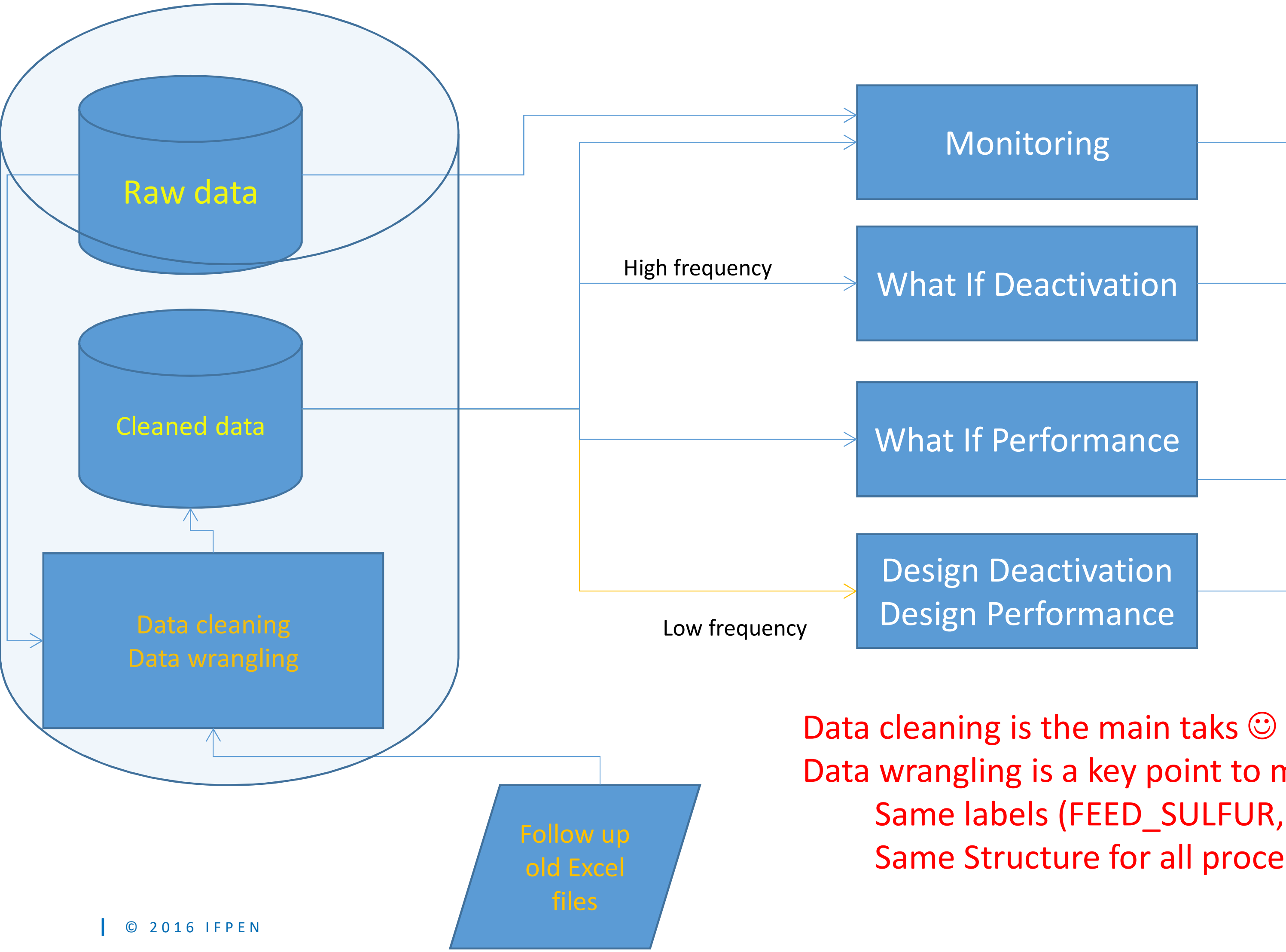
- Monitoring/Alerting models
- Design models
  - Design performances: predict Start of Run performances (yield, product properties…)
  - Design Deactivation: predict average deactivation during the whole cycle
- What if models
  - What If performances: what happen if I change a feedstock, a flow rate ….
  - What if Deactivation: predict impact of operating condition/Feed/Performances changes on cycle length ?

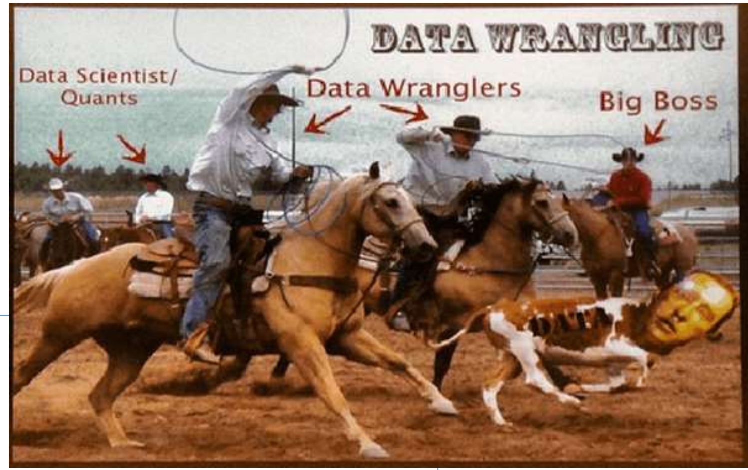
AGENDA

RESPONSIBLE
OIL AND GAS

- What Do We Mean by Digital Transformation ?
- Needs for process simulation ?
- **Some trends**
  - **Diagnosis**
  - **Models**
    - **Yields**
    - **Product properties**
    - **Deactivation**

GLOBAL ARCHITECTURE

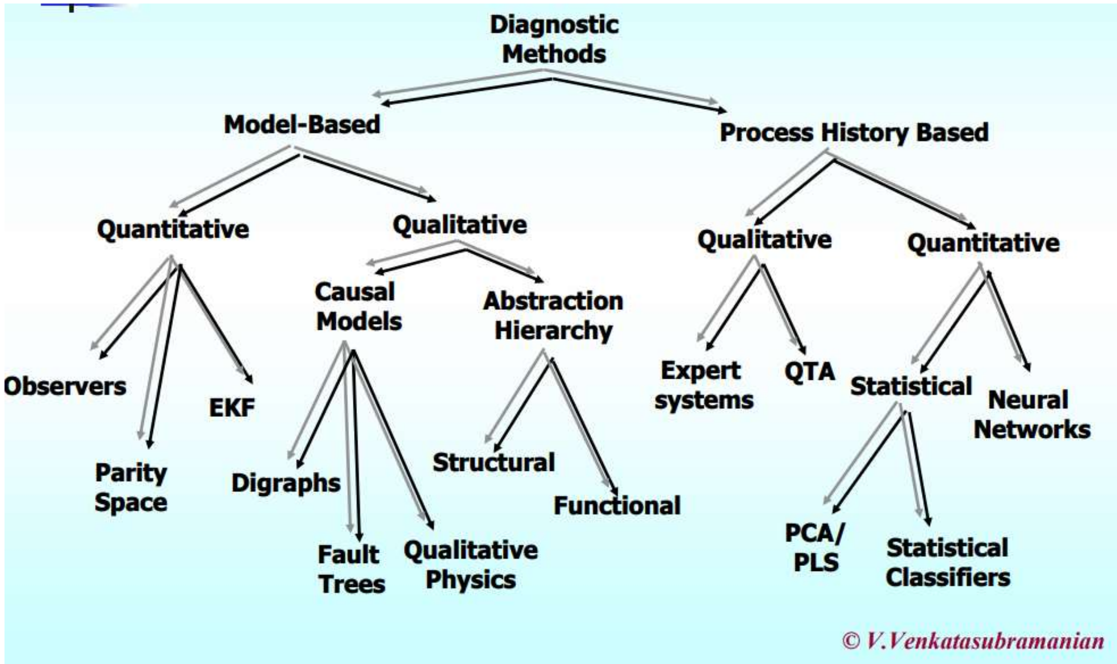




Data cleaning is the main taks ☺
Data wrangling is a key point to merge data for every processes
Same labels (FEED_SULFUR, AZUFRE_CARGA, Soufre_Charge……)
Same Structure for all processes (VGO_HDT, VGO_HCK…)

MONITORING

- Fault Detection and Isolation
  - Fault detection: Detect malfunctions in real time, as soon and as surely as possible
  - Fault isolation: Find the root cause, by isolating the system component(s) whose operation mode is not nominal
  - Fault identification: to estimate the size and type or nature of the fault
- FDI methods :
  - Model free methods (based on data)
  - Knowledge based methods
  - Model based methods
- Known techniques since 15 years !



V. Venkatasubramanian, R. Rengaswamy, K. Yin and S. N. Kavuri, "Review of Process Fault Diagnosis - Part I, II and III", Computers and Chemical Engineering, 27(3), 293-346, 2003.

MODELS

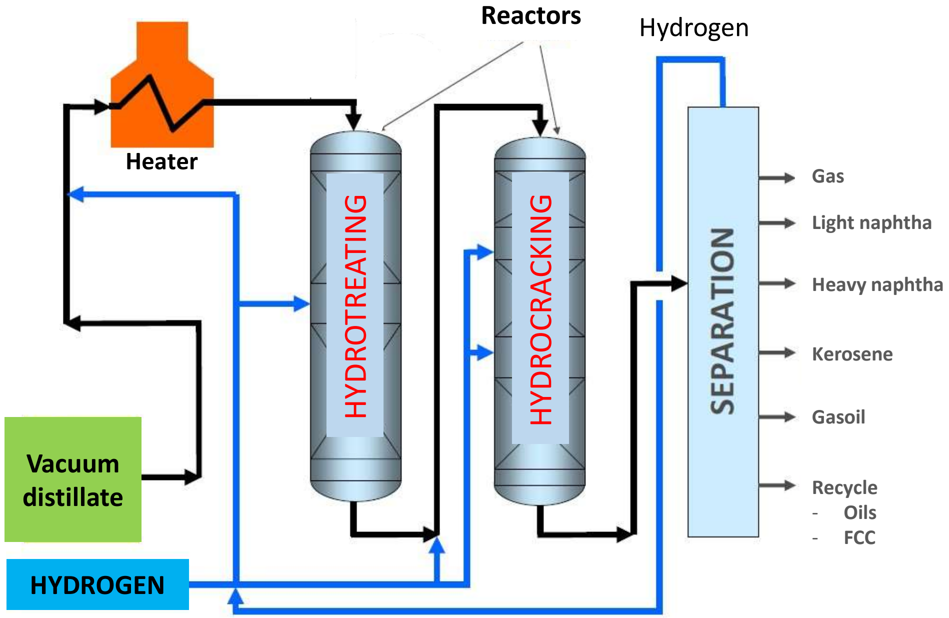
RESPONSIBLE OIL AND GAS

- Key points for digital twin in order to carry out what if analysis
- Design model
  - What will be the performances -> yield, product properties
  - What will be the cycle length -> deactivation
- What if model (during the cycle)
  - What will be the performances (yield, product properties) knowing the first part of the cycles
  - What will be the cycle length knowing the first part of the cycles
- -> Challenge: include online data in our models to update them

HYDROTREATING & HYDROCRACKING

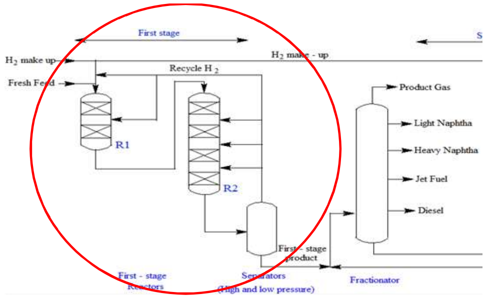
RESPONSIBLE
OIL AND GAS

- Hydrocracking process (HCK): fixed bed catalytic process with H_2 consumption

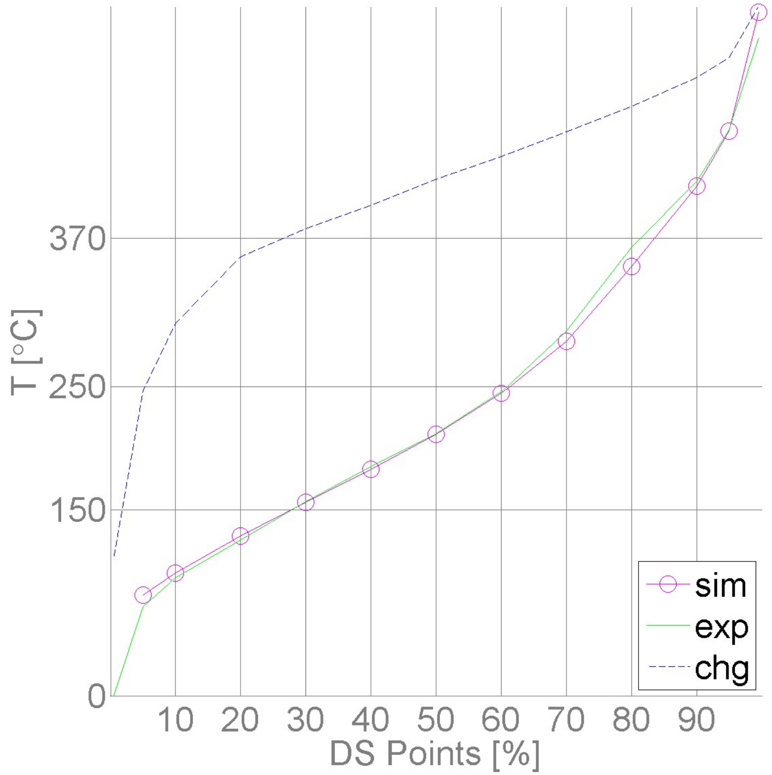


- Using one or several catalysts ($NiMoS_2/Al_2O_3$ for HDT; $NiWS_2/Al_2O_3$-SiO_2 or $NiWS_2$/zeolite for HCK), high pressure (50-180 bar), moderated temperature (350-430 °C)

HYDROCRACKING R1 + R2

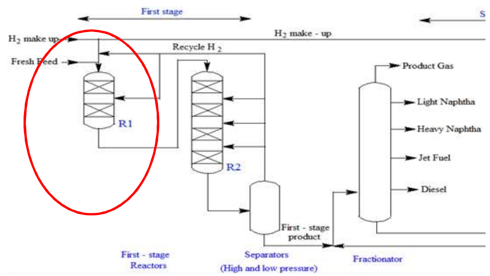


- Main Measurements
  - Feed, Intersampling reactor, Effluent
    - Density
    - SIMDIS
      - yield structure (Naphtha, KERO, GO, RS)
      - Gas, where applicable
    - Nitrogen and Sulfur Content
    - Aromatic Carbon Content
  - Operating Conditions
    - Temperature [°C]
    - LHSV [h-1] (holdup time)
    - Ptot, ppH2, ppH2S, ppNH3 [bar]



 |

HDT MODEL



RESPONSIBLE OIL AND GAS

- Several models
  - HDN -> estimate operating conditions in order to reach the N slip target (less than 10 ppm)
  - HDS, HDA -> strong impact on H2 consumption
- Very difficult task
- Classical model[1]
  - Expertise to estimate the main descriptors

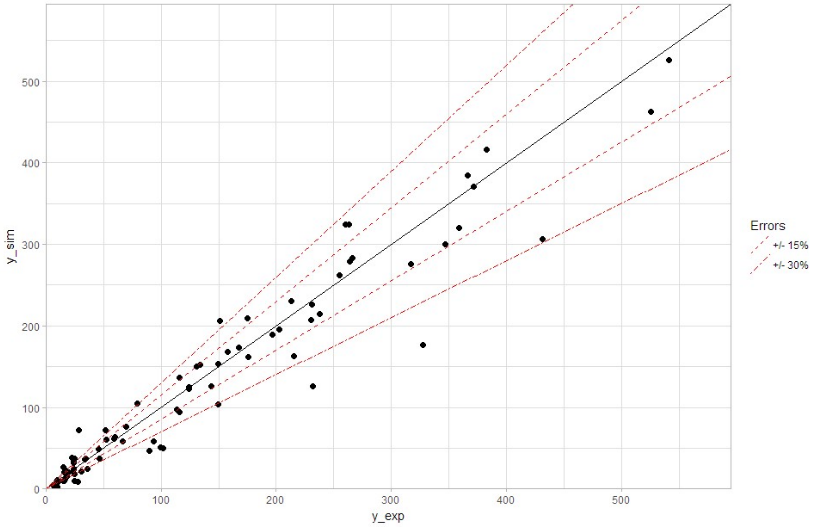
$$\frac{dy}{dt} = -k_0 exp\left(-\frac{Ea}{R.T}\right) ppH_2{}^m . y^n . f(Feed\_Descriptors)$$

  - Similar form for HDS and HDA
  - Conversion modelization -> Continuous lumping[2] (cf. next)
  - Experimental design can be used to decrease the required number of experimental points[3]
  - Machine learning techniques can also be used[4]
    - Less Robust ☹

(1) Celse et al., Chemical Engineering Journal 278 (2015) 469-478
(2) Becker et al. , Fuel, Vol 139, 2015, pp. 133-143
(3) Celse et al., International Journal of Chemical Kinetics, Wiley, 2016, 48 (11), pp.660-670
(4) Celse et al., Aiche Conference, San Antonio (2017)

RESULTS AND FUTURE IMPROVEMENTS

- Results



- Future improvements
  - How to add catalytic descriptors to reduce required time to develop new models?
  - Integrate feed inhibitors by using machine learning techniques

$$\frac{dy}{dt} = -k_0 exp\left(-\frac{Ea}{R.T}\right) ppH_2{}^m . y^n . f(Feed\_Descriptors)$$

IFP Energies nouvelles

NEW TRENDS FOR PERFORMANCES MODELS

RESPONSIBLE
OIL AND GAS

- Pure machine learning methods (N = $M_{\text{machine Learning}}$(Feed, COP))
  - ➢ Tested at IFPEN, Results not so bad but problems with Extrapolation + constraints (N > 0). Regularization should be used[1].
- Using Machine Learning to fit correlative terms in the kinetic models

  (N = M_{kinetic} (Feed, COP, $M_{\text{machine Learning}}$(Feed, COP))
  - ➢ Most “simple” approach: no modification of the basic model structure required
  - ➢ More confidence in model parameters, but phenomena not captured by the model not accounted for
- Using Machine Learning to determine correction terms to be applied to the simulation outputs from kinetic models (N = Mkinetic o Mmachine Learning)
  - ➢ Retaining basic structure of the kinetic model(s)
  - ➢ Allows poorly understood phenomena to be described
- Constructing multiple kinetic models (based on different proposed physical mechanisms) and apply Machine Learning algorithm(s) to determine an aggregated model

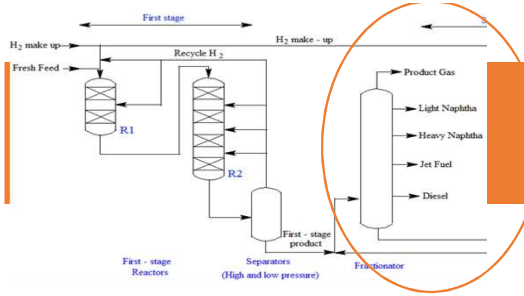
  (N = M_{Agg} ($M_{\text{kinetic,1}}$, …, $M_{\text{kinetic,n}}$, $M_{\text{ML,1}}$, …, $M_{\text{ML,n}}$)
  - ➢ Ensemble methods, See https://mlens.readthedocs.io/en/0.1.x/

➢ **Key point: How to deal with deactivation[2] ?**

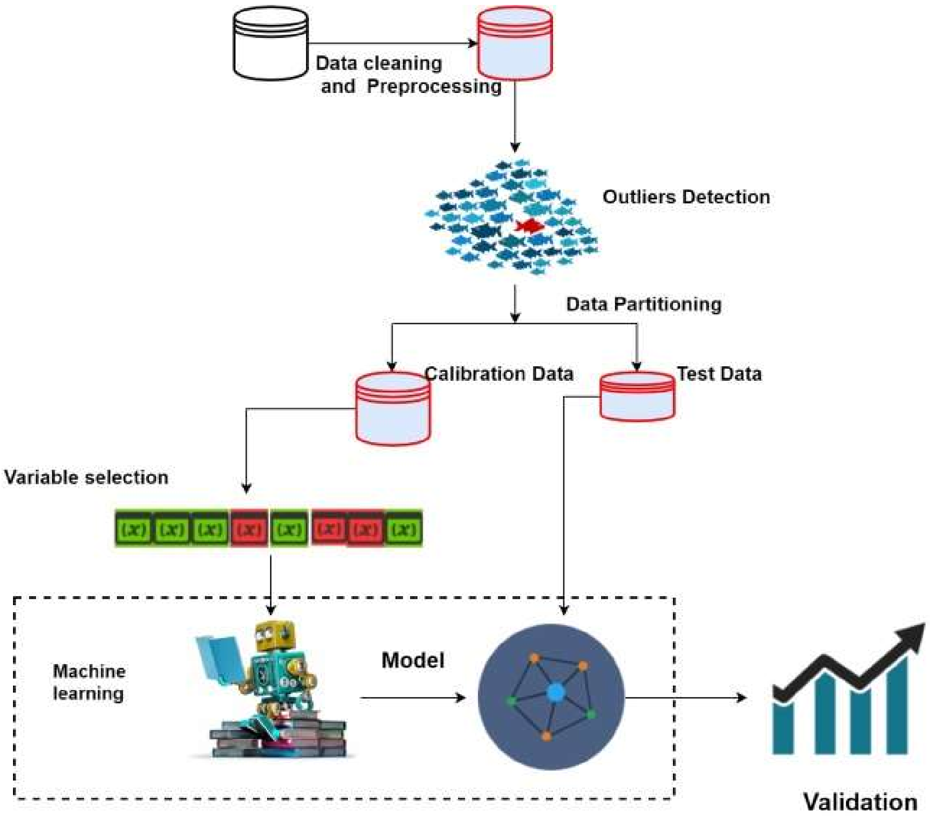
(1) A. Karpatne et al., *“Theory-guided Data Science: A New Paradigm for Scientific Discovery from Data”*, Computer Science and Engineering, 2017
(2) E. Iplik, Hydrocracking: A Perspective towards Digitalization, Sustainability 2020, 12, 7058; doi:10.3390/su12177058

IFP Energies nouvelles

NEW TRENDS FOR PRODUCT PROPERTIES PREDICTION



- How to take advantage from digitalization ? (i.e. more points)
- At least, use data from product properties prediction



PRODUCT PROPERTIES PREDICTION

- Which outlier detection method ? -> LOF[1]
- Which partitioning method ? -> Kennard & Stone
- Which variable selection ? -> leaps[2]
- Which model ?
  - Machine learning model for product properties
    - Linear
    - Random forest[3]
    - XG Boosting[4]
    - Support vector regression[5]
    - Kriging (Gaussian Process)

(1) Ding, H.et al. (2018). Solar Energy, 164(July 2017), 139–148.
(2) Furnival et al. (1974). Technometrics, 16(4), 499–511.
(3) Breiman L. Machine Learning. 45 (1), pages 5-32, 2001.
(4) Friedman, J. H. (2001). The Annals of Statistics, 29(5), 1189–1232.
(5) Vapnik, V. N. (2000). *The Nature of Statistical Learning Theory*

RESULTS

RESPONSIBLE
OIL AND GAS

diesel statistics (Training set)

Linear Model is the worst model

	Linear Model	Universal Kriging	Random Forest	Gradient Boost Method	SVR
Nb points	1432	1432	1432	1432	1432
MAE	0.0035	0.0000	0.0026	0.0026	0.0020
RMSE	0.0051	0.0004	0.0041	0.0038	0.0034
Pts +/-0.01 (%)	93.9	100.0	96.5	97.3	98.1
Pts +/-0.005 (%)	78.7	99.9	86.0	87.3	92.7
Pts +/-0.0025 (%)	51.8	99.6	65.6	63.1	78.8

diesel statistics (Test set)

Indicators are similar to the one obtained on the training set

	LM	UK	RF	GBM	SVR
Nb points	626	626	626	626	626
MAE	0.0035	0.0022	0.0024	0.0029	0.0023
RMSE	0.0050	0.0034	0.0038	0.0042	0.0034
Pts +/-0.01 (%)	93.9	97.9	97.0	95.8	97.9
Pts +/-0.005 (%)	78.1	90.1	86.9	83.2	89.5
Pts +/-0.0025 (%)	51.6	72.4	68.5	59.6	69.2

NEW TRENDS FOR DEACTIVATION PREDICTION

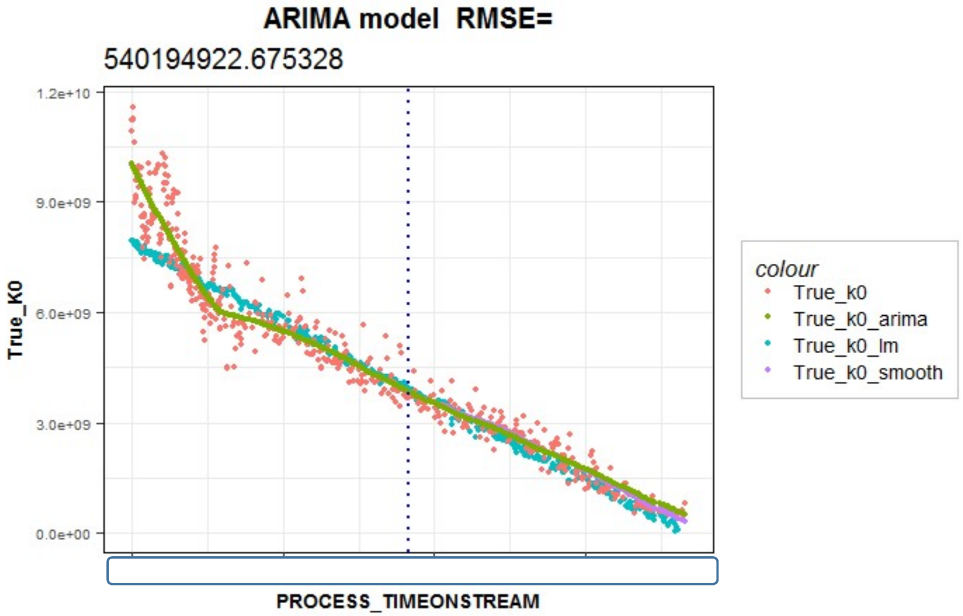
RESPONSIBLE OIL AND GAS

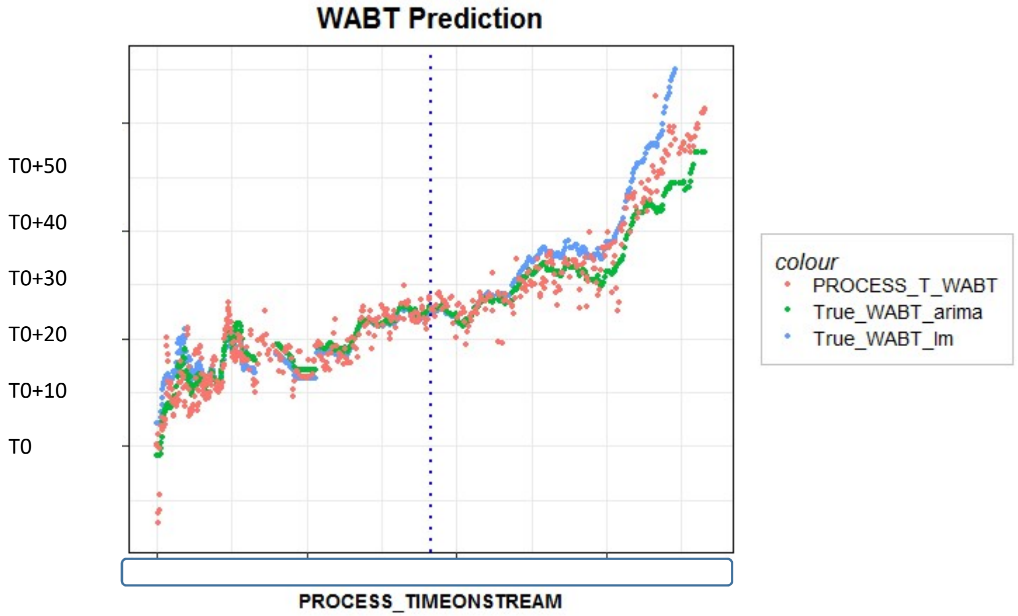
- Aim: Forecast WABT evolution knowing current part of the cycle
- Training data set (others cycles + current part of the cycle)
    - Remove the outliers
    - Estimate catalyst state : k0 using classical kinetic model (remove operating conditions & feed effect)
    - Fit a machine learning model on k0 based on several variables (feedstock, operating condition ….)
    - Fit an ARIMA (Autoregressive integrated moving average) model based on machine learning model covariate
- Test Data Set
    - Forecast k0 using Arima
    - Predict WABT using kinetic model and the previous k0

DEACTIVATION APPLICATION

RESPONSIBLE
OIL AND GAS

- Good prediction on WABT Prediction

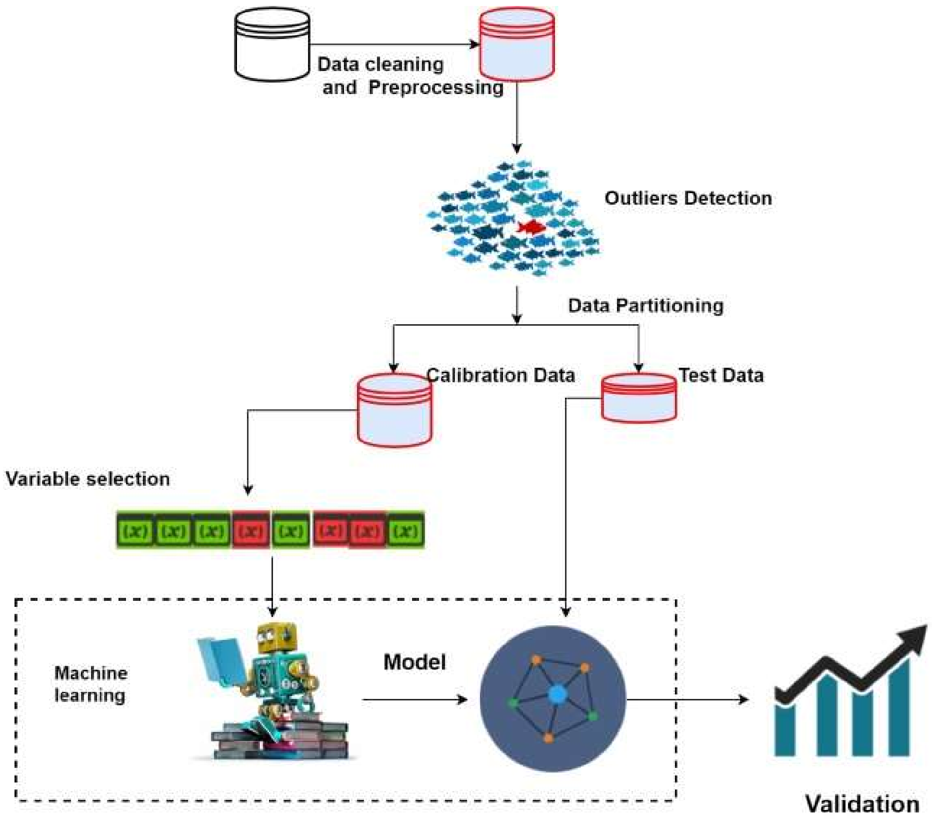




CONCLUSION

RESPONSIBLE
OIL AND GAS

- New trends for process simulation:
  - Design and what If models updated very often
  - Several kind of models for what if analysis
    - Performances: Hybrid model (kinetic + machine learning)
    - Product properties : machine learning techniques
    - Deactivation: When my cycle will finish ? Should I change the kind of feedstocks
      - Estimate catalyst state using kinetic models
      - Forecast using machine learning techniques
- What if models should be updated at high frequency (at least once a month) which is possible thanks to
  - Industry 4.0
  - Machine learning for Outliers Detection…
- This can be achieved if and only if data management is optimal



QUESTIONS?

RESPONSIBLE
OIL AND GAS

Benoit.celse@ifpen.fr