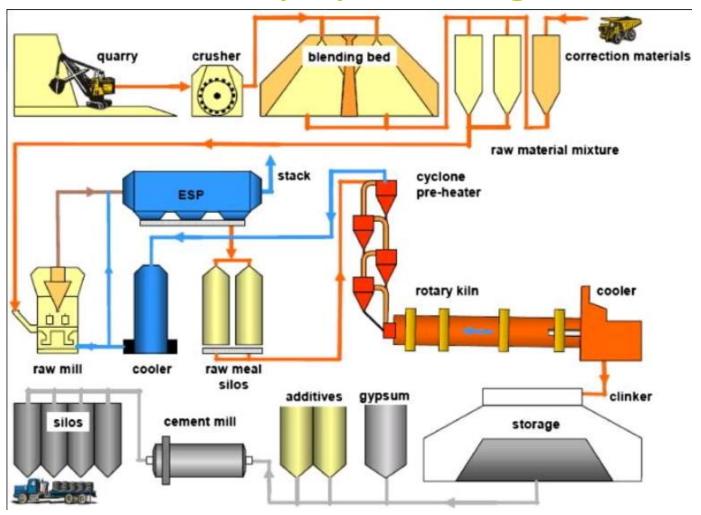
# Real Time Optimization of the cement production process

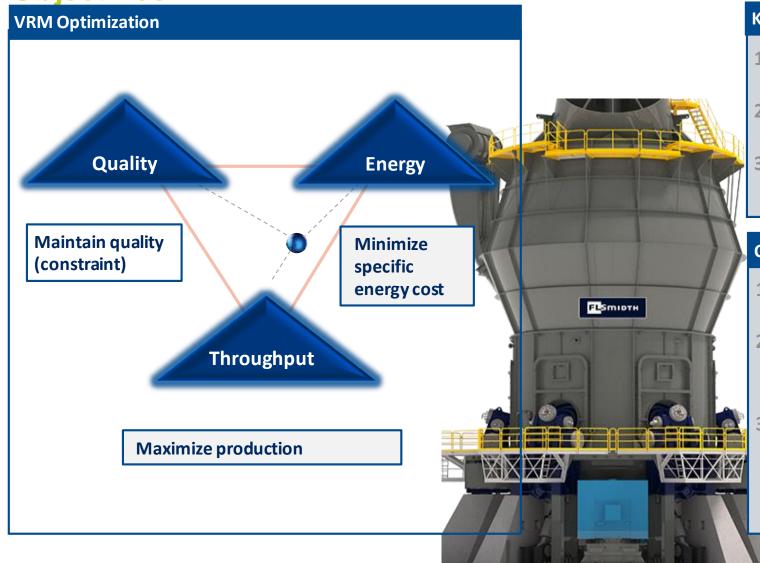
## The cement production process

#### **Cement Plant simple process diagram**



- Raw Mill is the equipment used to grind raw materials into "rawmix" during the manufacture of cement
- 2. Rawmix is then fed to a **Kiln**, which transforms it into clinker
- 3. The **Cement Mill** grinds the hard, nodular clinker from the cement kiln into the fine grey powder that is cement

**Objectives** 



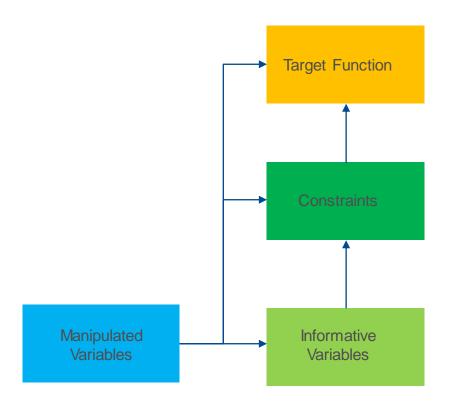
#### **Key Targets:**

- **1. Maintain Quality:** Minimize standard deviation of quality KPIs
- **2. Throughput (Feed Rate):** Increase mill productivity in tons/hour
- **3. Energy:** Minimize specific energy cost for given throughput

#### **Constraints to be considered:**

- **1. Quality:** Maintain material fineness standard deviation at target levels
- **2. Throughput:** Maintain and ideally reduce the unscheduled shut downs due to operational reasons
- 3. Energy: Maintain or increase mill operating time during off-peak hours (with lower energy cost), minimize the operation of mill during peak hours

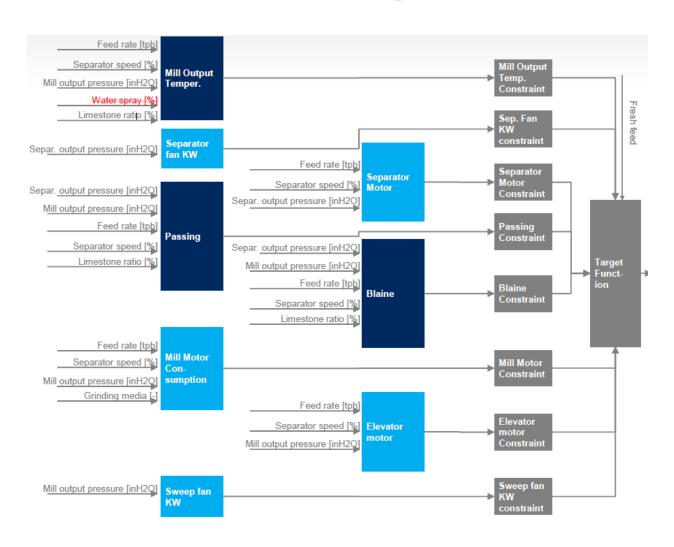
#### **Composite Model design**



An RTO should be able to suggest at specified time the values of the manipulated variables that maximize/minimize our target function keeping the operational constraints that the plant has set

In this optimization problem we use machine learning in order to predict the outcome of key variables according to given operating conditions

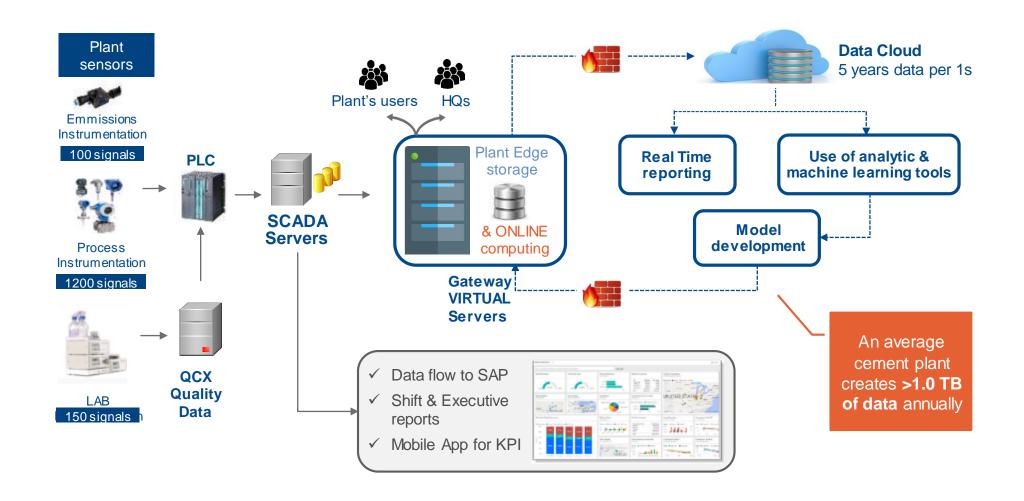
#### **Composite Model design**



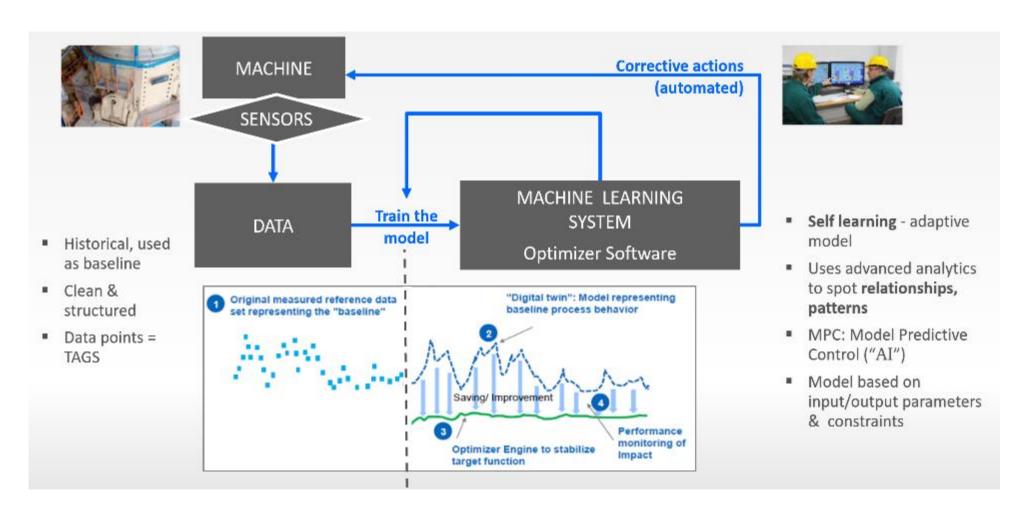
An RTO should be able to suggest at specified time the values of the manipulated variables that maximize/minimize our target function keeping the operational constraints that the plant has set

In this optimization problem we use machine learning in order to predict the outcome of key variables according to given operating conditions

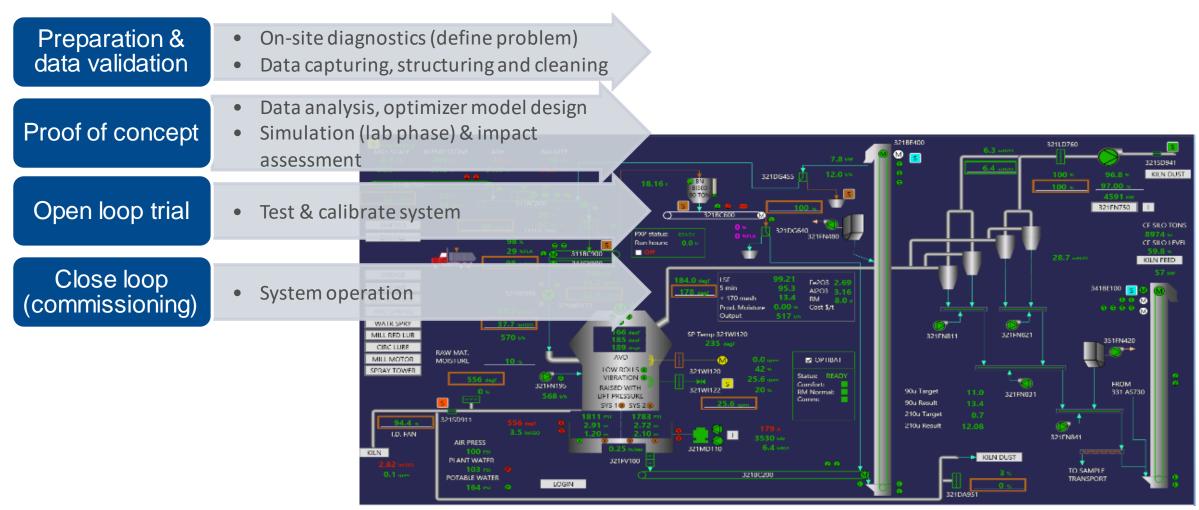
## **Upgrading our Infrastructure & Data Management**



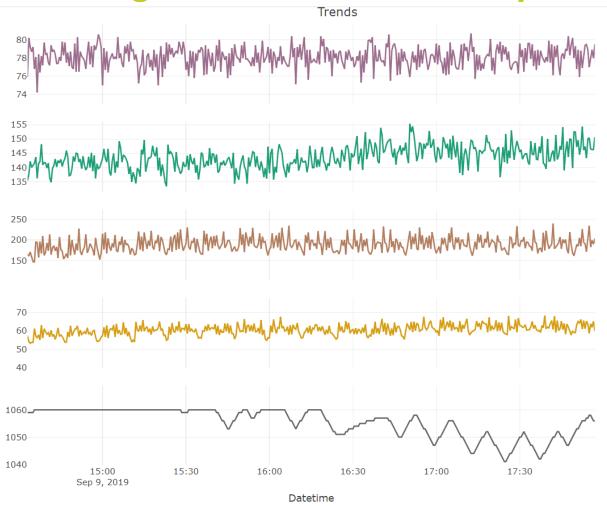
How does it work: use of AI in a machine learning system



**Project implementation steps / methodology** 



How a ML algorithm learns from noisy data?

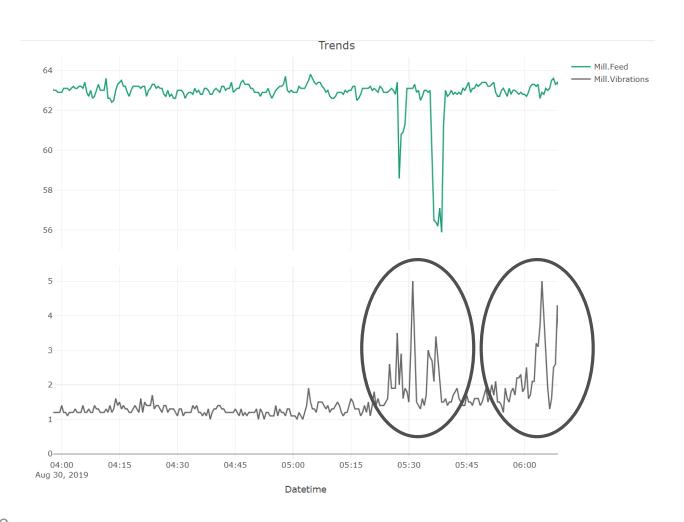


Most sensor data are noisy variables with high SD even on stable operating conditions.

**-** F10IT269

Appropriate data preprocessing is needed in order to smooth the data without loosing important information.

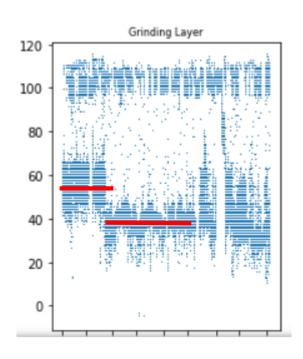
#### Do we really need data on such a high granularity?

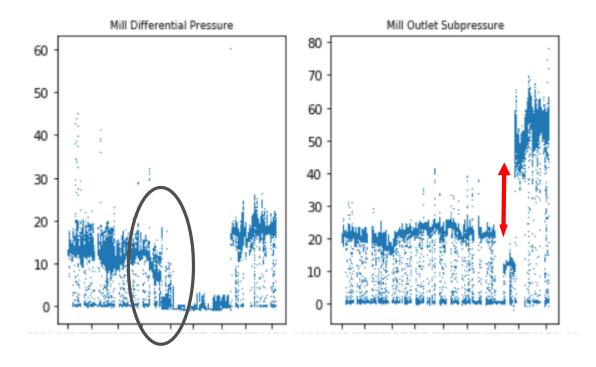


Vibrations can cause mill stoppages resulting in high downtimes.

Vibrations can occur in less than a minute, it is crucial an RTO to be able to predict and avoid them.

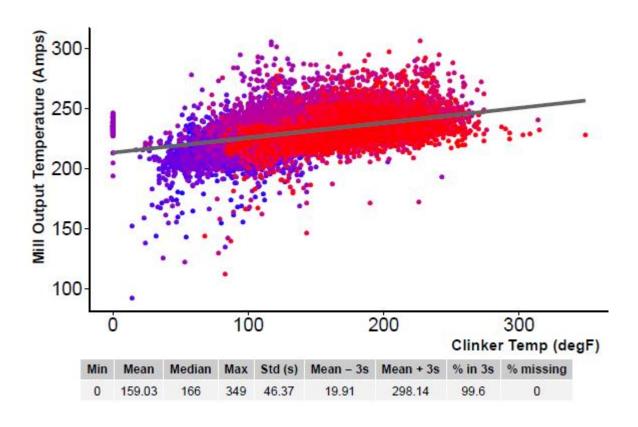
#### Are your data reliable?





Sensors may malfunction at spontaneous times or need maintenance and recalibration. Data quality checks should be done not only before model training but also when RTO is in operation.

#### Remove the outliers! Or not?

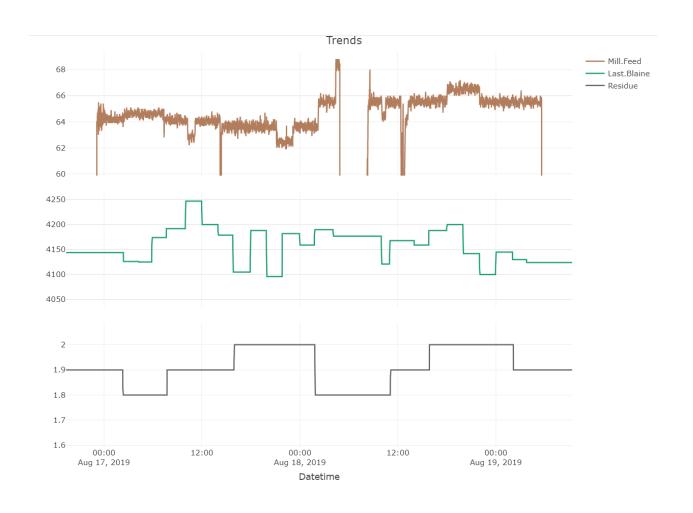


Outliers usually can harm you ML algorithm.

However a plant operates most of the time in the same conditions generating data in a specific space.

Can these extreme cases help your algorithm learn the real relationships between the variables or they are abnormal operating conditions that you cannot model?

#### How do you handle lab measurements?

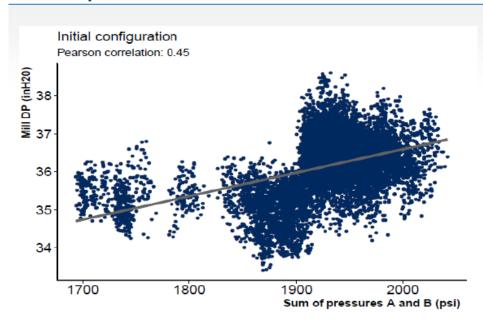


Blaine and Fineness are the most important quality characteristics of the end product.

Blaine and Fineness are measured in the lab from samples taken from the mill usually every 1-2 hours.

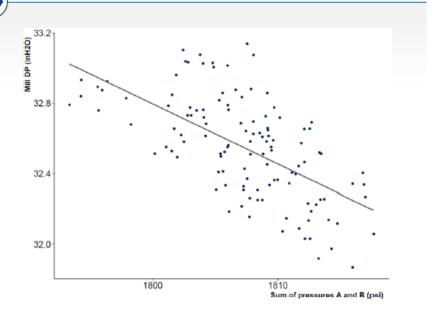
#### Are the correlations you observe correct?

From a positive correlation...



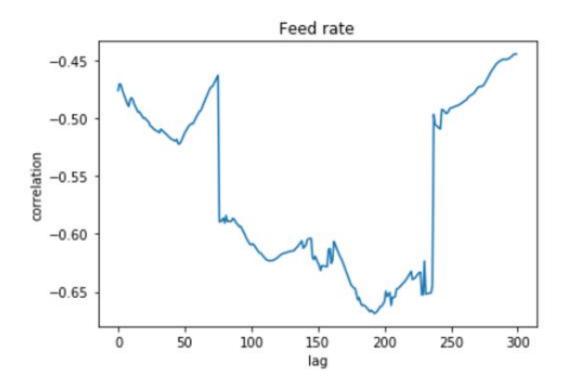
Looking at the full dataset, the sum of pressures seemed to have a positive correlation with mill DP, which was in the opposite direction of process knowledge

...To a negative one



Through isolation of representative data periods, the team was able to confirm the process knowledge and recreate the expected variable relationships<sup>1</sup>

## Synchronize your signals!



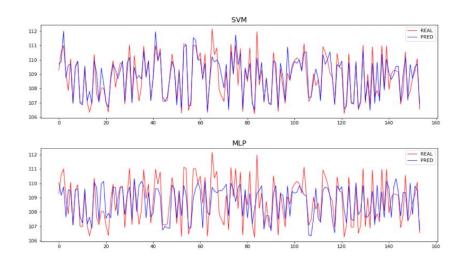
The material we put on the mill needs significant time to become end product.

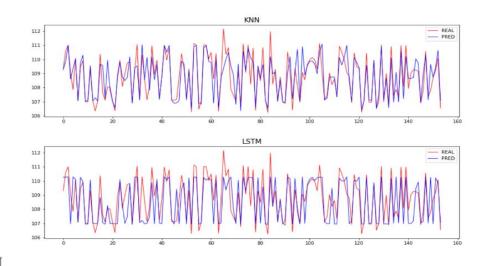
E.g. the blaine measurement of a sample we collect at time t is a result of the feed rate at time t-n.

It is important to estimate as accurately as possible these time delays om the variables in order to get meaningful correlation between them.

#### Use of machine learning algorithms

•Below we can have an example of Mill Exit Temperature block with the predicted versus the actual values graphs, along with the metrics that will help us to identify the best algorithm:



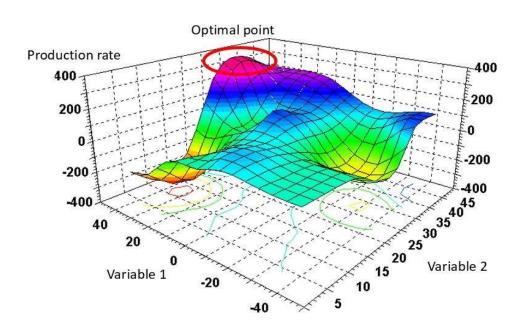


CPII

ALGORITHM	MAPE	MASE	MAE
SVM	0.2325	31.6868	0.252
MLP	4.9349	666.8841	5.3026
KNN	0.2634	35.8739	0.2852
LSTM	0.9176	8925.0351	0.8186

#### **Optimal vs fast solution**

The RTO is designed to provide values for the manipulated variables every 30 seconds.



The choice of the ML and the optimization algorithm is done taking into account this constrain

An ensemble model may give accurate results but an MLP can make predictions really fast.

A genetic algorithm can avoid local maxima but brute force on a constrained search space may be also sufficient.

#### **Summary**

- We do spent 90% of our time cleaning and preparing our data
- We use and test several approaches but we select the one that satisfy the business needs
- We collaborate closely with our automation engineers and process experts

