

Data-driven CFD simulation of an industrial semi-batch mixing process

C. Alfaro*, I. Viejo*, S. Izquierdo*
Corresponding author: sizquierdo@itainnova.es

*Instituto Tecnológico de Aragón (ITAINNOVA), Zaragoza, Spain

Abstract: We compare three approximations for predicting the viscosity of a material resulting from blending a main raw material and additives in an industrial semi-batch mixer. These approaches are: a data-based model build from experimental data collected in a real process; a theory-based model build by means of an integral analysis of transport equations describing the mixing process; and CFD simulations of this process. The comparison leads to a model proposal in which the selection of principal components for building the data-driven model is guided by the theory-based model. This theory-based model is enriched with results from the CFD simulations, which allows an efficient and accurate inverse analysis for fitting experimental data and numerical models. The proposed model can predict, with a minimal amount of process information and on-line, the final viscosity with the same uncertainty of the current experimental off-line quality approach.

Keywords: data mining, mixing, multifidelity.

1 Introduction

Mixing operation efficiency is typically critical to the successful manufacturing of a product. At industrial scale, efficient mixing can be difficult to achieve but failing in providing the necessary mixing could lead to significant increase of manufacturing costs[1]. That is why industries are investing in gaining a deeper control and understanding of mixing processes by installing sensors that provide more information along the process.

Thanks to sensor technology advances, data acquisition and storage is cheap and data is becoming much more abundant and accurate, therefore data-driven methods are gaining attention in the field of scientific computing versus model-driven methods [2], [3]. But data can be very often distorted by noise and its availability is limited depending on sensor system design and process characteristics.

On the other hand, progresses in computer-aided engineering (CAE) are making possible to implement models that simulate complex processes, like multiphase flows. But these models usually are not in full agreement with physical reality. This gap between experimental and computed results is due to either simplifications of the model, made to avoid unaffordable computational costs, or to other physical phenomena occurring during the process of whose relevance (or existence) the model was unaware of.

In finding the best blending among data-driven models and CAE models arises the opportunity to overcome each approach's limitations. Aiming at discovering this optimal way of combining data-based and model-based estimations, in this work three different approaches are compared,

focusing on improving data-based model. The interest in this analysis is caused by the widely use of mixing as unit operation in chemical process engineering [4], and the possibility of applying the described method to analyze additional unit operations.

This work has been carried out using real process data to take noise into account the way it is presented in actual working conditions. Experimental data is provided by a private company; therefore most of experimental values are not displayed, or are normalized when it is required.

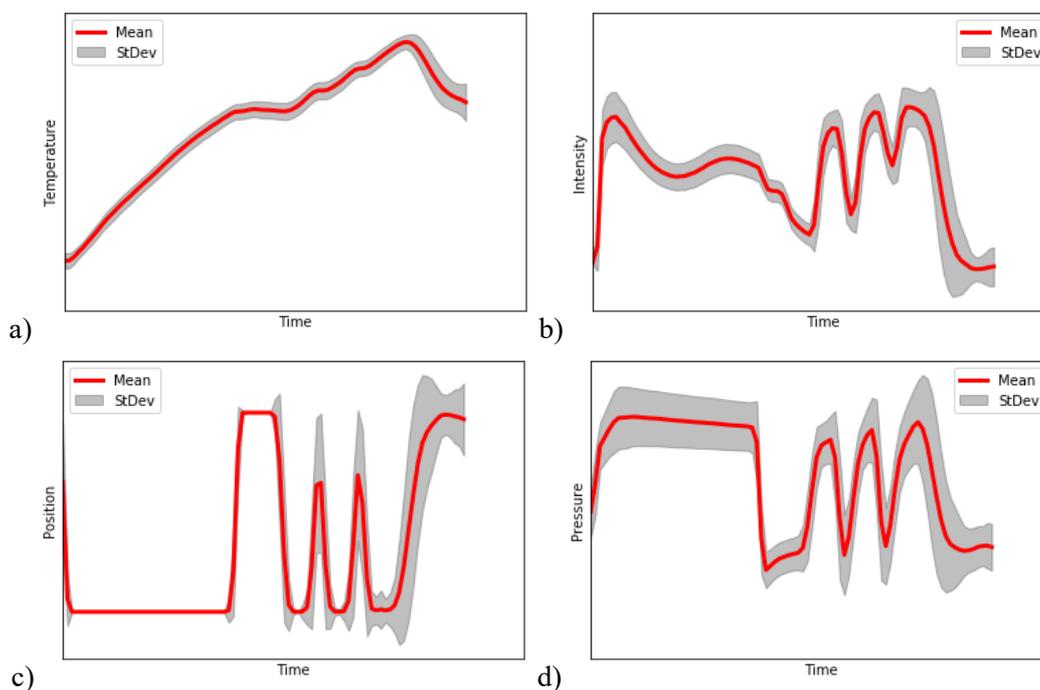
The present paper includes: description of data analysis and model methods (section 2); main results of the different models and comparison among them (section 3); and final conclusions of the work (section 4).

2 Methods

2.1 Data-based model

The aim is to build a data model of mixing process. Data-based model intends to relate viscosity of material after processing to measured process variables. Viscosity is represented by ML. Data was collected from approximately 1500 mixing cycles. For every cycle, mixing data includes several process variables (every 3.0 or 0.125 seconds, depending on the sensor), ordered/served compound fed to the mixer, and processed material tests, carried out in a moving die rheometer. Data cleaning eliminates the cycles where data recording was noticeable mistaken (like empty recording periods or test's results out of rheometer limits).

Among the recorded process variables, intensity, pressure, rotor speed, mixing temperature, cylinder intensity and position are selected.



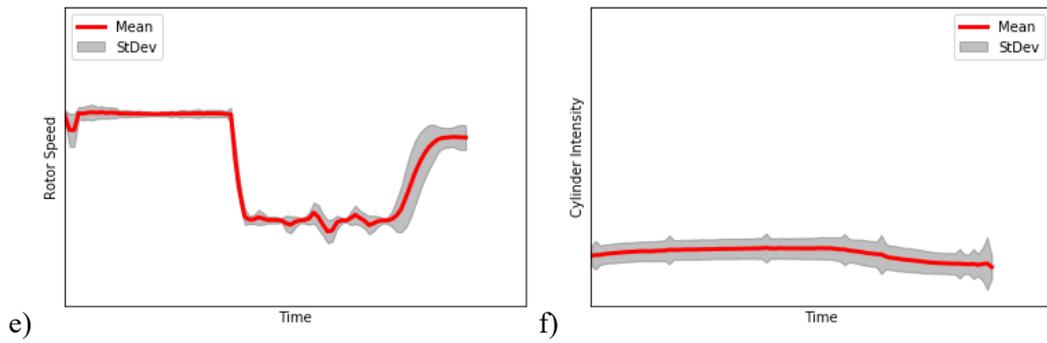


Figure 1: Mean and standard deviation of measured data:

a) Temperature, b) Intensity, c) Position, d) Pressure, e) Rotor speed, f) Cylinder intensity

They build a tensor of shape (variable, lecture, cycle), which is flattened in x-axis; therefore, it is turned into a matrix of shape (cycle, variables x lectures).

The served weight of each compound in each cycle is also considered. Together with flattened tensor of process variables, they constitute the input for the data model.

From the rheometer results of curing curve of processed material of each cycle, ML is chosen as output, since it is related to viscosity of the compound heated to vulcanization temperature.

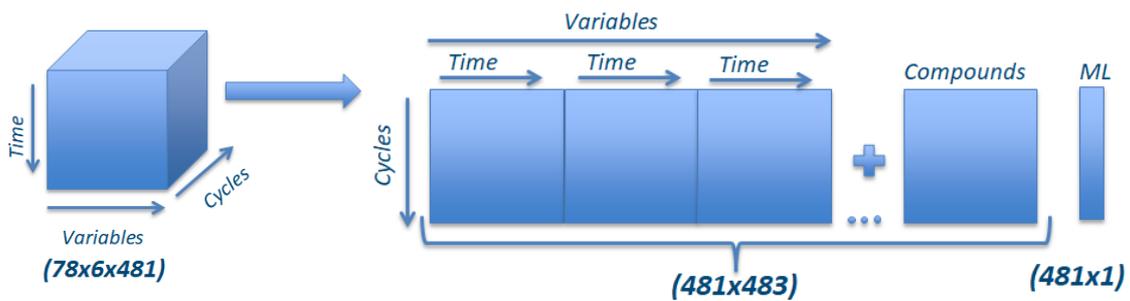


Figure 2: Input to data-based model

The input data matrix's shape is 481x483. To reduce the dimensions of the problem, Principal Component Analysis (PCA) is applied. As previous step, optimal number of reduced components is studied, regarding not only explained covariance but also the coefficient of correlation of a simple linear regression.

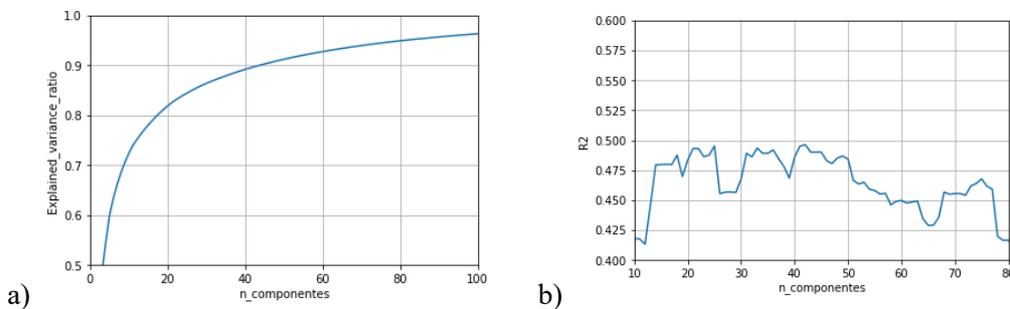


Figure 3: a) Sum of reduced components explained variance ratio vs number of reduced components; b) Coefficient of correlation (R^2) of linear regression of reduced components model vs number of reduced components

It is observed that highest coefficient of correlation is in the range of 15-50 reduced components.

Less components do not contain enough information (<80%) and including more components means introducing noise from the process into the reduced model. Considering explained covariance, coefficient of correlation and computational cost, 21 is selected as optimal number of reduced components for PCA in this case.

Reduced components are ordered by explained covariance; in this case, the first two reduced components contain approximately 40% of the unreduced information. When representing these two components for each cycle, it is observed that cycles clearly gather in two clusters, which are also grouped in periods of time.

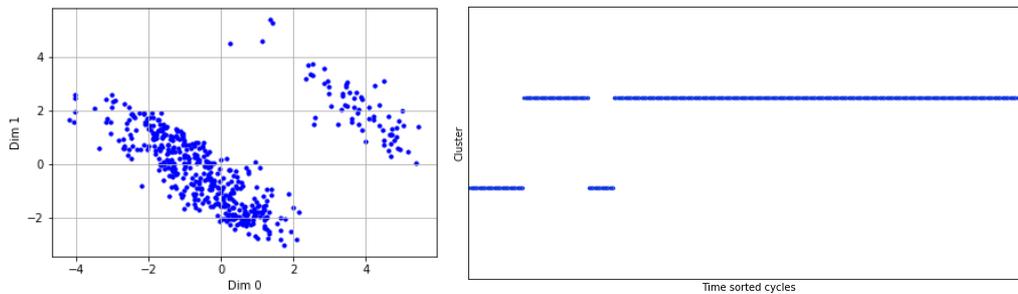


Figure 4: Division of cycles in two clusters

2.2 Theory-based model

The aim is to develop a system of ordinary differential equations that simulate and predict process variables's evolution in the mixer.

The model mixer is composed by a single cylindrical cavity of radius R_w and height L , representing one of the chambers of the Banbury mixer. The model mixer dimensions are chosen to make the volume of the model half of the total fluid volume in the real mixer (real mixer dimensions are unknown). In this approach, it is assumed that there is no free volume inside the mixing chamber. The characteristic, spiral-shaped, blades of the Banbury mixer are replaced in the model mixer by two simple radial impellers of radius R_i , height L and negligible thickness.

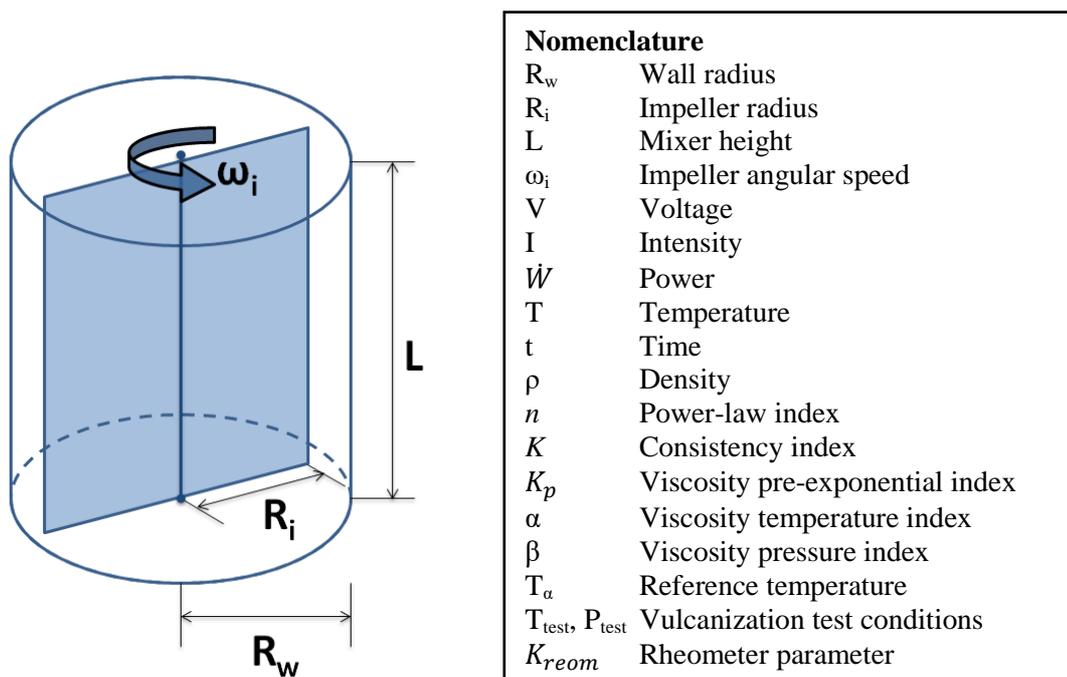


Figure 5: Model mixer

It is considered that fluid within the impeller radius rotates at the same speed than the impeller, ω_i . Fluid velocity between impeller's tip and wall is calculated solving Navier-Stokes equation and imposing no slip condition on the wall.

$$\omega(r) \begin{cases} \omega(r) = \frac{\omega_i R_i R_w}{R_w - R_i} \frac{1}{r} - \frac{\omega_i R_i}{R_w - R_i} & R_i < r \leq R_w \\ \omega(r) = \omega_i & 0 \leq r \leq R_i \end{cases} \quad \text{Eq. 1}$$

The momentum's conservation law shows that the system's power consumption is due to:

$$\dot{W}_{\text{input}} = VI = \dot{W}_{\text{rotor}} + \dot{W}_{\text{friction}} \quad \text{Eq. 2}$$

The power consumed by rotor speed is derived from the variation of kinetic energy in the mixer:

$$\dot{W}_{\text{rotor}} = \frac{\pi L \rho}{6} R_i^2 R_w (2R_i + R_w) \omega_i \frac{d\omega_i}{dt} \quad \text{Eq. 3}$$

On the other hand, the power dissipated by viscous effects is obtained by integrating the shear stress in the control volume:

$$\dot{W}_{\text{friction}} = \pi L R_k K \omega_i^{n+1} \quad \text{Eq. 4}$$

$$R_k = R_i^2 + \left(\frac{R_i}{R_w - R_i} \right)^{n+1} (R_w - R_i)^2 \quad \text{Eq. 5}$$

Parameter K is the flow consistency index of power law viscosity model. It varies with temperature and pressure, according to an exponential Arrhenius and exponential relation, respectively, as follows:

$$K = K_p e^{\alpha \left(\frac{1}{T} - \frac{1}{T_\alpha} \right) + \beta P} \quad \text{Eq. 6}$$

The correlation among process variables is hereby presented and every variation in any of them would cause a response in the power consumption in the form of intensity change.

$$V \frac{dI}{dt} = \frac{d\dot{W}_{\text{rotor}}}{dt} + \frac{d\dot{W}_{\text{friction}}}{dt} \quad \text{Eq. 7}$$

$$\frac{d\dot{W}_{\text{rotor}}}{dt} = \frac{\pi L \rho}{6} R_i^2 R_w (2R_i + R_w) \left(\frac{d\omega_i}{dt} \right)^2 \quad \text{Eq. 8}$$

$$\frac{d\dot{W}_{\text{fricción}}}{dt} = \pi L K R_k \omega_i^n \left((n+1) \frac{d\omega_i}{dt} + \omega_i \left(\beta \frac{dP}{dt} - \frac{\alpha}{T^2} \frac{dT}{dt} \right) \right) \quad \text{Eq. 9}$$

On the other hand, ML is related to viscosity of processed material through a proportionality constant that depends upon the rheometer and working parameters. Considering that all tests are carried out under the same experimental conditions:

$$ML = K_{\text{reom}} K = K_{\text{reom}} K_p e^{\alpha \left(\frac{1}{T_{\text{test}}} - \frac{1}{T_\alpha} \right) + \beta P_{\text{test}}} \quad \text{Eq. 10}$$

This model is applied to every cycle and fed by experimental process variables of temperature, pressure and rotor speed. Regarding material properties, density is selected among typical rubber density values, but five more parameters are required for the viscosity model: K_p , n , α , β , T_α . T_α

is set to room temperature, pressure viscosity dependence (β) is obtained from bibliography [5] and n from a generic rubber characterization, while K_p and α will result after applying a least squares optimization to minimize the difference among experimental and calculated intensity.

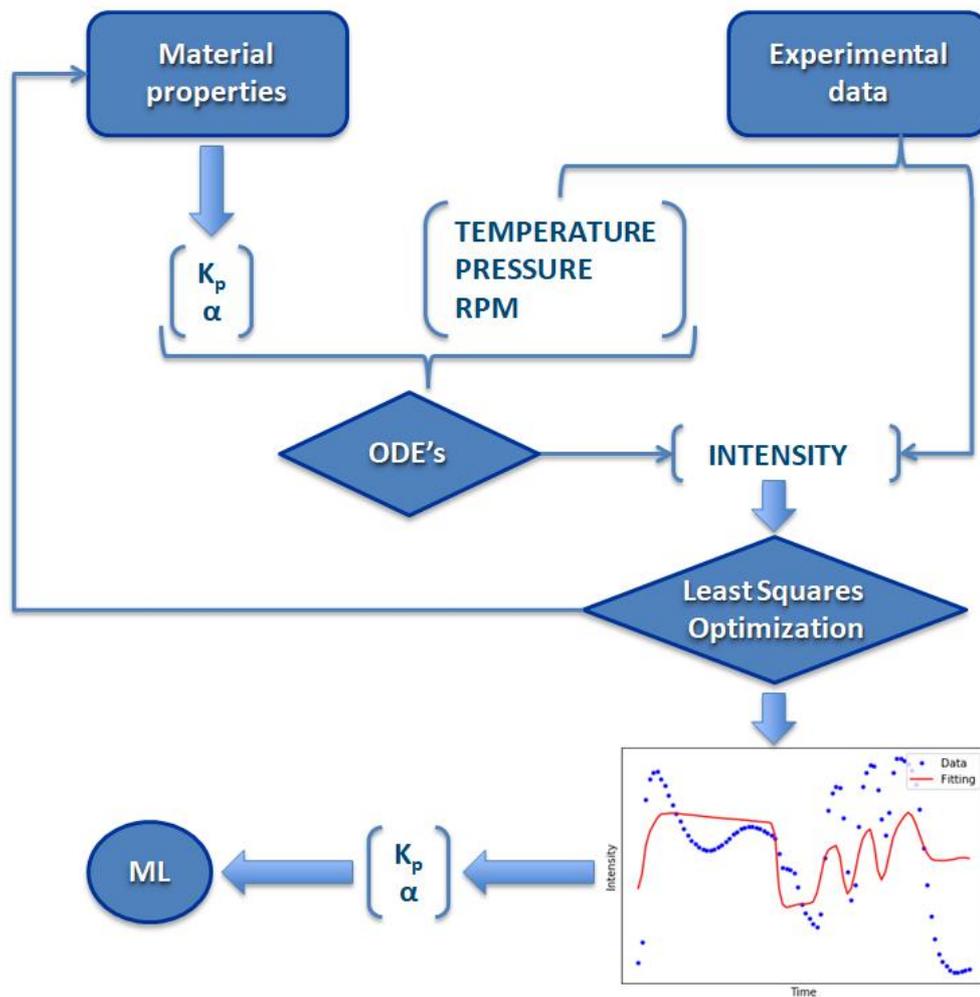


Figure 6 : ODE's model scheme

The complete process is shown in the above figure. Finally, calculated K_p and α through optimization can be related to ML as in Eq. 10.

2.3 CFD model

The flow inside the mixer is simulated using Ansys Polyflow to validate ODE's based model. The geometry consists of two tangential counter rotating rotors of the same size (R_i , R_w , L) as the ODE's mixer model, which results in a 40% increase in the fluid volume compared to mixer model's volume.

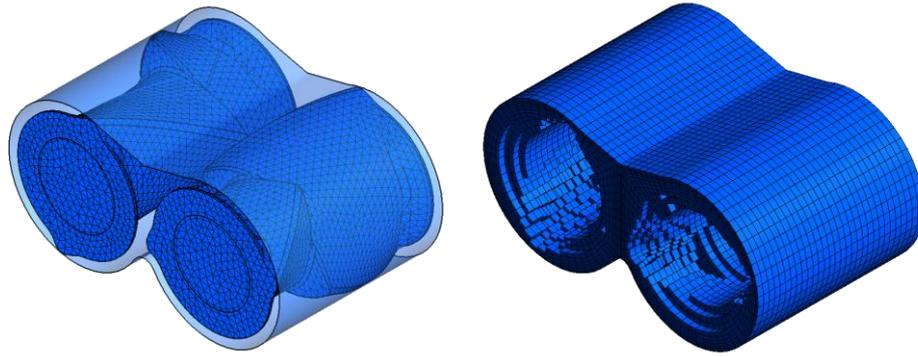


Figure 7: Mixer mesh

The fluid volume is initialized at constant temperature, calculated as mean initial temperature from experimental data, and adiabatic conditions are imposed.

Rotor speed is constant and equal to initial value of mean experimental rotor speed. It is assumed that the fluid sticks to the walls and to the rotor surfaces, and that the mixer is fully filled. At this stage actual material properties are unknown, so they are chosen from a generic rubber. Since a simulation of the full mixing cycle is computationally very expensive, only the first five revolutions are calculated.

The interpolation scheme is mini-elements for velocity with linear pressure, Picard iterations on viscosity and quadratic elements for temperature. The selected solver is AMF direct solver with secant iterative process and implicit Euler method for transient integration.

2.4 Data-driven physically-informed model

Complete experimental data allows to build a data-based model that predicts ML values through PCA and regression. On the other hand, by applying transport equations an ODE's model is developed and few specific process variables are required to get to ML prediction. By relating these two models, both of them can improve if information is wisely observed, beyond merely comparing its prediction capabilities.

While data-based model uses six process variables (intensity, pressure, rotor speed, mixing temperature, cylinder intensity and position) and nine compounds, ODE's proves that only the first four of them are relevant. At the same time, if theory's based prediction is worse than data-based model's, that would mean that experimental data still contains some information that ODE's model is omitting. Afterward, these missed variables can be found through data analysis and introduced in the ODE's model or input.

As mentioned above, CFD model's computational cost is very expensive, so only the first few seconds of the process are simulated. CFD provides very accurate but very small amount of information, therefore the output is employed to validate the ODE's model.

Later on this investigation and bearing in mind that mixer model dimensions are only an estimation, CFD model will be used to account for geometric uncertainties; this means, study the influence of fluid volume and mixer dimensions on the output.

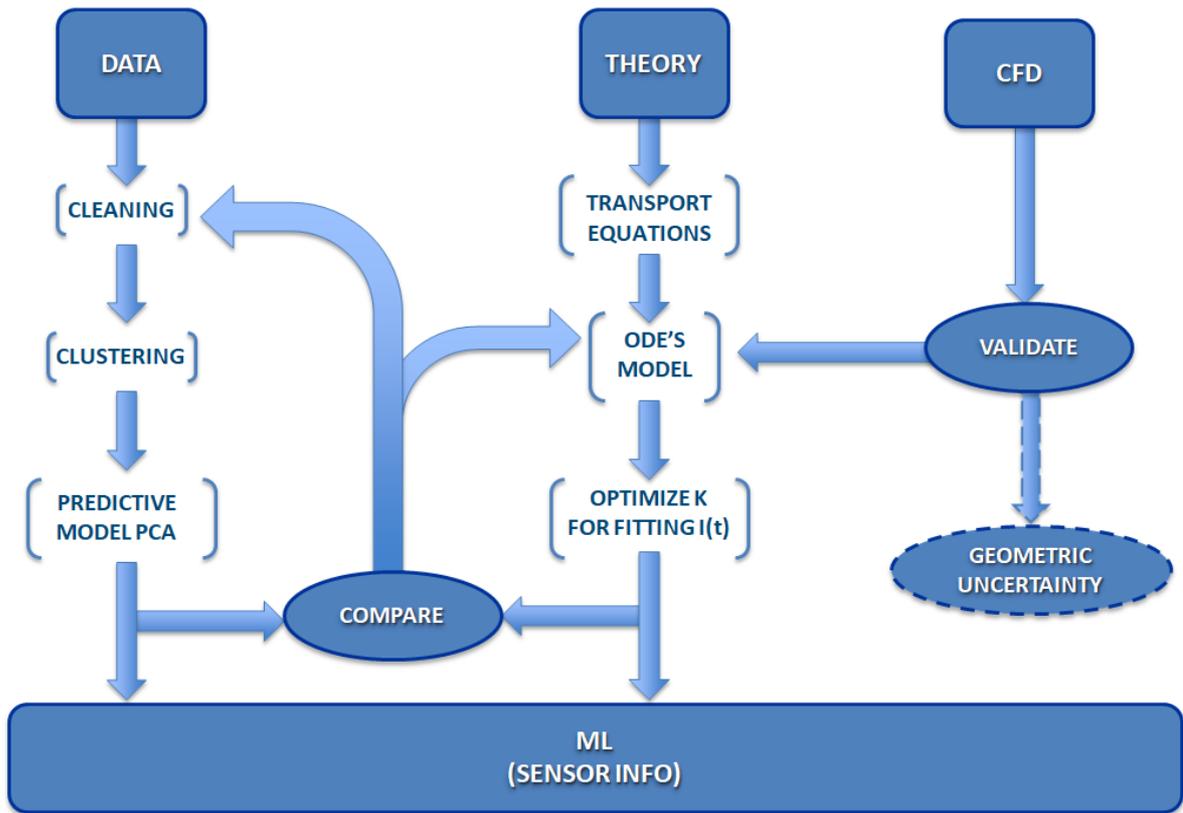


Figure 8: Data-driven physically informed model scheme

3 Results

3.1 Data-based model

First of all, low-pass filter is applied on data using a cut-off value based on measurement variance estimation, acting similarly to a Kalman filter. Explained variance ratio and coefficient of correlation depending on number of reduced components is studied for data before and after filtering.

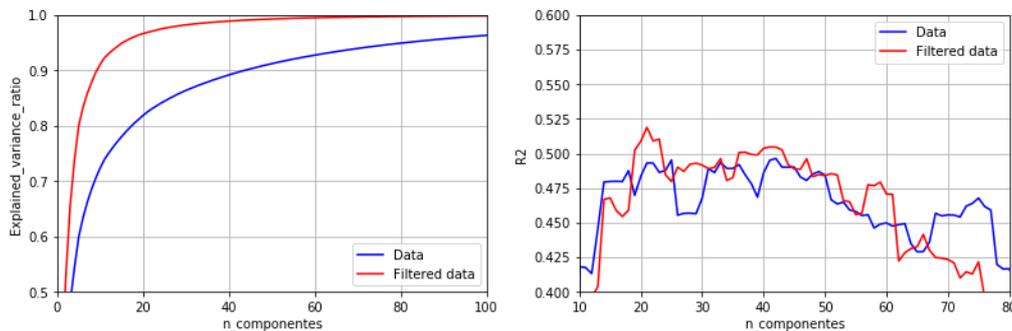


Figure 9: a) Comparison of sum of reduced components explained variance ratio vs number of reduced components; b) Comparison coefficient of correlation (R^2) of linear regression of reduced components model vs number of reduced components

The figure shows that the main improvement using filtered data is that the information gathered by the reduced components is larger, while filtering does not seem to affect significantly to coefficient of correlation; therefore filtering does not result in a better data model.

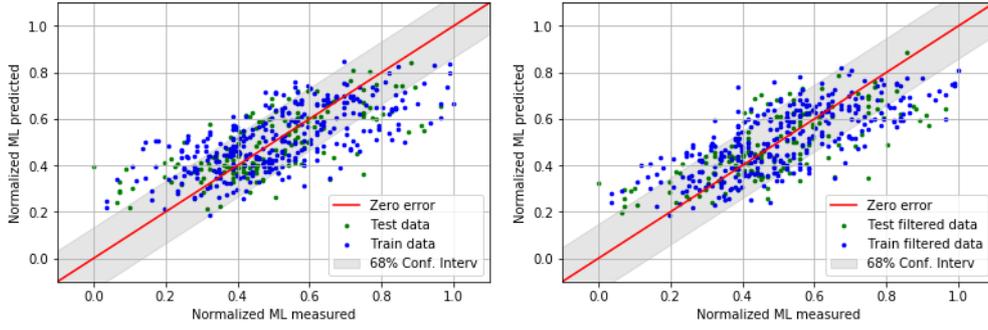


Figure 10: Linear regression to data and filtered data

Gaussian process regressor is applied to reduced data and reduced filtered data. Two thirds of the data are used to train the data model and one third to test it. The coefficient of correlation increases compared to previous linear regression; so does the computational cost.

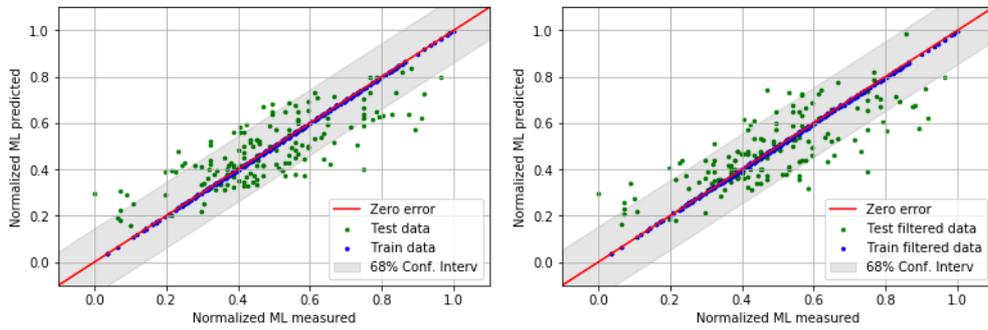


Figure 11: Gaussian regression to data and filtered data

The results of applying PCA and linear regression (LR) or gaussian regression (GR) to data and filtered data are summarized in the next table:

	Data	Filtered data
Reduced components	21	21
Explained covariance	82.5%	96.9%
Predicted data st. deviation (LR)	0.134	0.144
R2 (LR)	0.493	0.519
MSE (LR)	0.019	0.018
Predicted data st. deviation (GR)	0.142	0.150
R2 (GR)	0.611	0.594
MSE (GR)	0.014	0.015

Table 1: Comparison of PCA and linear regression (LR) or gaussian regression (GR) results among data and filtered data

Gaussian regressor increases the coefficient of correlation in 24% and 14% for data and filtered data, respectively, while decreasing the mean squared error (MSE) in around 20%. In this case, there is not significant differences between applying a filter previous to PCA, besides the already mentioned increase in explained covariance.

3.2 Theory-based model

In order to test the ODE's model, the input process variables in a first approach are the mean experimental values: mean temperature, mean pressure, mean rotor speed; in the same way, least squares optimization proceeds comparing calculated intensity to mean experimental intensity.

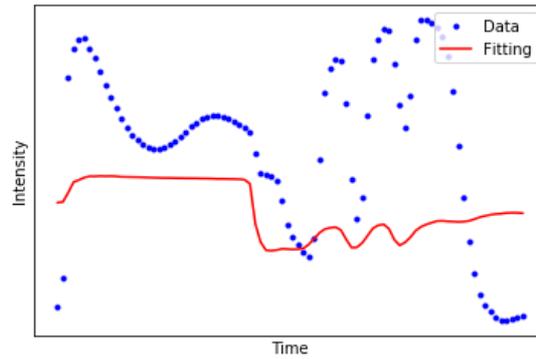


Figure 12: Fitting using mean experimental values

Experimental data is filtered previously to be introduced into the least squares optimization, to smooth the minimization, since the provided intensity data has zero precision. Once the optimization is runned for all cycles, and a pair of values $K_p\text{-}\alpha$ is obtained for each one of them, K is calculated at test temperature and pressure (K_{test}) as in Eq.6.

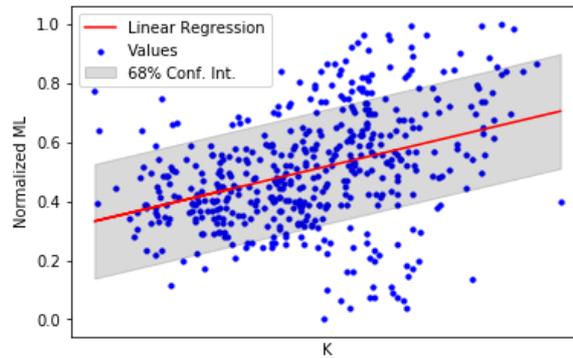


Figure 13: K_{test} vs ML – Linear regression

In the above figure, paired values K_{test} -normalized ML for each cycle are represented. According to Eq. 10, direct proportionality was expected. However, results fit a linear regression with significant Y-intercept. It is reasonable to consider that the rheometer has a working range and that the offset is related to the lower working limit of the instrument.

A gaussian regressor is applied, as in the case of data-based model, but only one dimension is required, the ODE's model output (K_{test}).

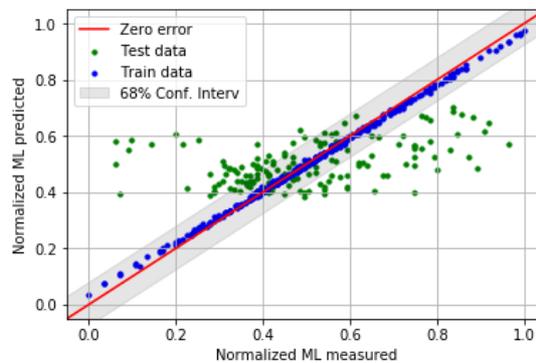


Figure 14: ML prediction from K_{test}

Regression	Linear	Gaussian
Predicted data st. deviation	0.194	0.078
R^2	0.141	0.161
MSE	0.032	0.029

Table 2: Linear and gaussian regression results of ML prediction from K_{test}

While the standard deviation of predicted data is similar or lower than the data-based model, the coefficient of correlation and squared error worsen noticeably.

3.3 CFD model

As stated before, CFD is the most computationally expensive model. It was not feasible to run a full-time mixing cycle simulation, but only 5% of the process. Nevertheless, it is still worthy to evaluate the results, since ODE's and CFD model share some hypothesis, like mixer dimensions.

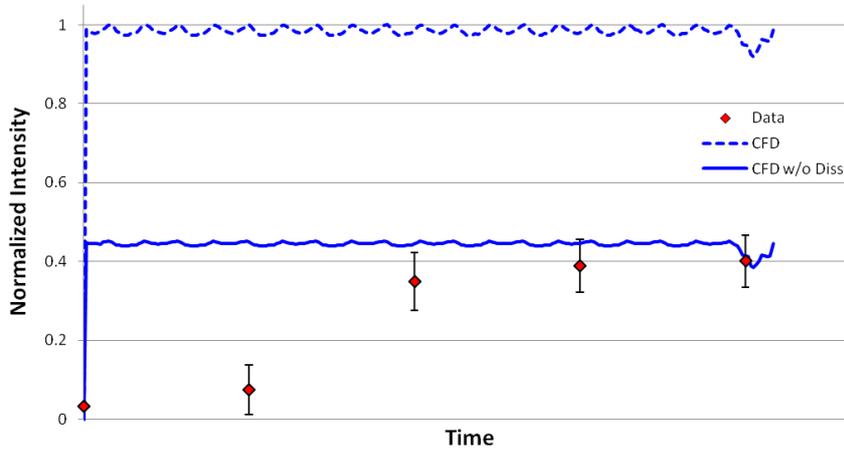


Figure 15: Comparison of intensity

Raw intensity from CFD-simulation is over twice times experimental intensity. But when removing viscous dissipation contribution to power requirement, experimental intensity values reach, after certain transient period, computed intensity. Viscous dissipation rate depends highly on material properties which, for CFD simulation, are chosen from a generic rubber. In addition, this effect was not taken into account in ODE's model due to its complexity.

3.4 Data-driven physically-informed model

Although at this point ODE's model has not improved ML prediction in comparison to data-based model, theory-based model provides useful insight into mixing process: intensity is highly correlated to ML. In a first approach, intensity is integrated for every cycle and represented against its respective ML value.

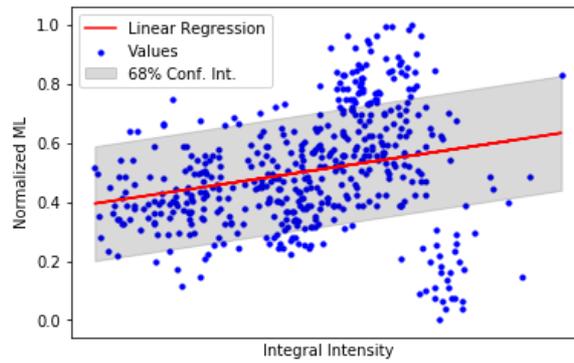


Figure 16: Integral intensity vs Normalized ML

The figure above highly resembles to Figure 13. Once again, a gaussian regressor is applied, using only a single dimension as input (integral intensity), to predict ML.

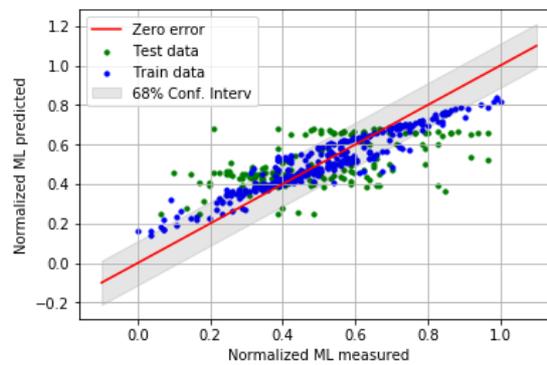


Figure 17: ML prediction from Integral Intensity

Still the results are not as good as the data-based model, so it is assumed that ODE's model is missing some relevant information to ML.

The PCA created in the data-based model allows to analyse the importance of each variable: by representing the first two dimensions and detecting the heaviest weights, which belong to the most influential variables.

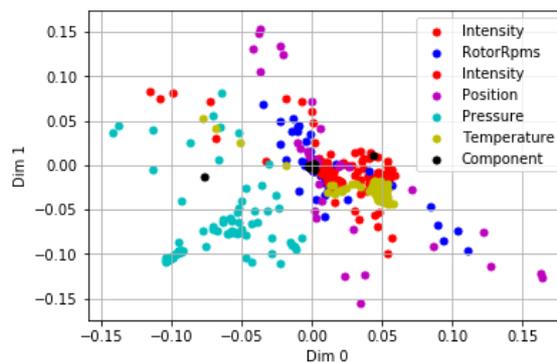


Figure 18: Variables' weights of first two dimensions of PCA

Components are not included in ODE's model. By depicting component's weights, it is observed that the most important component in PCA is component D.

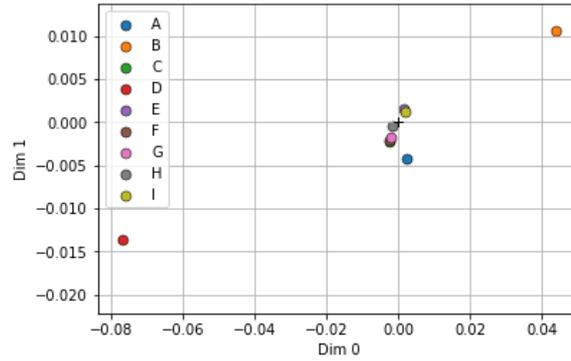


Figure 19: Components' weights of first two dimensions of PCA

Including component D in the previous regressor results in a simple model with two dimensions (Integral Intensity, component D) as input, but high correlation, as shown in figure below:

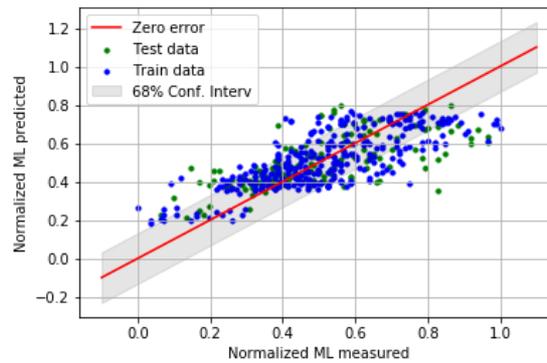


Figure 20: ML prediction from Integral Intensity and Component D

Dimensions	1	2
Predicted data st. deviation	0.111	0.131
R^2	0.283	0.470
MSE	0.027	0.020

Table 3: Gaussian regression results of ML prediction from Integral Intensity, and Integral Intensity plus Component D

Using only two dimensions (integral intensity and compound D), results are almost as good as the data based-model (21 dimensions) in terms of coefficient of correlation and error, with significant reduction in computational cost.

3.5 Comparison among models

The table below summarizes the best results of each model: reduced data without filter for data-based model, theory-based model and Data-Drive Physically-Informed (DDPI) with two dimensions; using gaussian regressor in all of them.

Model	Data-based	Theory	DDPI
Dimensions	21	1	2
Predicted data st. deviation	0.142	0.078	0.131
R^2	0.611	0.161	0.470
MSE	0.014	0.029	0.020

Table 4: Comparison among data-based, theory and DDPI model

Although data-based model shows the highest correlation and lowest error, it should be outlined that data-based requires 21 dimensions. On the other hand DDPI model reduces the dimensions in 90%, but loses 21% of correlation and increases the error in 40% compared to data-based model.

4 Conclusions

A data-driven physically-informed method is presented where theory and numerical simulations help to clean and select the most relevant information to build a data-driven model. This is a novel approach where instead of focusing on data-assimilation into numerical simulation, we focus on improving the classical PCA approach to decide most relevant information based on physically-informed restrictions. The main conclusion is that, for operation-units analysis of material manufacturing process, the information provided by the integral analysis of transport process is relevant for limiting the information used in data driven models. We relate this limitation to the thermodynamic consistency that a predictive model (without regards of its origin) should have.

Acknowledgments

This work was partially funded by the Spanish Industrial Research Programme RETOS-COLABORACIÓN RTC-2016-5438-4.

References

- [1] E. L. Paul, V. A. Atiemo-Obeng, and S. M. Kresta, *Handbook of Industrial Mixing: Science and Practice*. Wiley, 2004.
- [2] J. N. Kutz, *Data-Driven Modeling & Scientific Computation: Methods for Complex Systems & Big Data*. OUP Oxford, 2013.
- [3] Chao Shang, *Dynamic Modeling of Complex Industrial Processes: Data-driven Methods and Application Research*. Springer, 2018.
- [4] Nobuyuki Nakajima, *Science and Practice of Rubber Mixing*. RAPRA Technology, 2000.
- [5] K. D. Housiadas, ‘An exact analytical solution for viscoelastic fluids with pressure-dependent viscosity’, *J. Non-Newton. Fluid Mech.*, vol. 223, pp. 147–156, Sep. 2015.