ALMA MATER STUDIORUM · UNIVERSITÀ DI BOLOGNA

SCUOLA DI INGEGNERIA E ARCHITETTURA

Dipartimento di Ingegneria Civile, Chimica, Ambientale e dei Materiali

CORSO DI LAUREA MAGISTRALE IN INGEGNERIA CHIMICA E DI PROCESSO

TESI DI LAUREA

in

Process Analysis for Energy and Environment

Application of multivariate statistical methods

to the modelling of a flue gas treatment stage

in a waste-to-energy plant

CANDIDATO
Muratori Giacomo

RELATORE
Ing. Antonioni Giacomo

CORRELATORE
Ing. Dal Pozzo Alessandro

Anno 2017/2018
II sessione

ii

Abstract

Among all the flue gas components produced in Waste-to-Energy plants, acid airborne pollutants such as sulfur dioxide, hydrochloric acid and other halogens hydrides (HF, HBr) have, depending on their toxicity and amount produced, the most rigorous emission standards provided by the European Parliament. Their removal is thus a key step of the flue gas treatment which is mainly achieved with the Dry Treatment Systems (DTS), technologies based on the direct injection of dry solid sorbents which is capable to subtract the acid from the gas stream with several important advantages and high removal efficiencies. However, the substantial lack of a deeper industrial knowledge makes difficult to determine accurately an optimal operating zone which should be required for the design and operation of these systems. The aim of this study has been therefore the exploration, while basing on an essential engineering expertise, of some of the possible solutions which the application of multivariate statistical methods on process data obtained from real plants can give in order to identify all those phenomena which rule dry treatment systems. In particular, a key task of this work has been the seeking for a general procedure which can be possibly applied for the characterization of any type of DTS system, regardless of the specific duty range or design configuration. This required to overcome the simple mechanical application of the available techniques and made necessary to tailor and even redefine some of the available standard procedures in order to guarantee specific and objective results for the studied case. Specifically, in this so called chemometric analysis, after a pre-treatment and quality assessment, the process data obtained from a real working plant (Silea S.p.A., Valmadrera, Como, Italy) was analyzed with basic and advanced techniques in order to characterize the relations among all the available variables. Then, starting from the results of the data analysis, a linear model has been produced in order to be employed to predict with a certain grade of accuracy the operating conditions of the system.

Sommario

Tra gli svariati composti volatili prodotti dai termovalorizzatori, gli inquinanti acidi come SO2, HCl e altri acidi alogenidrici(HF, HBr) hanno, data la loro tossicità e grandi quantità prodotte, i limiti di emissione più stringenti tra quelli previsti dal parlamento Europeo. La loro rimozione è quindi un passaggio chiave del trattamento dei fumi che si basa in gran parte sull'utilizzo di sistemi di trattamento a secco (DTS), cioè tecnologie basate sull'iniezione diretta di sorbente solido capace di sottrarre l'acido dalla corrente di gas con diversi vantaggi. Tuttavia, considerato anche il gran numero di possibili configurazioni, sui DTS non è ancora stata raggiunta quella precisa ed approfondita conoscenza industriale che sarebbe necessaria per un'ottimale progettazione e funzionamento degli stessi. Lo scopo di questo studio è stato quindi, partendo da una fondamentale base ingegneristica, l'esplorazione delle possibilità che l'applicazione di metodi statistici multivariati a dati ottenuti da processi reali possono dare per l'identificazione dei fenomeni che controllano il funzionamento di questi impianti. In particolar modo il lavoro si è concentrato sulla definizione di una procedura generale che possa essere applicata a prescindere dalla configurazione o dalla potenzialità dell'impianto e questo ha richiesto il superamento della semplice applicazione meccanica delle tecniche disponibili, rendendo anche necessario l'adattamento o la ridefinizione delle procedure in modo tale da garantire risultati specifici e oggettivi per il caso studio. Nello specifico, in questa cosidetta analisi chemometrica, dopo alcune valutazioni di tipo qualitativo, i dati di processo ottenuti da un impianto funzionante (Silea S.p.A., Valmadrera, Como, Italy) sono stati analizzati per mezzo di svariate tecniche in modo da poter caratterizzare le relazioni tra le singole variabili. Successivamente, partendo dai dati di questa analisi, un modello lineare è stato creato affinché possa essere impiegato per predire le condizioni operative del sistema.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Waste to energy plants

Volume and hazard reduction of the waste produced by human activities is fundamental to lower the environmental, social and economic impact that an untreated waste could induce. Concerning this, incineration is part of a wide range of complex treatment processes that allow the overall management of the of wastes that arise in society.
Incineration is a Waste to Energy (WtE) process that burn and decompose substances in waste in presence of oxygen while recovering energy and reducing thus the use of fossil fuels and the CO_2 production. This allows to pursue the volume reduction of waste together with the capture (and concentration) or destroying of potentially harmful substances such as the ones already present in the waste or the ones derived by the combustion process. Indeed, the waste incineration process transfers about the 80% of the burnt matter into the flue gas that is mainly composed of CO_2 and water, together with a mix of many other components, showed in the next table (Tab.1.1) as typical concentration ranges in waste flue gas.

CO	TOC	Dust	**HCl**	**HF**	**SO2**	NO_x	Hg	Cd	PCDD/F
<10-30	1-10	1000-5000	**500-2000**	**1-10**	**150-400**	200-500	0.1-0.5	0.1-0.5	1-10

Table 1.1: Daily average emissions concentration ranges (mg/Nm^3).

The atmospheric emission of these components is harmful for human health and ecosystem integrity and thus emission limits are provided from the qualified authorities. Each single pollutants in the flue gas have then to be reduced from the concentration of the combustion chamber to the required emission standards. This is done through a series of abatement stages that are called air pollution control (APC) systems and come right after the combustion chamber and before the emission in atmosphere. Commonly an APC train (Fig.1.1) starts with the removal of the fly ashes directly downstream the boiler, then the next steps are the neutralization of acid gases, the NO_x removal and a final polishing step aimed to remove all the left pollutants like mercury and low volatile organic compounds such as dioxins.

The European emission standards are given by the Industrial Emissions Directive (IED) of the European parliament[4, 2010/75/EU] together with the criteria for the determination of the best available technologies (BAT) which must be used for the APC systems to reach these standards.
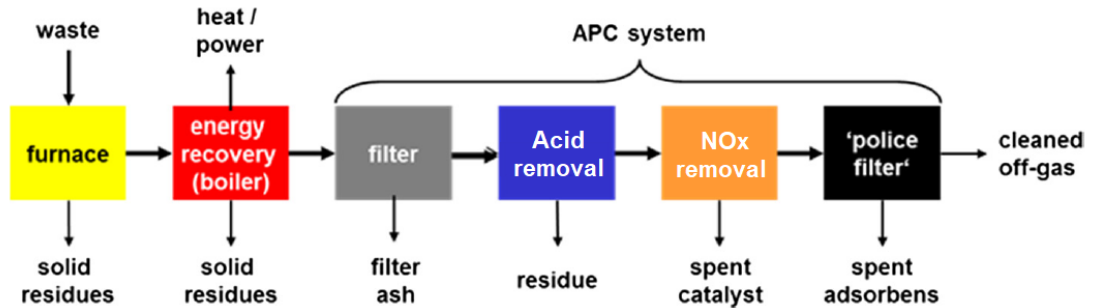
Figure 1.1: Scheme of a conventional waste incineration plant [1]

1.1.1   Acid gases

Among all the flue gas components, depending on their toxicity and amount produced, acid airborne pollutants such as sulfur dioxide, hydrochloric acid and other halogens hydrides (HF, HBr) are the ones listed in the IED which have the most rigorous emission standards. HCl especially, is obtained from the decomposition of inorganic chlorides and organic chlorine compounds (PVC etc.) that are highly present in wastes and, without a correct abatement strategy, incineration would be its major emission source to the atmosphere. Air emission limit values concerning waste incineration plants are reported from the IED in table 1.2 as daily average emission limit and half-hourly average emission limit for the 97% of the overall operation time (B) and with a possible exceeding in emission in the 3% of the measures given in (A).

Emission limit (mg/Nm^3)	Daily	A (100%)	B (97%)
Hydrogen chloride (HCl)	10	60	10
Hydrogen fluoride (HF)	1	4	2
Sulphur dioxide (SO2)	50	200	50

Table 1.2: Daily and half-hourly average emission limit (STP, standardized at 11 % oxygen in waste gas)

1.2   Acid gas removal: Dry treatment systems

Acid gas removal can be pursued by several processes which are classified as wet, if the acid react with a neutralizing agent in liquid phase, or dry/semi-dry, if the neutralizing agent is a solid or a wet solid. Until 2000s the process standard for this type of application was the wet scrubbing but, since mid-1990s, dry treatments have been demonstrated to be cost-effective, easy to operate and to be maintained[1] and then have gradually substituted the more obsolete technologies.

In DTSs acid pollutants are subtracted from the gas stream through the reaction with some alkaline powders which can be divided in two main categories, the calcium based and the bicarbonate based system, which have different costs, reactivity and working conditions.

Slaked lime ($Ca(OH_2)$) is obtained through the calcination and hydration of the limestone ($CaCO_3$) and due to its high availability and low cost it is the most used reagent in these type of applications. Lime has lower reactivities than other types of reagents but some high porosity refined types are produced at higher costs through specific processes which guarantee higher reactivities. For Ca based systems the optimal working temperature window is between 120°C and 160°C, with a performance increase proportional to the decreasing of temperature, due to the effect that the higher relative humidity induces on the particle core, enhancing the reactivity.

Bicarbonate based alternatives are largely widespread and rely on the Solvay registered NEUTREC™ process, which employ Solvay synthesized sodium bicarbonate ($NaHCO_3$) as reagent. The latter, despite its higher reactivity, is more costly than the Ca based reagents but its final process residues can be regenerated and re-used. These systems, differently from the Ca alternatives, follow a classical kinetic increase from 140°C to 300°C, due to classical kinetic factors. Find in table 1.3 a comparison recap of different technologies.

	T [°C]	SR (HCl)	SR (SO2)	M [kg/waste ton]
Wet systems				
	60-65	1.02-1.15		12-18
Dry systems				
Ca based	120-160	1.1-1.5	1.3-3.5	20-30
$NaHCO_3$	140-300	1.04-1.2		15-18

Table 1.3: Typical operational ranges for Temperature, stoichiometric ratios SR and reagent mass flow M [1].

1.2.1   The chemical process

Slaked lime is capable to neutralize acid gases with a single step reaction:

$$Ca(OH)_{2(s)} + 2\,HCl_{(g)} \rightarrow CaCl_{2(s)} + 2H_2O_{(g)}$$
$$Ca(OH)_{2(s)} + 2\,HF_{(g)} \rightarrow CaF_{2(s)} + 2H_2O_{(g)}$$
$$Ca(OH)_{2(s)} + SO_{2(g)} + \tfrac{1}{2}O_{2(g)} \rightarrow CaSO_{4(s)} + H_2O_{(g)}$$
$$Ca(OH)_{2(s)} + CO_{2(g)} \rightarrow CaCO_{3(s)} + H_2O_{(g)}$$
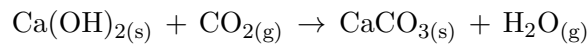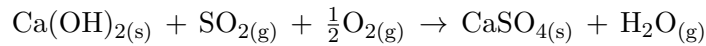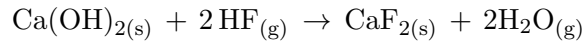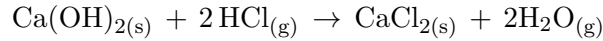
The NEUTREC$^{\text{TM}}$ process follows instead two steps where firstly, at temperatures above 140°C, NaHCO$_3$ releases CO$_2$ and water in a so called "pop corn effect" decomposition that gives a sodium carbonate (Na$_2$CO$_3$) powder which, given its high specific surface, is much more reactive than the starting bicarbonate and can then neutralize the acid gas.

$$NaHCO_{3(s)} \rightarrow Na_2CO_{3(s)} + CO_{2(g)} + 2H_2O_{(g)}$$

$$Na_2CO_{3(s)} + 2\,HCl_{(g)} \rightarrow 2NaCl_{(s)} + H_2O_{(g)}$$
$$Na_2CO_{3(s)} + 2\,HF_{(g)} \rightarrow 2NaF_{(s)} + H_2O_{(g)}$$
$$Na_2CO_{3(s)} + SO_{2(g)} + \tfrac{1}{2}O_{2(g)} \rightarrow NaSO_{4(s)} + CO_{2(g)}$$
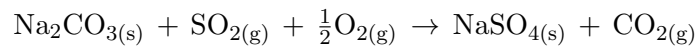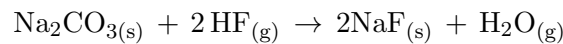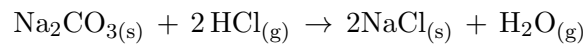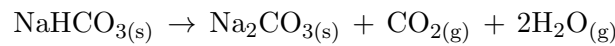
Solid-gas reaction model

For all types of reagents the heterogeneous gas-solid reaction can be easily carried-on in one single step with fast operations but, differently by wet scrubbing, kinetics are much more un-promoted. The process follows in fact a so called "shrinking core mechanism" (Fig.1.2) where the acid gradually consume the sorbent particle and creates a product layer on it.
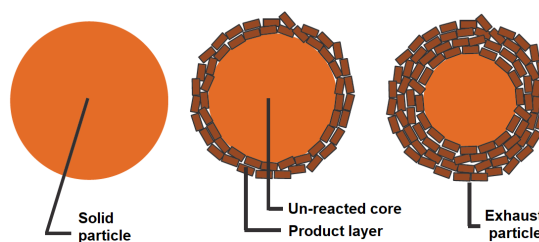
Figure 1.2: Gas-solid shrinking core mechanism [2]

The thickening of the gathering product layer slows down the fresh core reaction and does not allow the full depletion of the sorbent, making thus essential to account for a quantity of reagent which is bigger than the stoichiometric need, leading then to a larger reagent consumption and a higher production of final residues.

These complex phenomena have been studied[5, Antonioni, 2011] for the reactions of hydrogen chloride and sulfur dioxide with sodium bicarbonate, leading to the production of a conversion model that is capable to predict the removal efficiency of acid gases as a function of the rate of solid reactant:

$$\chi = \frac{rs^n - rs}{rs^n - 1} \tag{1.1}$$

As can be seen from the equation, the acid pollutant conversion χ is expressed as a function of two parameters:

- **rs**: is a dimensionless parameter correspondent to the ratio between the effectively dosed reagent and the one which is stoichiometrically needed. Starting from the reactions reported in the previous section it is possible to write it as:

$$rs = \frac{NaHCO_{3in}}{NaHCO_{3st}} = \frac{NaHCO_{3in}}{HCl_{in} + 2SO_{2in} + HF_{in}} \tag{1.2}$$

- **n**: is an adjustable parameter which is proportional to the removal efficiency of the system. This is a lumped parameter which gathers all the chemical phenomena occurring in the system and the operating conditions affecting the reaction (e.g. temperature, contact time, particle size etc.).
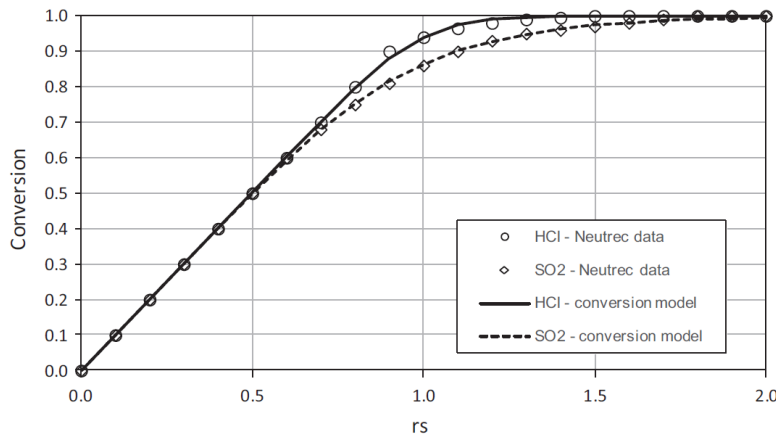
Figure 1.3: Conversion model validation vs. literature bench scale data [5, Antonioni, 2011]

The previous figure (Fig.1.3) refers to a model validation carried out on bench scale data[7, Brivio, 2007] and it shows a very good agreement of the model equation with experimental data, confirmed by a coefficient of determination $R^2 > 99\%$ for both the HCl and SO2 predictions.

1.2.2   DTS configurations

Possible DTS configurations are several and highly customized depending on the amount and types of waste burnt. Acid gas removal is usually the first pollution control step after the furnace and for this reason the injection of the powders occurs right after the air pre-heating system and, in some cases, after a humidification pre-conditioning. Then, once the stream has been sprayed, it is sent to the de-dusting step, commonly performed through a cyclone or a fabric filter, which can work also for the removal of the particulate and flying ashes coming from the furnace. In case that fabric filters are used alkaline reagents settle on the membrane while creating a fixed bed where the gas pass through and the reactions keep going on, leading thus to an overall increasing of the abatement performance. Starting from this basic configuration it is possible to customize the system for example by changing the injection point or with the use of a humidification pre-conditioning system. Actual design trends push then toward the use of multiple stages systems that use both bicarbonate and slaked lime and which, if properly operated, have been demonstrated to require a lower amount of solid reactants than a single-stage process[5].

1.3   Aim of the study

Generally speaking DTSs can be considered simple and versatile plants which are cost-effective, easy to operate and to be mantained. This, together with the easiness of solid residues disposal and the absence of operating issues (such as the "rain-out" at the chimney), make DTSs the preferable solutions in most of the cases.
However, particularly in case of multiple stages solutions, these systems suffer of a substantial lack of a deeper industrial knowledge[5] which makes difficult to determine accurately safe operating zones, leading then to the necessity to operate with high excess of reactants that leads naturally to a surplus production of residues and consecutive disposal costs. Actually, the reported model (eq.1.1) demonstrates its applicability for the prediction of the DTS conversion on large time averaged windows[5], but its single lumped parameter makes it not very flexible to briefly changing operating conditions and thus not capable to produce reliably instantaneous predictions. For this reason, even in light of the increasingly lower emission standards, the aim of this study is the exploration of some of the possible solutions which a chemometric analysis on real process data can give. Like all the chemical treatment processes in fact, air pollution control systems in WtE plants are equipped with process control instrumentation where data are measured at high frequencies in order to be employed in monitoring and control. These data however represent a great deal of information that by means of multivariate statistical methods can be extracted and employed to identify all the DTS ruling phenomena and try to produce a model which is capable to produce reliable instantaneous predictions.

Chapter 2

Material and method

Chemometrics is the application of mathematical and statistical tools on a chemical system in order to characterize the mutual relations among the available measures and study the state of the system. Usually this approach is applied to on-purpose designed experiments that are performed on controlled conditions and where high precision data are collected with high reliability. In this specific case however, process data are not produced primarily for the chemometric analysis, so they must be checked and cleaned since they must follow specific requirements of type, ordering and quality. After this, the big quantity of correlated and redundant informations characteristic of the process derived measures makes necessary, in addition to the classical visual and basic statistical analysis, the use of advanced statistical techniques that allow a deeper elaboration and are capable to extract the essential informations from the dataset, allowing then to proceed with a final regression analysis aimed to the estimation of a model for the process. This model can then be employed to predict with a certain grade of accuracy the system state allowing then to optimize the operating conditions. In this chapter, techniques which have been demonstrated to be useful in this work are introduced without any claim to be exhaustive in the presentation of the broad range of methods which are effectively available.

2.1   Data pretreatment

All the measurements from DCS are available in large quantities, are subjected to many types of error and, in case that these inconsistencies are not detected and correctly removed, this will end easily in a misunderstanding of the data trends which can even make impossible the analysis and modeling. In particular it is possible to spot four types of error which are characteristic of sensor measurements:

- Outliers: by definition these are the measures which fall out from the normal measures trend and for this reason they can be easily spotted and removed; They take part to this type of error any zeros and unjustified peaks measures.

- Missing measures: measures holes in the process data consist usually in $10 \div 20\%$ of the overall availability.

- Gross errors: are commonly related to a wrong calibration of the sensor and they are easily spotted and conceptually avoidable (e.g. measured inlet and outlet flow rate in an equipment don't close the mass balance).

- Random error: this is the so called noise, it is unavoidable and hopefully small so that a proper filter is capable to remove it without affecting the real measure.

2.1.1 Data cleaning

Before to start the analysis and modeling of the data, previous mentioned wrong data must be identified and corrected such that the final results are consistent and conciliated each other. Usually, wrong data can be manually spotted through a visual control on the data charts (Fig.2.1) but in case of large quantities of data or for on-line applications a computational approach is necessary. Hence, it is essential to create some algorithms that allow to detect, remove and substitute wrong measures with values that are much more likely to be consistent with the surrounding measures and process constraints and thus help the next analysis to be more correct. Big windows of missing measures cannot be reconstructed reliably and for this reason they force the subdivision of the dataset in more subsets which anyway could turn out to be useful in the regression analysis for the creation of training and validation sets.
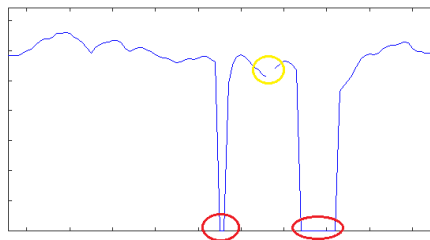
Figure 2.1: Error examples: Outliers (red) and missing measure (yellow)

2.1.2 Data filtering

Once data have been re-constructed it is then necessary to smooth out short-term fluctuations and thus remove the noise of the sensor through a filter that can have a more or less sophisticated approach.
The most simple filtering method is the moving average, where each measure is substituted by the average of the previous or next (or both) measured n-terms. The larger is the width of window average, the stronger is the filter and thus the higher is the capacity to remove random errors but with an higher loss in information. Filtering works also as a polishing step to assimilate better into the data trend those measures that have been added during the cleaning phase to substitute some wrong measures.

Find in the next figure (Fig.2.2) the effect of a moving average filter on the cleaned version of the previous example (Fig.2.1).
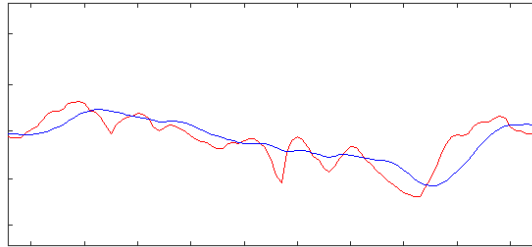
Figure 2.2: Final process data after cleaning (red) and filtering (blue)

2.2 Data analysis

High complexity of the process data can make very difficult to understand the background of the observed phenomena. This, in addition to the availability of advanced modeling tools, could induce the analyst to pass directly to the modeling phase without trying to give a meaning to the starting observation, ending thus in models that, despite even outstanding results, cannot be explained. Data analysis includes all those techniques that, through a more or less deeper elaboration, allow to explore also the most complex systems so that it is possible to:

- Suggests hypotheses about variables correlations which, on an expertise base, can be demonstrated as causalities;

- Help to make assumptions;

- Support the selection of more advanced statistical techniques;

- Provide a basis for further data collection;

In this specific case the main purpose of the analysis is to characterize the type and the influence of each variable on the system state and so can be considered successful if all the necessary statistical tools are correctly employed to characterize the system, understand what are the independent variables that truly affect its state and finally determine their type of influence on it.

2.2.1 Quality of the data

Chemometric analysis on real working systems finds an important advantage in the easy and the high availability of the data. The latter however are not produced intentionally to be analyzed and, depending on their characteristics, it can be more or less easy to obtain meaningful results. Indeed, the most common techniques employed in the regression analysis rely on the fact that each variable should vary enough so that it is possible

to study its variation related to the one of the other variables. This characteristic is not obvious for process variables which instead in some cases can assume almost constant values or are kept intentionally on specific set points by the control system. Moreover, it is common for many statistical methods to rely on the normal distribution of the analyzed data and, in case this is not verified as it is common for process variables, the final result ends to be biased. Nevertheless, these are intrinsic and unavoidable circumstances and in these cases the only possible solution is the preliminary characterization of all the variables probability distribution so that, at the end of the analysis, it is possible at least to discern real poor result from the ones dependent on the bad conditioning of the dataset.

The probability distribution

The distribution of the variables measures can be evaluated either visually with histograms or through specific parameters called simple moments of the distribution. The first two moments, the mean μ and the standard deviation σ, are the basic parameters which define the Gaussian normal distribution that fits at the best way the variable set. While the first gives the absolute order of magnitude of the measure, the second quantify the dispersion of the variable set.

In case of a multi-variable dataset it is better to refer to the relative standard deviation $\sigma^* = \frac{\sigma}{\mu}$ (called also coefficient of variation) which is a quantification of the scale of the variable and allows to compare different variables. Many mathematical and statistical tools are sensitive to the relative scaling of the variables and require to bring the whole dataset to a common scale by means of a so called normalization process. The most used normalization method is the standard scores evaluation where each variables sets is transformed through eq. 2.1 in order to have zero mean and standard deviation equal to one.

$$X^* = \frac{X - \mu}{\sigma} \tag{2.1}$$

The skewness γ (3rd moment) and the kurtosis k (4rd moment) allow to characterize the deviation of the variable distribution from the normality.
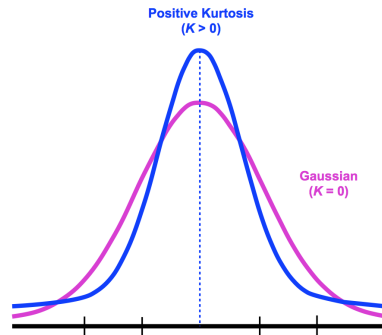
Figure 2.3: Kurtosis comparison

The kurtosis k in particular is an important parameter which allows to estimate, for a given variability expressed by σ^*, if the probability distribution is more or less squeezed toward the mean. If a distribution is characterized by an high kurtosis it means that, related to a normal distribution which have the same σ, a bigger part of the measures are concentrated nearer to the center and a small part of those are, as a counterbalance, as much far from the mean as low it is their number. In process derived variables, assuming that true outliers have been fully removed, high kurtosis are usually characteristic of measures which are mostly constant but experience few justified measure peaks. This makes the measure disadvantaged in a correlation analysis on the base of its intrinsic lack of variability and even more depending on the next normalization. Indeed, the fact that the normalization procedure relies only on the first two moments neglecting the possible absence of normality, makes high kurtosis a detrimental factor in the analysis and, being formally equal in σ^*, the variables which are characterized by an higher kurtosis, and are thus more squeezed toward the center of the distribution, are subjected to a variance reduction which is higher than the one which should be really needed due to their true distribution.

2.2.2 Preliminary evaluations

The regression analysis starts with the application of basic techniques aimed to the preliminary evaluation of the possible correlations/causalities and to start making hypothesis and assumptions. This, in case of complex systems, is not intended to fully determine the system but is essential so that, once the data are analyzed with advanced techniques, it is possible to understand better the results. In particular, this phase is necessary to identify those correlations that, although they lies naturally in the dataset, are supposed to be irrelevant for the study and need to be clearly recognized in the next steps of the analysis in order to avoid misunderstandings.

Data visualization

The first preliminary step for the analysis of a complex system is the representation of the data. Effective visualization in fact makes complex data more accessible, yet understandable and usable. A human can distinguish differences in size, shape, position, orientation and color readily without significant processing effort and thus a correct visualization set up allow to perform particular analytical tasks. It is easy, once data are represented through an effective way, to make comparisons, spot correlations, support the selection of more advanced statistical tools and confirm or refute any other results.

In this perspective, process data time-series (Fig.2.4), captured over a period of time, are the first step to understand how operations affect that measure, vice-versa and also to spot possible correlations in case that two or more variables are superposed in the same chart.
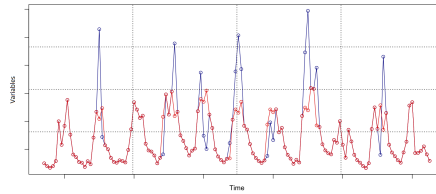
Figure 2.4: Time-series of two correlated variables

Then, if some correlations are spotted, they can be further studied with correlation plots, that are bivariate representations which, in case of multivariate evaluations, are called scatterplot matrices (Fig.2.5). Many other representations are possible through histograms, bar charts and other advanced representation tools in order to rank and categorize the effect of any real or on-purpose created variables or to quantify the frequency distribution of the latter.
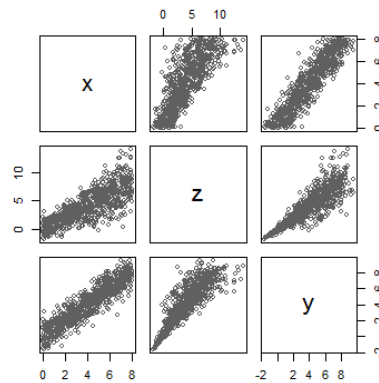
Figure 2.5: Correlation scatterplot matrix

Correlation indexes

Despite the proved effectiveness, plot shapes are not always fully understandable and, as it is common for bivariate representation of process data, can appear messy and confused. For this reason in order to better detect possible correlations it could result necessary to associate visual representation to an absolute correlation index. The correlation between two measures is a statistics which express how much, if the first variable varies, the second is likely to vary itself. The most used parameters capable to quantify the grade of correlation among two variables (x, y) is the Pearson's correlation coefficient, that represent how much the relation between x and y can be described by a linear function:

$$r_{x,y} = \frac{cov(x, y)}{\sigma_x \sigma_y}$$

This statistics is parametric, meaning that it is preliminary assumed that the analyzed variables are normally distributed. Of course this is not generally the case for process variables and for this reason it is worth to use the Spearman correlation coefficient, which is non-parametric and assesses how well the relationship between the two variables can be described using a monotonic function:

$$\rho_{x,y} = \frac{cov(rg_y, rg_y)}{\sigma_{rg_x} \sigma_{rg_y}}$$

Pearson and Spearman results ranges from -1 to 1 which implies respectively full direct or indirect monotonicity. In case that the value is zero it means there is no correlation.

2.2.3 Latent variable techniques: Principal Component Analysis

Chemical processes plant are heavily instrumented by sensors that measure variables at high frequencies. This makes available a large amount of data that, even though they can be considered a great resource, are characterized by a big quantity of correlated and redundant informations that even if relevant for some control systems may not affect significantly the whole process. For this reason a preliminary elaboration aimed to remove unnecessary informations and extract the essential ones is fundamental. A common approach to this so called "*data mining*" is to think the whole variables dataset as a n multi-dimensional space where each measure is represented on a different space dimension. In case that some correlation or redundancy exists, it is possible to state that the effective dimension of the space on which the process data vary can be reduced. Latent variables techniques (LVs) are some algebraic methods that allow to rearrange the multidimensional space to find, in order of importance, those new dimensions called latent variables, on which the information is maximized (Fig.2.6).
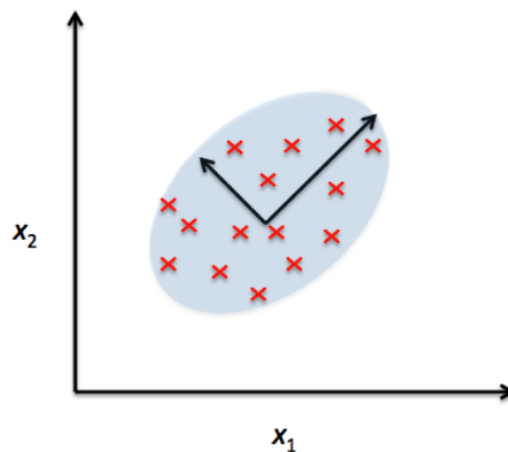
Figure 2.6: Principal components in a two variables data set

Different LVs approaches differ in the way the LVs are selected. After the determination of a proper LVs set it is possible to operate a compression while choosing only the first $(n - l)$ LVs that retain the largest amount of information. By this way problems of process analysis, monitoring, and optimization are greatly simplified and also multivariate systems, that differently are difficult to be represented, can be visually explored. Moreover, a deeper analysis on the LVs and on the transformation matrices allows to perform a correlation analysis on the raw variables.

Principal Component Analysis (PCA) is a popular multivariate latent variable technique that can be successfully applied to analyze and monitor continuous processes. Mathematically speaking, PCA relies on the eigenvector decomposition of the covariance matrix of the X dataset consisting of m measures of the n variables:

$$cov(X) = \frac{X^T X}{m - 1}$$

The eigenvectors of $cov(X)$, the so called principal components (PCs) and are usually determined through a singular value decomposition. PCs are ordered depending on the magnitude of the correspondent eigenvalue so that the first accounts for the space trajectory with the largest possible variance, and each succeeding component in turn, orthogonal to the previous, has the highest variance possible.

This basic approach is called *covariance* PCA but, in case that the X matrix is preliminary column-wise normalized to the standard scores, which is the most common case in process data which do not have a common scale, then all the raw variables are by definition characterized by unit variance. Then equation 2.2.3 coincides with 2.2.2, $cov(X)$ is called correlation matrix and the technique becomes the so called *correlation* PCA. In this case algebraic and geometrical interpretation are the same but the normalization allows to focus the analysis on the correlations without any bias given by the dataset scale.

Choice of the number of PCs

The size of each eigenvalue can be associated to the number of variables which have been lumped in that specific component. By following this definition it is possible then to calculate a PC percent retained variance as the ratio among the size of its eigenvalue and the overall sum of the eigenvalues. Then, if the singles retained variances (%RV) of each PCs are summed up in order, it is then possible to evaluate the cumulative retained variance (%CRV) (see the next example Tab.2.1).

PC	Eigenvalue	%RV	%CRV
1	2.3	46.03	46.03
2	1.61	32.11	78.14
3	0.58	11.68	89.82
4	0.42	8.44	98.27
5	0.09	1.73	100

Table 2.1: Eigenvalues and retained variance (example)

Once the retained variances have been calculated it is then possible to exploit the results to, firstly have an overall evaluation of the amount of the system redundancy and subsequently to select a suitable number of components that is capable to reproduce correctly the system without loosing essential information. Most common selection methods are:

- The first method relies on the CRV chart where, as can be seen from the next figure, in case that the raw system was highly correlated and redundant, it is possible to spot a "knee" in the curve. This knee correspond to a serious drop in retained variance, as can be confirmed in the table (Tab.2.1), and thus the number of PCs correspondent to it is conventionally chosen as a reasonable number of dimensions capable to compress the system.
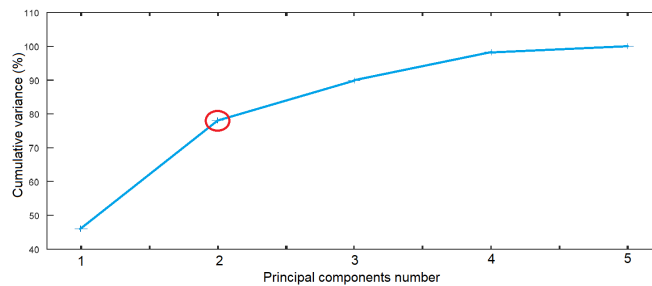
- The choice of the minimum number of the PCs can be done also through a cross validation. Here the principal components are firstly determined on a training set and then the Root Mean Square Error is calculated between the real variables of a validation set and their reconstruction (Eq.2.2) through a k specific number of PCs. The k number of PCs that gives the minimum error on the cross validation is then taken as reference.
- Two other possible "fast" methods define the last good component either as the last that have a retained variance bigger than 5% or the last which have the correspondent eigenvalue bigger than one.

The loading analysis

The PCA procedure allows to represent the X dataset through the new space basis as follows:

$$X = t_1 p_1^T + t_2 p_2^T + ... + t_k p_k^T + E \tag{2.2}$$

Here p are the eigenvectors which correspond to the direction cosines of the orthogonal rotation associated to the PCA transformation and for this reason they have unity norm. If the eigenvectors are multiplied by the square root of their corresponding eigenvalue then they are called *loadings* and they are proportional to the standard deviation expressed by that trajectory. The coordinates t represent the projections of the X raw components in the new reference system and are called *scores* while E, the *residual matrix*, contains all the information rejected while compressing the dataset to a lower dimensionality.

The loadings and the scores are fundamental parameters that allow to analyze the system and characterize correlations among the starting raw variables. Indeed, each element of the loading vector represents the effective contribution of that raw variable on that principal component and this allows first to label that PC as dependent from a single phenomenon (e.g. a PC which is highly contributed by all the temperatures of the process can be labeled as the "process temperature" PC) and then allows to spot correlations among the raw variables depending on the fact that variables which contributes strongly to the same PC (see stars in Fig.2.7) are for sure correlated (or anticorrelated).
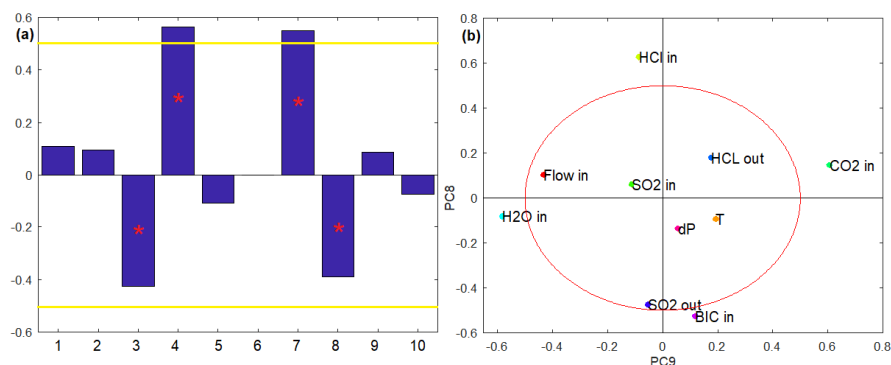
Figure 2.7: Loading plots: (a) Single PC (b) Bivariate

These evaluations can be facilitated with bivariate or 3D loading representation but, in case of high number of variables, it can still be difficult to interpret the contribution plot. In these cases, if two or more PCs appear to be strictly related depending on the mutual presence of important contributions of the same variables, it can be worth to perform a further rotation of that analyzed subspace in order to ease the final interpretation[18]. Among several types of rotations the most common is the orthogonal varimax one which works simply by maximizing the variance on the orthogonal axes of the analyzed subspace.

2.3   Modeling

The last step of the chemometric analysis is the production of a mathematical model for the analyzed chemical system. Modelling procedures are divided in white or black box approaches depending on the amount of a priori information which is available. White-box models are created only on the base of predetermined information while black-box models returns a results but their functional form is neither known nor predetermined. Chemometric modeling is an hybrid approach between the two where, once a plausible structure for the model has been designed starting from the preliminary assumptions, the available data are used to train the model and fit its parameters.

2.3.1   Design of the model

The design of the model usually starts once the variables which affect most the state of the system and their specific influence on the latter are determined. Nevertheless, it can happen that the data analysis is not capable to clarify completely the roles of some variables and in this case the modeling phase need to assume also a validating purpose. Variables which have a specific role in the prediction of the system state are called predictors and, once they are arranged inside the model, they build, together with the constant parameters, its structure. A part from these, the model is defined also by its hyperparameters, constant values that define its structure and, whereas they are external to the model, cannot be estimated directly from the data. This is the case for example of the number of features in a linear regression or the exponents grade in a polynomial regression.

Modeling procedure

Once all the internal and external parameters have been individuated they need to be estimated so that the model is capable to reproduce reliably future data.
This is done through three steps:

1. **Training:** the parameters of the all the possible models solutions are statistically fitted by minimization of the prediction error on a so called a training data set.

2. **Tuning:** the performances of the different solutions are then compared by evaluating their error on a new independent set called cross-validation set.

3. **Validation:** the solution that ensure the best performance on the cross-validation set needs finally to be tested on a third independent set called validation set in order to be sure that any over-fitting is avoided. The crucial point of this set is that it should never be used to choose among two or more solutions, so that the error on this set provides an unbiased estimate of the generalization error.

Modeling subsets

The previous listed process makes necessary the subdivision of the available dataset in order to check and counter-check all the assumptions on different and independent data series. A common rule for the dataset subdivision is that the training set is made up of the 60-80% of the available measures while the left part is divided among tuning and validations. This subdivision must be done consciously so that all the three independent sets are equally representative of the same dataset probability distribution because, differently, the performances on the singles dataset would not be comparable but dependent on the specific represented population.

Estimation of the model performance

During the training in order to minimize the prediction error or during the tuning and validations procedures, if two or more solutions are available and need to be compared or in case that some hypothesis need to be tested a tool for the estimation of the goodness of the model is necessary. This can be done with some indicators that have different characteristics:

- **Coefficient of determination R^2**
  This coefficient is largely used in the context of statistical models and accounts for the proportion of the variance of the real variable, taken from the dataset, which is predicted from the modeled variable:

$$R^2 = 1 - \frac{SS_{res}}{SS_{tot}} = 1 - \frac{\sum_i (y_i - f_i)^2}{\sum_i (y_i - \bar{y})^2}$$

  Where SS_{res} and SS_{tot}, defined as above, are respectively the residual sum of squares, which accounts for the model missed variance, and total sum of quares which accounts for the overall variance of the real predicted variable.
  R^2 gives an absolute estimation of the goodness of the model with a maximum of 1, which correspond to full prediction, decreasing values from 1 to 0 which stand for a decreasing capacities of prediction and values below 0 whcih mean that the model is not working.

- **Errors**
  Mean Square Error (MSE), Root MSE (RMSE) and Mean Absolut Error (MAE) are all possible estimation of the mean distance among the real measures and the ones predicted from the model. Errors are sensible to the scale of the measure and allow only to compare different solutions.
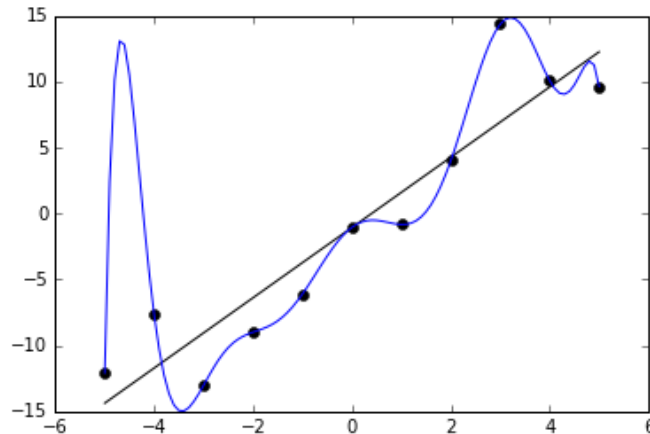

Figure 2.8: Real trend of the data (black) and overfitting model (blue)

Over-fitting is a common misfitting phenomenon (See Fig.2.8) which is spotted if the model is wrongly designed so that it produces optimal predictions on the training set but it predicts poorly new additional independent datasets such as future observations. This is common in those models where the number of features is over sized related to the effective variability of the modelled system.

2.3.2 Empirical model from the literature

Empirical models for the studied chemical system can be already available from the literature. In that case, if the features of the model can be estimated directly from the available dataset it is possible, through a statistical fitting procedure, to estimate its parameters. Then its performance can be evaluated through the mentioned methods and, in case of bad or misleading results, the model can be uploaded and further validated depending on available informations.

2.3.3 Multiple linear regression

If an empirical model is not available, among the broad range of multivariate modeling techniques, Multiple linear regression (MLR) is one of the most simple and easy interpretable models. It is a hybrid model where the structure is assumed to be linear a priori but the constant parameters are estimated from the process data:

$$Y = \Theta_0 + \Theta_1 F_1 + \Theta_2 F_2 + ... + \Theta_n F_n$$

Table 2.2: Linear regression model

In the previous equation Θ_i are the model parameters while F_i are the so called features of the model. Notice that Θ_0 is a constant parameter called 'bias term' which does not refer to any feature.

Partial least squares regression

Linear regression models that works with large numbers of variables are often subjected to over-fitting phenomena which, given that the model features should be the on-purpose selected predictors, cannot be solved with the discard of any of those, making then impossible to prevent this misfitting phenomenon.

This problem can be solved while taking advantage of the possible rotation of the dataset space in order to concentrate as much information as possible in the first k principal components which guarantee the model to not overfit the predicted variable. The key point of this linear modeling procedure, called Partial Least Square regression (PLS), is that the rotation is performed on the predictors variables space not only while pursuing the maximization of the variability on this space but also while maximizing the correlation of the the found PCs with the matrix of the dependent variables, which in case they are more than one are decomposed on their principal trajectories too.
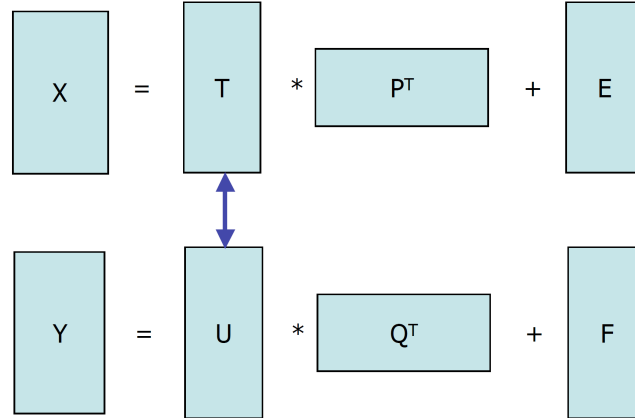
Figure 2.9: PLS matrix representation

Formally, if X and Y are the independent and dependent variables sets represented on the raw reference system, then they can be re-written generically as $X = TP^T + E$ and $Y = UQ^T + F$. In case of PLS, T and U are the variables sets represented in the new compressed reference system which is chosen so that:

$$max[corr^2(U, T), var(U) * var(T)] \tag{2.3}$$

Once this new "modeling conscious" reference system has been determined, the number of features of the model is chosen as an hyperparameter and all the discarded information is condensed inside the residual matrices E and F.

PLS produces very simple models that, if correctly validated, can give good results with a very low computational expense. Moreover, in addition to the system state prediction, the model features and parameters can be exploited for an integrative investigation of the dataset through the analysis of the loadings and of the θ weights referred to the starting raw variables, which quantify the grade of proportionality among that predictor and the system state.

The θ weights are obtained while considering that the PLS model written as $Y = HQ^T$ (with Q the set of chosen PCs and H their respective weights) can be written equivalently as $Y = X\theta^T$. From this equivalence θ is seen as the projection of Q on the raw system and can be written as $\theta^T = PQ^T$, where P is the loading matrix compressed to the maximum k PCs spotted with the tuning of the model.

Chapter 3

Case study

This work is based on the process data obtained from the Silea S.p.A. waste to energy plant located in Valmadrera, Como, Italy. All the reported plant specifications are obtained from [10]. The plant works on two lines (line 1 and line 3) which can treat respectively 6 waste ton per hour and 9.5 waste ton per hour, mainly divided in municipal waste and treated biomedical waste. Both lines are based on the same layout: furnaces with heating recovery systems and APC systems which are composed by a DTS, a NO_x removal section and then a final wet scrubbing section.

3.1 Acid gas removal

Acid pollutants are removed in both lines with a double reagent injection (Fig.3.2) but only one de-dusting stage. The first injection consists of a solid alkaline powder named Depurcal^{TM} MG, a dolomitic sorbent where $Ca(OH)_2$, thanks to the properties of magnesium oxide (MgO), which actually does not take part to the neutralization reaction, can be sprayed directly inside the furnace at high temperatures (800-1400°) without the relative drawbacks. The pre-treated flue gas with exhausted lime is then sent, after heat recovery, to the reaction tower of the DTS where activated carbon and bicarbonate powders are sprayed with a pneumatic transport. Here after a mean residence time of 2 seconds, treated stream is then sent to a fabric filter for de-dusting. Fabric filter in this case is constituted by several internal cells which works independently, allowing then to exclude one of them for maintenance without stopping the equipment. Singles cells cleaning is done instead during the operations with compressed air through pulse jet technique that uses full immersion valves. The reaction continues on the solid which accumulates on the membrane until the residues are removed, gathered in a heated hopper and carried to a storage silo. This work is aimed to the analysis of the bicarbonate DTS stage comprehending the plant section which starts after the pre-heating system and ends after the fabric filter.
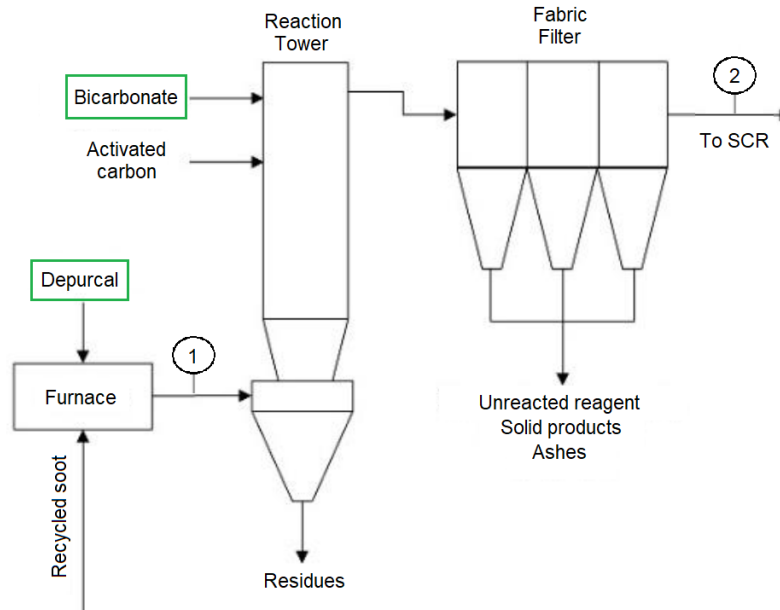
Figure 3.1: DTS section of the Valmadrera plant

3.2   Process data

The chemometric analysis performed in this work is based on the availability of process data from line three. Here, in addition to many process variables measurement systems, three FTIR sampling points are positioned after the furnace (before the DTS reaction tower), after the fabric filter and before the chimney. The first (FTIR1) and second (FTIR2) in particular, which are reported in figure (Fig.3.2), are exploited in this work for the analysis of the inlet and the outlet composition to the bicarbonate DTS stage. FTIR instrumentation is capable to detect several different compounds compositions which are used by the Distributed Control System (DCS) to control the dosage of the reagents and then sent to the Emission Monitoring System (EMS). Other essential process data for this work are stored in the DCS and include the flue gas flow rate, its temperature, the bicarbonate feed flow rate and the pressure differential at the fabric filter, actually a measure of solid residues accumulation on the membranes, see the recap in the next table (Tab.3.1). All the available data have been sampled from 27/07/2015 h1.01 to 12/08/2015 h23.38, with a frequency of one minute, for an overall amount of sampling for each variable of 24419 measures.

T	Gas	$NaHCO_3$	dP	HCl	SO_2	HF	CO	NO_2	CO_2	O_2	H_2O
°C	Nm3/h	kg/h	kPa	mg/Nm3	mg/Nm3	mg/Nm3	mg/Nm3	mg/Nm3	% vol	% vol	% vol

Table 3.1: Process variables and units

3.2.1 Data pre-treatment

This work has started with a preliminary analysis on the quality of available data. From a first evaluation it is clear that several measures are missing due to temporarily mis-functioning, voluntary stops or lack in the storage of the data. This is true especially in case of FTIR measures which are characterized by several windows of missing numbers or zeros measures (spotted by the outlier seeking algorithm), which last from some minutes to even 6 hours. In particular in the analyzed period it is possible to spot:

- A big series of very brief stops which are randomly encountered in both the FTIR sensors.

- The periodical occurrence, each 646 minutes of windows of missing numbers of 9 minutes firstly in FTIR1 and then after 53 minutes of 10 minutes in FTIR2. Those are probably, due to the regular periods and windows length, an automatic reset or calibration.

- Starting from the 03/08/2015 h 09.30.00 to 06/08/2015 08.40.00 a series of long missing measures windows which occur alternatively or simultaneously in both the FTIR sensors. These are probably due to an unwanted stop and necessary fix or to a wanted maintenance on the instrumentation.

While the first two types of windows can be substituted with the reconstruction algorithm quite accurately, the long windows cannot obviously be reconstructed because of their width and for this reason the whole dataset between the 03/08/2015 h 09.30 to 06/08/2015 h 08.40 need to be discarded. This so created big hole in the data availability however may be exploited in a modeling perspective for the creation of the needed modeling subsets.

Outliers removal, reconstruction and filtering algorithms

After the selection of the dataset, its cleaning has been pursued through two main algorithms aimed to the detection and conversion of outliers in missing measures (NaN) and to the subsequently substitution of all the NaN values with a new value, consistent with the previous and the next terms. This procedure results to be necessary only on the FTIR compositions measures since process data from the DCS (T, gas feed, $NaHCO_3$ feed, dP) are highly accurate and do not seem to present anomalies.

The outlier removal is done in two main steps. Firstly, for each variable set, a statistical distribution of the measures is estimated and then, once the probability distribution function is available, it is possible to choose, either manually or through an algorithm, a tolerance factor which is the minimum probability for a measure to not be an out lier. In this case the tolerance factors has been chosen manually for each variable starting from a visual representation of the data. See the algorithm in appendix B.

Compositions measures, due to the high heterogeneity of the burned waste, are highly variable and usually present high and sharp peaks while, on the contrary, it is common for outliers to show up as zeros or very low unjustified measures. For this reason the probability distribution function is fitted as a Rayleygh function which, as can be seen from the next figure (Fig.3.2), is zero in zero and more tolerant on the high measures.
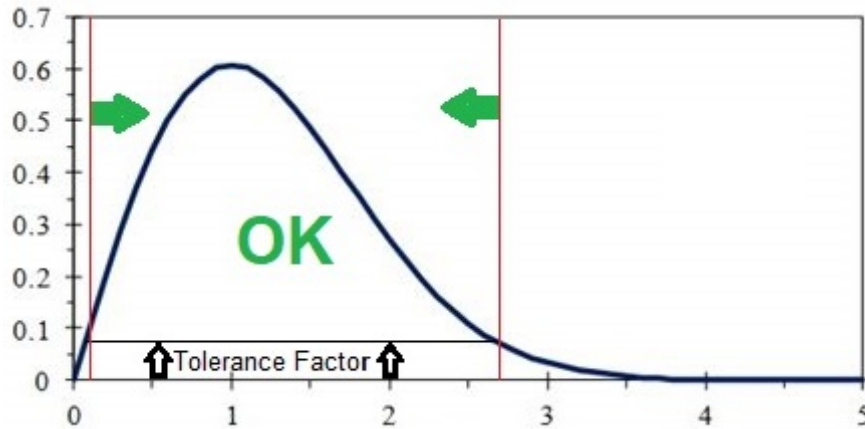
Figure 3.2: Rayleigh probability distribution function

The reconstruction of the data   is then done through the scanning of the singles variable sets. Every times the NaN removal algorithm found a an hole in the dataset it measure its width and it choose three variables before the starting and after the finish of the window. This data are then used to determine an interpolation curve (with interp1 Matlab function) that allows finally to calculate all the missing values as the ones which are more likely to joint the hole extremes. Nota bene that this algorithm is tailored on this specific dataset where it is known that data holes are rare and not nearer than three measures. This allows to estimate the interpolation curve without any check on the effective availability of three measures before and next the hole but, in a more general interpretation, the latter must be provided. See the algorithm in appendix B. Find in figure A.1 an example of data reconstruction.

Data filtering   is finally applied on the data after the data reconstruction to remove random errors. This is done through a moving average with a window width of 10 minutes, a good compromise between noise removal and the left true signal. Find in the table (Tab.A.1) a recap of the cleaning operations.

3.2.2 Dataset subdivision

Starting from what is stated in section 2.3 the dataset needs to be divided in three independent subsets that are equally explicative of all the possibles states of the systems. Nevertheless, this requirement follows a general modeling standard which, given that in this case the available data are time series not produced on purpose, cannot be precisely satisfied in terms of represented population if the three subsets are obtained by the simple splitting in consecutive time series.

In order to obtain the best possible subdivision a distribution analysis needs to be performed on such variables that are as most explicative of the process as possible. At this purpose the assembled variables defined in 1.2.1, the *rs* and the conversion, are supposed to be appropriate.

The dataset, as it is justified in 3.2.1, is already subdivided in two main subsets due to the missing of important informations from the 03/08/2015 h 09.30 to the 06/08/2015 h 08.40. Unfortunately, the distribution analysis on these two preliminarily subsets gives soon a first problem and, as can be seen from the figure (Fig.A.2) especially in the case of the conversion, the two subsets cannot be overlapped in terms of probability distribution. This misfitting, as it is clearer in the time series, is due to an anomaly in the second set, from the 10/08/2015 h22.36 to the 11/08/2015 h07.36, due to unusual low dosage of bicarbonate related to the amount which should be needed depending on the inlet amount of acid. Here in fact the *rs* persists stable under the unity for about nine hours and as a consequence the DTS experiences a substantial drop in conversion. After this phase a sharp peak in the *rs* is experienced, meaning that at that point the systems restart to work, it detects an anomalous amount of acid in the outlet and thus it feed an high amount of bicarbonate.

Concerning the modeling procedure this part of the dataset is not in compliance with the required standard and for this reason needs to be discarded. These so created subsets have the probability distributions represented in figure A.3. As it is clear subset 2 is still biased from some operations problems which do not allow to have a stable distribution of the rs and thus of the final conversion. After a detailed study these anomaly has been ascribed to a possible inconsistency among the measure of the fed bicarbonate, which is an input to the process, and the real dosed bicarbonate which, in case of malfunctioning of the dosing system can be different. Starting from this possible explanation subset 2 have been completely discarded with a final result which can be seen in figure 3.3 which represent three main subsets:

1. **Training set:** from 27/07/2015 h1.01 to 04/08/2015 h7.48 and from 06/08/2015 h 08.40 to 08/08/2015 h 11.29 correspondent to the 66.7% of the dataset.

2. **Cross-validation set:** from 08/08/2015 20.07 to 11/08/2015 h14.46 correspondent to the 22.3% of the dataset.

3. **Validation set:** from 11/08/2015 h14.47 to 12/08/2015 h 23.38 correspondent to the 11% of the dataset.
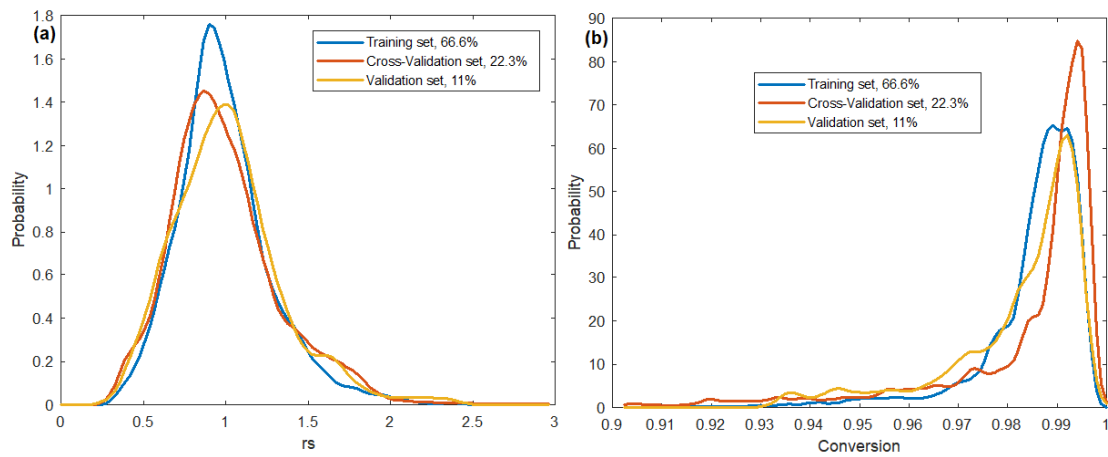
Figure 3.3: Probability distributions: (a) rs (b) χ

This so obtained distributions are, as expected, not perfectly comparable but this result is considered acceptable and the 66.7-22.3-11% dataset subdivision is perfectly within the standard ranges. However, further checks will be carried on once more advanced results are available from the data analysis.

Chapter 4

Data analysis

The data analysis starts after the assembling of the dataset based on some preliminary assumptions:

- All the non-acid outlet compounds (CO, NO_2, O_2, CO_2, H_2O) and the inlet CO and NO2 are irrelevant for the process and are then neglected.
- The oxygen, even if it takes part in the sulfur dioxide neutralization, it is usually neglected on the base of its high excess. In this case however it is maintained because it is an important estimator of some correlations trends hidden among the inlet variables which depends on the combustion process and need to be spotted and eventually ignored if irrelevant.
- The carbon dioxide is involved in the bicarbonate decomposition and it is then necessary to evaluate if its presence among the combustion products is relevant for the DTS performance.
- Water is involved in many of the process side reactions which, together with the moistening of the powder, could influence the DTS performance and need then to be analyzed.
- The temperature and the pressure differential have been selected from the available process database because they could represent a good information source to be exploited respectively in the reactor kinetic analysis and in the influence of the fabric filter on the whole abatement performance.
- All the left variables are supposed to be influential on the process.

4.1   Distribution and preliminary correlation analysis

The left variables are supposed to be influent on the DTS system but needs to be analyzed through a distribution probability analysis in order to evaluate their measure quality, the amount of carried information and thus determine their capability to be correlated each other. In the next table (Tab.4.1) the mean (μ), absolute (σ) and relative standard deviation (σ^*) together with the kurtosis (k) are reported to assess the statistical distribution of each variable.

Section	Variable	Unit	μ	σ	σ^*	γ	k
Inlet:							
	Gas	Nm3/h	$6.17*10^4$	$4.65*10^3$	0.06	-1.17	8.6
	T	°C	190	2.16	0.01	-0.10	1.6
	HCl	mg/Nm3	$1.00*10^3$	209	0.21	1.05	3.4
	SO_2	mg/Nm3	18.8	18.8	1.00	2.90	13.6
	HF	mg/Nm3	4.14	4.80	1.16	6.92	51
	CO_2	% vol	12.3	0.57	0.05	-0.04	1.4
	O_2	% vol	6.23	0.60	0.10	0.14	1.6
	H_2O	% vol	18.0	1.74	0.10	-0.33	0.3
Reactor/outlet:							
	$NaHCO_3$	kg/h	144	53	0.37	0.72	0.6
	dP	kPa	150	3.3	0.02	0.26	31
	HCl	mg/Nm3	15.9	14.5	0.91	2.50	7.1
	SO_2	mg/Nm3	2.52	1.85	0.74	2.42	14.6
	HF	mg/Nm3	0.12	0.81	6.46	7.83	65.5

Table 4.1: Variables and their statistics parameters

After a standard score normalization the mutual correlation can be visually analyzed through a scatter-plot matrix (Fig.A.4) and parametrically by means of bivariate correlation Spearman coefficients. Find in the next paragraphs the results and comments of these applied techniques.

Inlet gas flow

	T	$HCl_{(in)}$	$SO_{2(in)}$	$HF_{(in)}$	$CO_{2(in)}$	$O_{2(in)}$	$H_2O_{(in)}$	$NaHCO_{3(in)}$	dP	$HCl_{(out)}$	$SO_{2(out)}$	$HF_{(out)}$
$Gas_{(in)}$	0.32	0.21	-0.03	-0.06	0.01	-0.03	-0.04	0.17	0.28	0.07	0.07	-0.03

From a first visual inspection it is clear that the inlet gas flow measures lies generally on a squeezed band. The low variability, confirmed by the high kurtosis, is due to the process control operations and makes more difficult to spot correlations even with more advanced techniques.

From the correlations analysis it is possible to spot that the only clear but obvious effects are experienced on the pressure drop in the fabric filter and on the bicarbonate dosage in the reactor while all the left measures, in particular T and $HCl_{(in)}$, are possibly related but need a further analysis.

Inlet acids

	$Gas_{(in)}$	T	$HCl_{(in)}$	$SO_{2(in)}$	$HF_{(in)}$	$CO_{2(in)}$	$O_{2(in)}$	$H_2O_{(in)}$	$NaHCO_{3(in)}$	dP	$HCl_{(out)}$	$SO_{2(out)}$	$HF_{(out)}$
$HCl_{(in)}$	0.17	0.06	1	0.20	0.23	-0.16	0.016	-0.22	0.45	0.11	0.18	0.13	0.13
$SO_{2(in)}$	0	-0.16	0.20	1	0.49	0.18	-0.15	-0.08	0.30	-0.05	-0.48	0.13	0.01
$HF_{(in)}$	-0.06	-0.13	0.23	0.49	1	-0.07	0.09	0	0.14	-0.03	-0.23	0.06	-0.01

The three inlet acids compositions are analyzed together depending on their high mutual correlation, confirmed also by their obvious direct proportionality with the bicarbonate dosage due to the control system based to one or all these three compositions. HCl, as it is confirmed from table 4.1, is present in larger quantities than the other two acids and its kurtosis confirm the visual impressions of a good variability.

On the contrary $SO_{2(in)}$ and HF result from the distribution chart lower and squeezed, which is confirmed, especially in case of HF, by the very high kurtosis. This means that burnt waste is less rich in compounds which produce these kind of products but many peaks of concentration are experienced. As can be clearly spotted from the bivariate plots, while each single acid is obviously correlated with its outlet measure, $HF_{(in)}$ and especially $SO_{2(in)}$ demonstrate a strong anti-correlation with $HCl_{(out)}$.

Other inlet variables

	$Gas_{(in)}$	T	$HCl_{(in)}$	$SO_{2(in)}$	$HF_{(in)}$	$CO_{2(in)}$	$O_{2(in)}$	$H_2O_{(in)}$	$NaHCO_{3(in)}$	dP	$HCl_{(out)}$	$SO_{2(out)}$	$HF_{(out)}$
$CO_{2(in)}$	0.05	0.13	-0.16	0.17	-0.07	1	-0.93	0.44	0	-0.09	-0.18	-0.13	0.21
$O_{2(in)}$	-0.07	0.16	0.02	-0.15	0.09	-0.93	1	-0.27	-0.07	0.07	0.14	0.14	-0.15
$H_2O_{(in)}$	-0.04	-0.04	-0.22	-0.08	0	0.44	-0.27	1	-0.15	-0.14	-0.03	-0.08	0.23

Among these variables carbon dioxide and oxygen show the most eye-catching relation in the whole matrix plot and it is also clear that a linearity of these two takes place with the water inlet composition. All the three measures, accordingly to their distribution parameters and to the visual inspection, have a good variability. This results in many spotted possible relations that needs a further analysis.

Temperature and pressure differential in the fabric filter

	$Gas_{(in)}$	T	$HCl_{(in)}$	$SO_{2(in)}$	$HF_{(in)}$	$CO_{2(in)}$	$O_{2(in)}$	$H_2O_{(in)}$	$NaHCO_{3(in)}$	dP	$HCl_{(out)}$	$SO_{2(out)}$	$HF_{(out)}$
T	0.31	1	0.06	-0.16	-0.13	-0.13	0.16	-0.04	0.03	0.15	0.16	0.07	-0.02
dP	0.28	0.15	0.11	-0.05	-0.03	-0.09	0.09	-0.14	0.04	1	0.09	0.02	0.12

The temperature, even though it is low variable, given its low k and γ can be considered well conditioned. On the contrary, the pressure differential has a very high k and this can justify the poor results for correlation coefficients. In addition to the ones previous mentioned (dP vs $Gas_{(in)}$ and T vs $Gas_{(in)}$) some other possible, but not clear, relations are experienced but not confirmed with many variables. The two variables do not demonstrate preliminarily an influence on the outlet acid but a further analysis is required.

Outlet acids

	$Gas_{(in)}$	T	$HCl_{(in)}$	$SO_{2(in)}$	$HF_{(in)}$	$CO_{2(in)}$	$O_{2(in)}$	$H_2O_{(in)}$	$NaHCO_{3(in)}$	dP	$HCl_{(out)}$	$SO_{2(out)}$	$HF_{(out)}$
$HCl_{(out)}$	0.07	0.16	0.18	-0.48	-0.23	-0.19	0.14	-0.03	-0.12	0.09	1	0.61	0.16
$SO_{2(out)}$	0.07	0.06	0.13	0.13	0.06	-0.13	0.14	-0.08	-0.14	0.02	0.61	1	0.01
$HF_{(out)}$	-0.03	-0.02	0.13	0.01	-0.01	0.20	-0.15	0.23	0.01	0.12	0.16	0.01	1

Outlet acids are obviously related each other and are, as a natural consequence of their abatement which makes all the three measures are highly squeezed toward low measures, very bad conditioned. This can produce doubtful or, as in case of HF, useless results. HCl and SO_2 confirm all the correlations already mentioned.

4.1.1   Discussion

The distribution analysis suggests an overall good quality of the dataset which implies that many of the correlation results can be considered reliable. This is not the case of some variables which, due to process control (inlet flow and dP) or intrinsic properties of the waste and abatement process (Inlet/outlet SO_2 and HF), are almost constant and make the correlation analysis more doubtful. In case of HF the variable results almost useless and, depending on its low average quantity and the high mutuality with the SO_2, makes its removal from the dataset highly suggested. The Spearman coefficients shows several possible correlations among the inlet variables which deserve a further and deeper analysis in order to spot, if possible, eventual phenomena that takes place before the system and can then be isolated to perform a more precise regression analysis on the outlet variables. Concerning the reactor it is clear that the inlet SO_2 and HF are anti-correlated with $HCl_{(out)}$ and thus with the DTS performance. This is indirectly dependent on the control system operations because these two acids, evaluated as equivalent HCl, induce an over feeding of bicarbonate which however, due to the higher affinity of the HCl for the sorbent, ends in an increased abatement of the latter. Both O_2 and CO_2 show a comparable but opposite correlation with the outlet acids that suggests an possible indirect involvement of these two in the overall performance.

4.1.2 The inlet stream

Given the high number of variables and the clear existence of several correlations among the inlet ones ($Gas_{(in)}$, T, $HCl_{(in)}$, $SO_{2(in)}$, $CO_{2(in)}$, $O_{2(in)}$, $H_2O_{(in)}$), the latter are processed independently by means of a PCA so that it is possible to have a quick and preliminary characterization of all those phenomena which do not depends on the DTS operation but are a consequence of the furnace working, allowing then to identify and isolate those in the next part of the analysis. The loading analysis, as confirmed from the spotted knee in the next chart (Fig.4.1), is performed on the first three principal components which have the eigenvalue larger than one, with an overall retained variance of 74%.

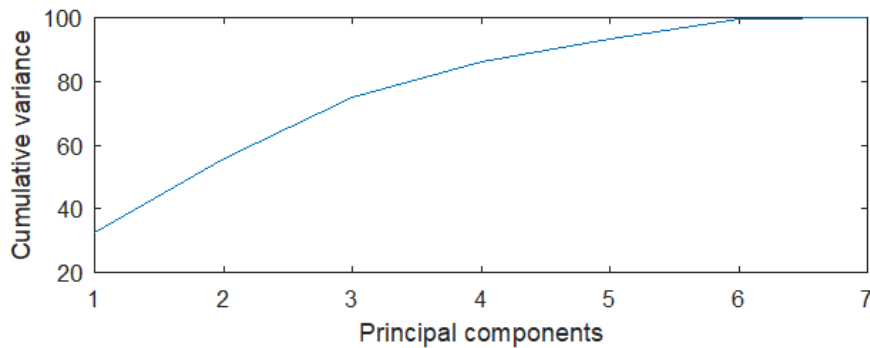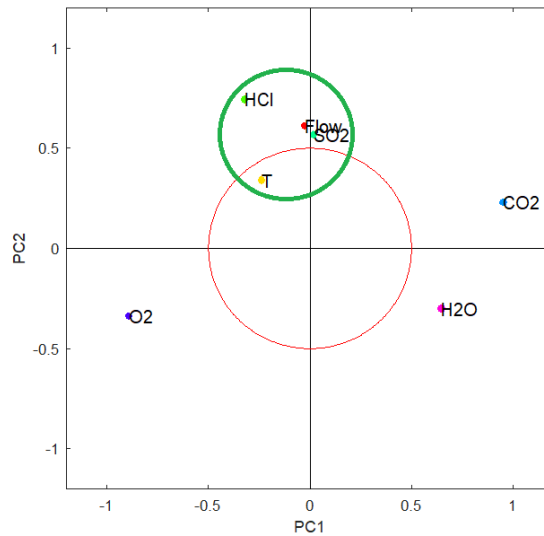| N°PC	EigenVal	Variance (%)	Total Variance (%)
1	2.29	32.7	32.7
2	1.61	23.1	55.8
3	1.36	19.4	75.1
4	0.77	11.0	86.1
5	0.51	7.2	93.4
6	0.44	6.2	99.6
7	0.03	0.41	100

Table 4.2: PCA results

Figure 4.1: Total retained variance

Figure 4.2: PC1 and PC2 bivariate loading contribution plot

In the first PC the anti-correlation among $CO_{2(in)}/H_2O$ and $O_{2(in)}$ is associated to the excess dosage of air in the furnace. In the second PC the clear anti-correlation among the water/oxygen and the spotted green cluster (acid composition, temperature and inlet flow) could be associated to the furnace operations evolution and the water content of the waste. In the furnace in fact, after the loading of new waste, the combustion energy is primarily employed for the vaporization of the excess water, with then a lower oxygen consumption, a lower amount of flue gas, lower temperatures and acid production. The third and last analyzed PC features the depth development of green cluster and highlights the anti-correlation among the acid composition $(HCl_{(in)}+SO_{2(in)})$ and the amount and temperature of the flue gas. This PC, in contrast with PC2 that seems to be more linked with the amount of burned waste, represents a quality benchmark of the type of burned waste and its calorific value. Find in figure A.7, A.8 and A.9 the associated time series.
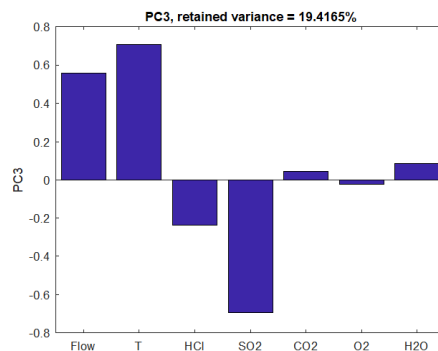

Figure 4.3: Inlet stream: PC3 loading contribution plot

4.2 The reactor

After the analysis of the inlet measures the first step is to proceed with the investigation of the whole dataset in order to characterize the DTS reactivity and produce a conversion model. This would imply, as natural prosecution of the work, the principal component analysis of the whole dataset added of the missing variables which have not been analyzed in the previous section. Here however, given the availability of a conversion model from the literature (section 1.2.1), the first step has been the preliminary application of the latter on the available process data in order to look for any possible interesting outcome for the next analysis.

4.2.1 Applicability of the available conversion model

The results of the available model effectively suggest that Eq.1.1 is suitable to describe solid-gas reactions at the bench scale but its applicability on process scale data needs to be confirmed.
Experimental data in fact come from a controlled environment where reactions are monitored and stable while, in case of process data, operative conditions are not stable and the analyzed information comes from a continuous data stream. Indeed, the applicability of the model has been demonstrated[6, Antonioni, 2013] on process data averaged on a one-day base (Fig.A.11) but, although this is confirmed also for this specific case (Fig.4.4), it does not actually validate the performance for on-line applications.
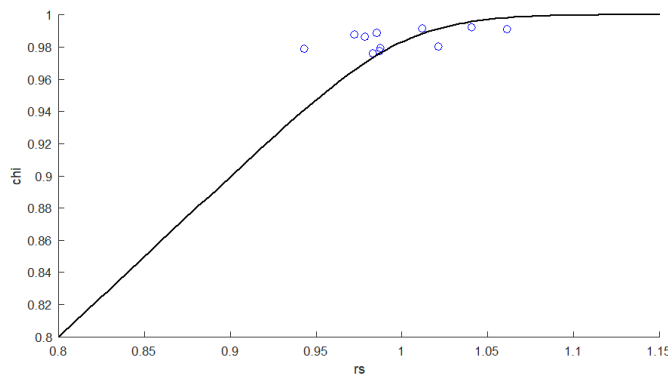

Figure 4.4: Model calibration vs. case study process data (1 day average)

In figure 4.5 the instantaneous rs (1 minute sampling) of this analyzed case study are compared with their statistically fitted conversion model (see appendix B).
It is clear that, even though the proportionality among χ_{HCl} and rs is effectively confirmed, the model cannot be validated. Firstly, data experience a high scattering probably due to the only single fitting parameter which does not allow to consider all the physical phenomena and and briefly changes in operating conditions. Subsequently, as it is clear especially from the fitting chart, the output of the model results to be largely underestimated for a big part of the rs.

This underestimation has been ascribed to the model unsuitability to work with instantaneous measures. Due to its bench scale derivation in fact, this model can work only basing on the preliminary assumption that the reacting acid entering at time t stays in contact with the solid sorbent dosed at time t, and exclusively with it, till the reaction ends. This can still be assumed for large time windows averages, as it is for the previous reported case, but cannot be applied for real time measures since the acid that enters the DTS at time t, due to the low contact time, does not fully exhaust the bicarbonate dosed at the same instant but then, on the contrary, once the gas reaches and passes through the bag filters it reacts with the cake made up by all the reagent dosed, partially unreacted and settled before.
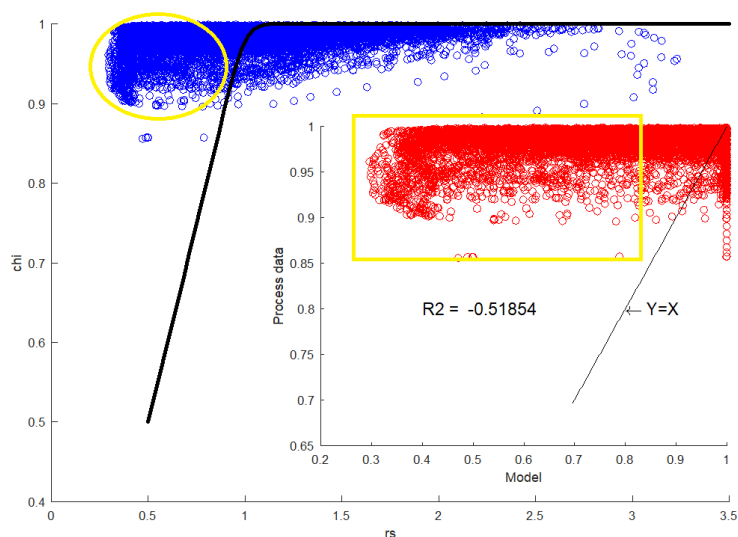

Figure 4.5: Case study instantaneous measures: (a) χ_{HCl} vs. rs chart and (b) Process data vs. model prediction

4.2.2   Filter reactivity potential

What is clear from the application of the available conversion model is that the instant bicarbonate feed recorded by the DCS cannot be directly associated with the acids compositions recorded at the same instant but it is necessary, in order to produce a model which is capable to predict instantaneous measures, to account for the effective overall reactivity potential given also by the settled solid in the fabric filter, which need then to be characterized depending on the capacity and the operations of the filter.
The settled solid reactivity however, given that the filter is divided in more sections that are cleaned depending on an unknown procedure, cannot be determined accurately for each instant but need to be characterized instead through a statistical parameter while thinking at the settled reacting sorbent as a stacked sequence of the k preceding inlet bicarbonate feeds, leading then to the definition of an aggregate bicarbonate variable (see next figure 4.6).
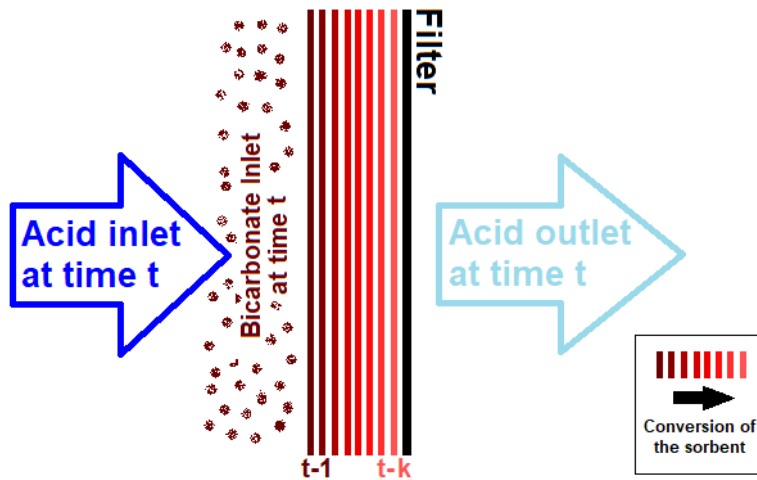
Figure 4.6: The bicarbonate accumulation at time t

By means of this constant parameter the filter becomes then a simplified but active element of the model which needs to be statistically fitted by maximizing the correlation among the aggregate bicarbonate feed and the DTS performance. This statistical fitting has been pursued while evaluating how the singles variables varies their mutual correlation depending on k. In particular a PCA has been reiteratively performed by substituting the instant bicarbonate with the aggregate bicarbonate variable calculated with a larger k at each iteration.

Preliminary PCs labeling

The first step of the characterization has been the labelling of the main PCs of the full dataset. In the aggregate variance plot (Fig.A.12) there are no noticeable knees but the one, barely visible at PC4, which is confirmed in table A.2 where PC5 has the first eigenvalue lower than the unity. These four main PCs which account for the 67.8% of the overall dataset variance and can be labeled, depending on their mono-variate contributions A.13, as follows:

- **PC1:** The first PC is ascribable, as it was for PC2 in the inlet variables analysis, to the furnace operations evolution as it is clear from the anti-correlations spotted in the inlet variables and confirmed by the time series trend.
- **PC2:** The second PC represent the the reactor operations. Here in fact the outlet acids are anti-correlated with the bicarbonate inlet and, as a consequence of the phenomena described before, with the SO_2 inlet.
- **PC3:** The third PC is linked with the quality of the waste and is linked to the PC3 spotted in the inlet variables analysis. As in the previous case in fact, it shows a clear anti-correlation among the acid presence in the inlet and the furnace output in terms of temperature and flow.
- **PC4:** This PC is highly contributed especially by the carbon dioxide.
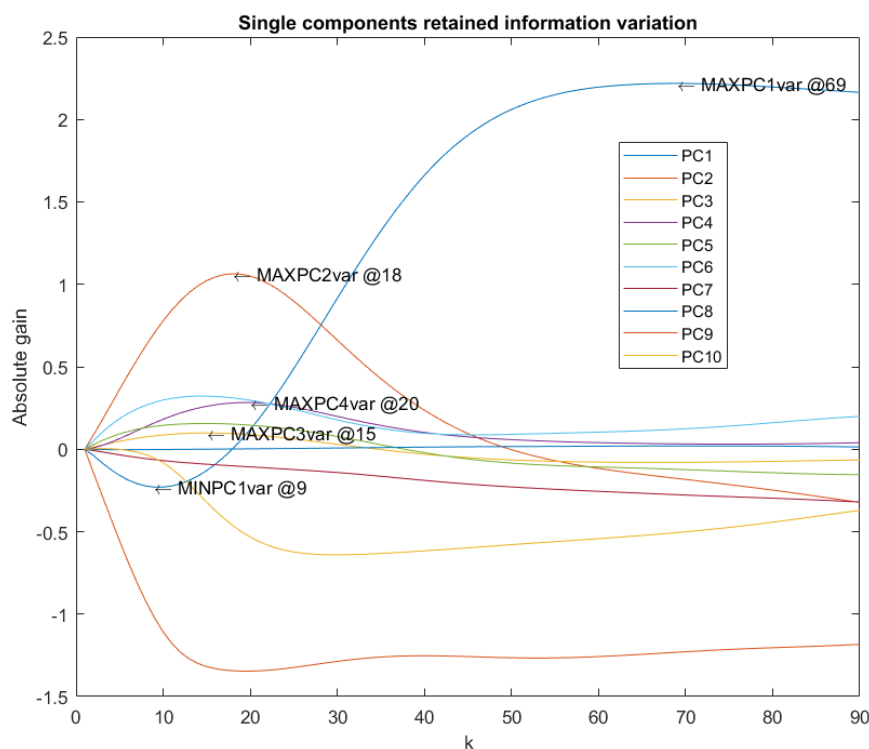
Figure 4.7: PCs absolute gain vs k

The k parameter characterization

The previously mentioned approach is visualized in terms of absolute gain of the
singles PCs in the previous figure (Fig.4.7) which confirms the starting hypothesis. PCs
retained variance is demonstrated to be highly dependent on k and for specific values of
this parameter, as can be seen in figure A.10, the overall information is shifted from the
last "noisy" PCs to the main ones, meaning an overall better correlation of the dataset.
The better consistency of the aggregate variable is in particular confirmed by the trend
of PC2 which is subjected to the highest gain at $k = 18$ where, as evident from the
comparison loading plot (Fig.4.8), the loadings contribution which undergo the highest
gain are the one correspondent to the bicarbonate and the outlet acids. This validates the
employment of k as a parameter necessary to account for the overall reactivity potential
of the DTS. Looking at figure 4.7 together with the loading comparison plot of PC1
(Fig.4.9) it is clear that the behaviour of the first PC is in contrast with PC2. Here
the retained variance primarily drop down till $k = 9$ and then starts to rise till values
of k around 70. From figure 4.9 the key factor of this trend appears to be again the
correlation among the aggregate bicarbonate and the outlet acids. Indeed, the positive
contributions of the outlet acids drops down till $k = 18$ and this effect, given that the
bicarbonate contribution in both PC1 and PC2 is increasing and that the outlet acids
contributions in PC2 is rising, is ascribed to the shift of weight from a positive but false
correlation in PC1 to a true negative correlation in PC2.
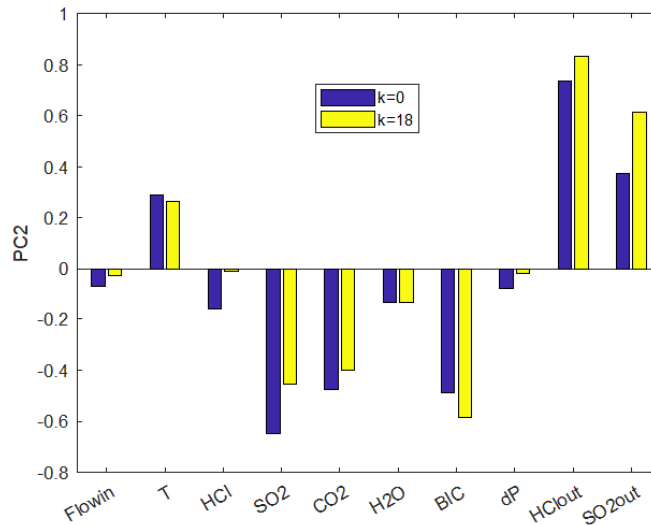
Figure 4.8: Loading comparison for PC2 at $k = 0$ and $k = 18$

The fact that for $k > 18$ PC2 starts to drop variance indefinitely while PC1 has the opposite effect is supposed to be an algebraic effect of the principal component decomposition which, depending on the decreasing benefit of aggregate bicarbonate on the reactivity potential, starts to bring back variance to a stronger but more useless component. Depending on the spotted algebraic phenomena k must be chosen in an interval between 0 and 18 and, given that PC2 is the component which is mainly linked with the reactor potential a value of k around 18 appears to be the most suitable characteristic time to be chosen for the reactivity potential of the filter.
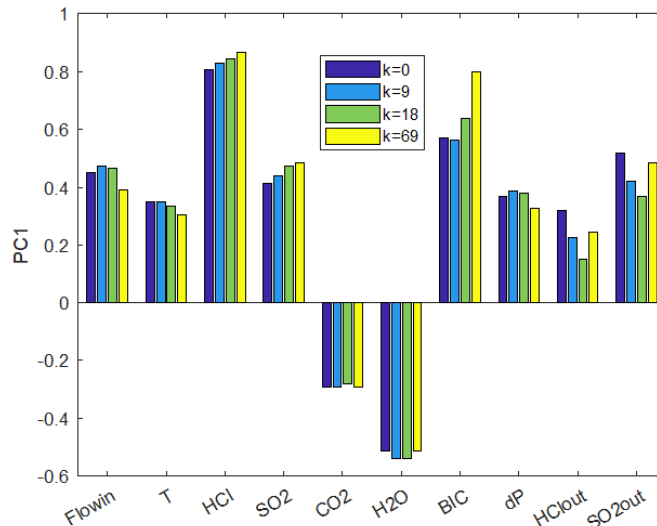
Figure 4.9: Loading comparison for PC1 at $k = \{0, 9, 18, 69\}$

A further validation of this result has been obtained by evaluating the performance of a mono-variate linear regression among the aggregate bicarbonate variable, calculated at different k, and the outlet HCl. As can be seen from the reported figure (Fig.4.10), the R^2 of the regression has a peak at $k = 17$, which is slightly lower than the k spotted previously. This allows to think that some secondary hindered phenomena are present and looking only at PC2 is not fully exhaustive.

The k parameter is then chosen for the next analysis and modeling to be equal to 17 but, given its statistical derivation and all the possible fluctuations in the dataset, its nature should be more likely characterized as a range. In this case then, starting from the peaks spotted for the different PCs in 4.7, the time range among which the settled solid in the DTS fabric filter is supposed to have a good residual reactivity without been cleaned goes from 15 to 20 minutes.
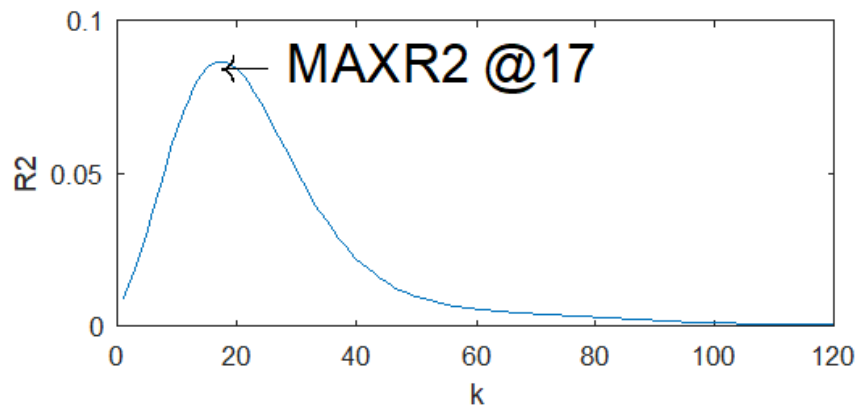

Figure 4.10: R^2 of the monovariate regression of HCl_{out} on BIC_{in}

4.2.3 Principal component analysis

The four PCs maintain the same references as before but now the effective contributions of all the raw variables can be clarified and better analyzed. All the retained variances are reported in the next table 4.3 and plotted in figure A.14. The use of the aggregate bicarbonate lead to a net gain of 2.2% with an overall explained variance of 69.9% in the first four PCs.

N°PC	EigenVal	Variance (%)	Total Variance (%)
1	2.34	23.3	23.3
2	1.87	18.7	42.1
3	1.47	14.7	56.8
4	1.31	13.1	69.9
5	0.91	9.1	79.0
6	0.70	7.0	86.0
7	0.62	6.2	92.2
8	0.39	3.9	96.1
9	0.22	2.2	98.3
10	0.17	1.6	100

Table 4.3: PCA results

Labeling

The preliminary labeling and the associations with the starting inlet analysis are confirmed even after the k characterization and now a deeper analysis is possible:

- The second principal component (RV=18.7%) is confirmed as the only suitable for the description of the DTS performance while the others (1,3,4), given the low contribution of the HCl outlet composition, are associated to the inlet flue gas.

- PC4 (Fig.A.17) (RV=13.1%) is early dissociated from the other two depending on the high carbon dioxide contribution and on the lack of substantial contributes in the inlet flow and temperature. If looking back to the inlet analysis 4.1.2, carbon dioxide contributed strongly only the air dosage PC1 and this, whereas only PC4 looks to be strongly contributed by carbon dioxide, suggest that PC4 inherits the variability trajectory that was of pertinence of PC1 in the inlet analysis. The removal of the oxygen naturally induces a net drop in retained variance and an algebraic rearrangement in the trajectory which is demonstrated by the different contributions of variables such as sulfur dioxide.

- The first and the third PCs (Fig.A.15,Fig.A.16) (RV1=23.3%,RV3=14.7%) confirm their associations to the variability trajectories which were of pertinence of PC2 and PC3 in the inlet analysis. Here however the addition of the outlet and reactor variables decrease the resolution of the labeling and a large series of mutual mid-contributing variables do not allow to understand clearly the real nature of the PCs. This circumstance suggests the application of varimax rotation on the PC1-PC3 plane so that it is possible to maximize the variance on the main axes of the plane with a final bivariate loading contribution plot which is reported in figure 4.11. The orthogonal axes of the new reference system, called T1 and T3, account respectively for the 20.3% and the 17.7% of the overall variance and, depending on the spotted contributions, can be considered representative for two independent characteristics of the inlet flue gas, reflective of the quality and the amount of the gas exiting from the furnace.
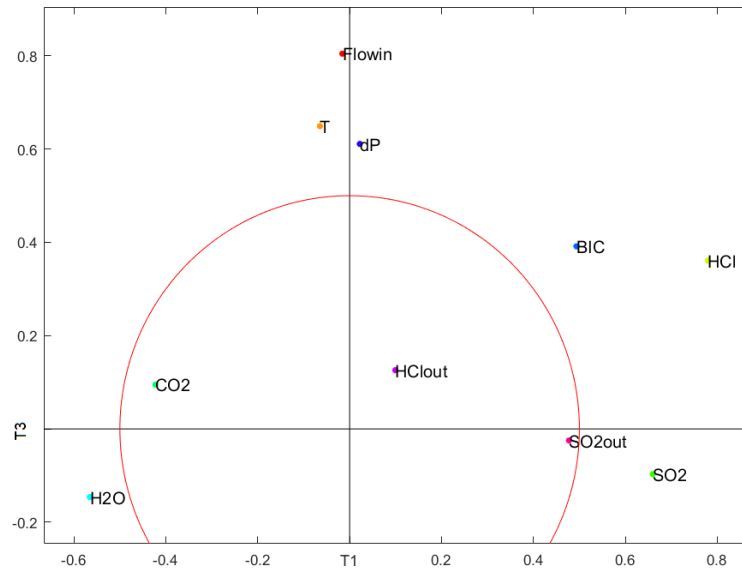
Loading analysis

Figure 4.11: Result of the varimax rotation

The first trajectory T1 (see also fig.A.18) sees the predictable counteraction of the inlet acid compositions to the water and carbon dioxide as a natural effect of the type and amount of burned waste and the related products of combustion.

T3 is instead characterized by the high positive contributions of the inlet flow (with pressure differential) and the temperature, which are assumed to be estimators of the combustion regime in the furnace and the consecutive production of flue gas. The HCl inlet positive correlation on T1 suggests either an higher combustion enthalpy of HCl-producing wastes or that higher temperatures are needed for the combustion of that kind of waste.

The bicarbonate dosage, as a natural consequence of the proportional control, have equal loadings for both trajectories. Together with the outlet acids, the contribution of the carbon dioxide is the lower among the analyzed variables confirming that its composition is almost independent from the other and that its variability is largely described by PC4.
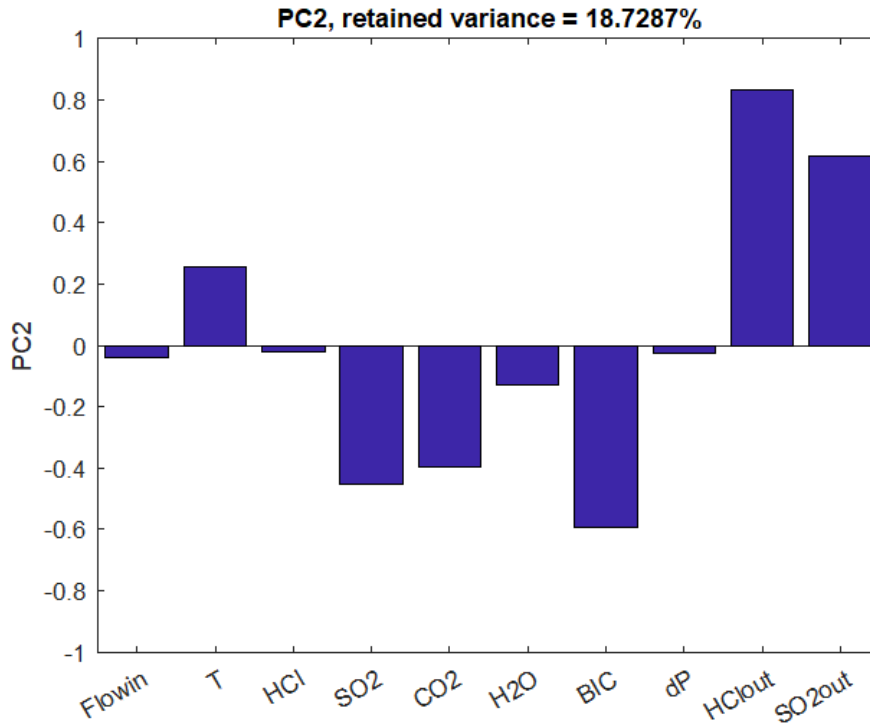

Figure 4.12: PC2 mono variate contribution plots

PC2 represents the abatement performance of the DTS and in particular the loading analysis allows to associate the positive values of this component to the lower DTS working regime. The multivariate analysis substantiate the primary role of the sulfur dioxide, which strong contribution however definitely highlights a basic incoherence problem of this trajectory where the sulfur dioxide participates thus both as removal enhancement agent and as an agent to be removed. On the PC2 trajectory, in addition to the already justified contributions (HCl_{out}, $NaHCO_{3(in)}$, $SO_{2(in)}$), also the carbon dioxide appears to vary strongly, with a clear negative contribution. This cannot be justified on a chemical basis that instead suggests an opposite effect due to the partial employment of the bicarbonate for its neutralization. The contribution should be then ascribed instead to some other hidden phenomena, or pure algebraic effects, which correlates directly or indirectly this variable to the abatement performance.
All the remaining variables contribute with loadings lower than 0.4, meaning that their influence on the final reactivity is not sure. The inlet hydrochloric acid seems to not influence the performance but is obviously maintained as predictor depending on its intrinsic influence on the outlet acid.

4.3   Discussion

The analysis allows to spot and characterize several known and unknown phenomena and, through the principal component analysis, to condense most of the DTS ruling phenomena in four variables which allow a general overview of the DTS system. The assembling of the dataset for the analysis has demonstrated to be fundamental because the choice or the omission of each single variable can lead to difficult interpretation or even upset completely the analysis, making then difficult to obtain clear and absolute results. Two examples of this have been clearly spotted in the roles of the oxygen, for the characterization of the furnace operation, and the one of the sulfur dioxide in the DTS performance. In the first case the preliminary choice to maintain the oxygen in the inlet analysis has been proved to be essential since it allows to highlight a variability trajectory (PC4) which otherwise would have been impossible to label correctly. Differently, in case that the oxygen had been left also in the final analysis it would have contaminated the results. In second place, the fact that in PC2 some variables have doubtful roles, is in part due to the mutual presence of both the outlet acids concentrations in the dataset. PCA in fact, depending on its intrinsic algebraic nature, builds PC2 as a compromise among the abatement potential on the two acids even though each single variable could have different and even opposite effects. This fact is well exemplified by the role of the inlet sulfur dioxide which presents a strong negative contribution on PC2 which has been ascribed to its positive effect on the HCl abatement. Looking at the preliminary analysis in 4.1 however, it is clear instead that its correlation with the outlet SO_2 is naturally positive and hence, depending on this specific effect, its contribution on PC2 should be positive.

These spotted covering effects could be potentially affect each variable and, if not spotted, makes the analysis more difficult. The chemical and process expertise becomes then fundamental to evaluate each result, while the next modeling phase assumes a crucial validating role because it allows to focus independently on each single outlet acid while allowing a final predictors characterization through the analysis of the weights of the model.

Chapter 5

Modeling

The aim of the modeling phase is to produce a mathematical model that is capable to predict with an acceptable grade of accuracy the future data while being interpretable for the validation and improvement of the results obtained in the previous analysis. The partial least square linear method is a good compromise in this sense and allows, through the study of the model parameters and features, a final characterization of the effective predictors influence on the system state. Its simple structure allows an easy and direct design procedure (see chapter 2) which entails only the uniquely definition of the independent and dependent variables matrices, assembled from the available dataset, and the determination of the exact number of features as an hyperparameter through a cross-validation. Nevertheless, both the assembling and the cross-validation require the possibility to test and compare the possible solutions.

5.0.1 The dependent variables

The instantaneous DTS performance is univocally determined by the reactor conversion on the singles acid pollutants. However, the conversion is an assembled variable which, depending on the not completely reliability of the dataset, sums up all the possible errors of the starting raw variables, leading then to a possible drop in model performance. For this reason the final choice is to model directly the compositions of the outlet acids. Whereas the linear model function does not have any co-domain restriction, then the outlet prediction can assume any value. In this particular instance however the prediction must lie within a maximum composition equal to the inlet and a minimum equal to zero. Then a check and correction of the outlet predictions is necessary and lead to a natural increase of the model performance.

5.0.2 The independent variables set

The direct modeling of the acid output suggests the transformation of all the compounds compositions in flows which are directly relatable with the inlet bicarbonate. Nevertheless, the inlet gas flow has been demonstrated to be low variable and almost irrelevant on the DTS performance and this, together with the bad conditioning of some of

the compositions measures, make its embedding almost useless and possibly deleterious. The two approaches are thus tested in parallel. Then, for each of the two methods, the independent dataset need to be correctly assembled only with the sure predictors because useless variables which do not affect the system can still contaminate the PLS prediction. Indeed, even though PLS works on those trajectories that are mainly correlated with the modeled variables, this does not exclude that other variables, which are not predictors but are correlated with one or more predictors, maintain a certain grade of variability on those trajectories while even influencing the spatial direction of the latter. If this is true, then the weights assigned to the modelling PCs are automatically shared also with the contributions of the above mentioned variables, that will then assume a not-null role in the model, while the weights associated to the real predictors can instead being badly influenced. In this case the final assembling of the real predictors set is done starting from the data analysis results with a trial and error procedure while removing gradually, and with different combinations, all those variables which demonstrate to give a worse fitting performance if left in the variables set.

5.0.3   Tuning and Validation

From the PCs time series (Fig. from A.20 to A.23) is clear that the preliminary stackability assumption of the probability distributions (Fig. A.24, A.25) of the three modeling subsets cannot be confirmed. The available dataset does not allow the correct creation of independent and equally representative subsets and this upsets the conventional tuning-validation procedure, making thus the model design more difficult because different configurations can perform alternatively better on different populations. Indeed, the only fitting performance evaluation on the cross-validation set does not allow absolute choices and it is then necessary to integrate the procedure and carry out further checks. In this case, given that any further rearrangement would be ill-judged, it has been decided to evaluate the performance with a "population conscious" approach, meaning that the tuning evaluations are done on the two validations sets, both joined and separated, while privileging those solutions that fits better operative populations of higher interest.
By this way it is possible to have a more objective design without loosing the theoretically required independence among the singles subsets. In this case, in addition to a general predilection for the more variable subsets, depending on the DTS purpose to fulfill upper emission limits, the operative populations of higher interest are identified as the lower DTS regimes that produce higher concentrations.

Subset	μ	σ	k
Training	0.09	1.16	0.90
Cross-Validation	-0.25	1.74	0.24
Validation	-0.04	1.59	-1.08

Table 5.1: PC2 subset statistics

Starting from the data analysis results, PC2 is known to be directly related with the reactor regimes, then its probability distribution chart (Fig.A.25) and its simple moments (Tab.5.1) are analyzed in order to characterize how the DTS works on each single subsets:

- The training set appears to be statistically representative of a smaller process sample, but with an average lower working regime than the validations set. The statistical fitting will then produce a final model suitable for the prediction of a good working regime but with a low reliability on extreme measures.

- The cross-validation set is made up by a larger amount of measures but it is averagely representative of an higher working regime than the validation set. This, together with its lower variability (slightly larger σ but much higher k), makes the performance on the validation set preferable.

5.1 HCL model

All the results are proposed while starting from the evaluation of the model with the complete set of available independent variables with the direct use of the concentrations. This allows to have a preliminary overview of the capacity of the model and, by means of the analysis of the weights of the singles variables, allow to proceed with the next trial and error procedure. As above mentioned, the range correction procedure works by bringing back the output predictions to the allowed range $[0 \; C_{in}]$, the latter however has been demonstrated to act only on the underestimations, while the upper concentrations are always below the maximum allowed values.

5.1.1 Features tuning
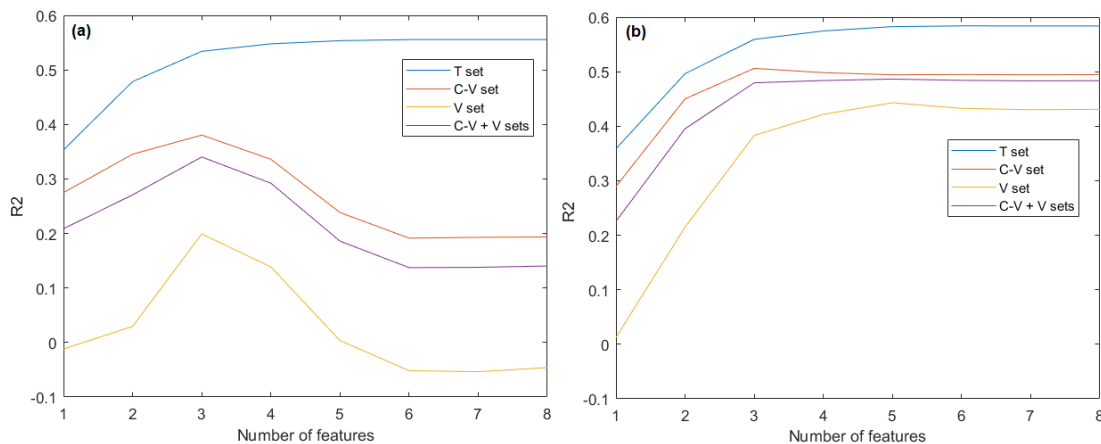
Figure 5.1: Fitting performance vs N°of features: (a) No range correction, (b) With range correction

From the first part of the previous chart it is clear that on the raw output prediction the over-fitting rises on both the validations subsets after three features. Then on the right part the results for the range corrected predictions are reported. Here, in case that the misfitting due to the larger number of features was generalized on the whole concentration range, no variations of shape would be experienced. What appears clear from the spotted plateau is instead that the correction of the underestimations is decreasing and almost eliminating the over-fitting problem. This means that, as can be clearly seen from figure A.26, the increase of features is effectively increasing the misfitting phenomena, but that the latter is localized on the low-concentrations. Moreover, depending on the relative amount of the singles specific populations (high or low concentrations) the singles subsets can even benefit of the higher number of features. This is demonstrated by the fact that, whereas the training of the model is performed on averagely lower regime population, those subsets which are known to be populated by higher concentrations (i.e. T and V sets) appears to increase their final performance (see Tab.5.2), while in case of the CV set a slightly lower performance is experienced. This means that an higher number of features is not only irresponsible of any substantial un-fixable over-fitting but, as it is exemplified in figure A.27, is averagely increasing the performance on those operative populations which are of higher interest. Depending on these observations the best configuration is chosen then while basing on the V set performance peak, with a number of features equal to five.

NF	T	CV	V	CV+V
3	0.534	0.381	0.199	0.340
5	0.554	0.238	0.004	0.186
With range correction				
3	0.559	0.506	0.383	0.480
5	0.583	0.494	0.443	0.487

Table 5.2: R^2 vs N°of features

Choice of the predictors

The variables weights associated with this configuration, reported in figure 5.2, essentially validate the results of the data analysis. Indeed, the most prominent contributions are given positively by the inlet hydrochloric acid and negatively by the inlet sulfur dioxide and bicarbonate. In addition to a conspicuous bias term, all the other variables appear to have negligible weights which are probably due to the residual contributions of the PLS space rotation. However, in order to check this conjecture and finally validate the analysis, a trial and error procedure is carried on while removing gradually, and with different combinations, the mentioned variables.
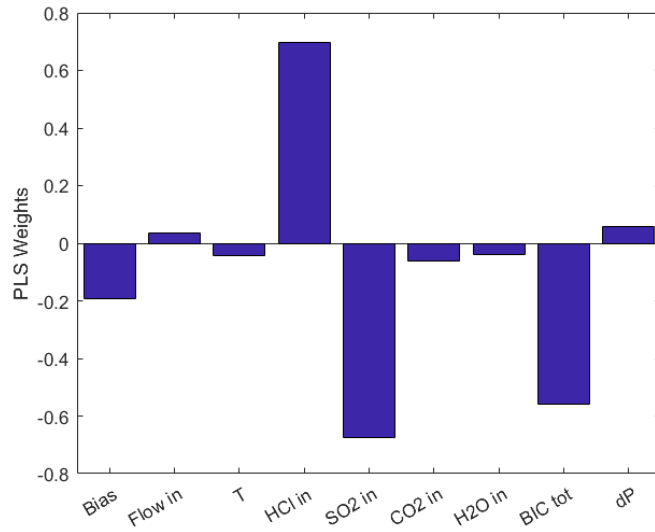
Figure 5.2: Variables weights for the five features - eight variables PLS model

Indeed, given their known mutuality in several directions, removing singles variables could not produce the same desired performance increase which is given by their combined removal. Each configuration is estimated after the cross-validation for the estimation of the optimal number of features which is simultaneously reported. All the differences independent sets performances are compared to the reference with the full set and, depending on the better or worse performance, are highlighted respectively in green or red.

Removed Var.	N°F	Concentrations				N°F	Flows			
Reference	5	0,583	0,494	0,443	0,486	5	0,584	0,494	0,441	0,486
Flow (F)	4	0,579	0,495	0,435	0,485	5	0,580	0,489	0,427	0,479
T	5	0,583	0,496	0,447	0,489	5	0,584	0,496	0,445	0,489
H2O	3	0,570	0,504	0,431	0,490	3	0,574	0,498	0,433	0,487
CO2	5	0,583	0,497	0,441	0,488	4	0,584	0,495	0,443	0,488
dP	5	0,582	0,491	0,440	0,483	5	0,583	0,491	0,444	0,485
T CO2	4	0,583	0,500	0,441	0,490	3	0,576	0,500	0,461	0,496
T dP	4	0,578	0,498	0,443	0,489	5	0,582	0,494	0,446	0,488
CO2 dP	4	0,579	0,499	0,459	0,494	4	0,580	0,499	0,453	0,493
T CO2 dP	2	0,563	0,504	0,465	0,499	2	0,553	0,499	0,467	0,497
F T CO2 dP	2	0,563	0,504	0,460	0,498	2	0,557	0,496	0,475	0,497
T H2O CO2 dP	2	0,562	0,513	0,447	0,501	2	0,554	0,508	0,452	0,499
F T H2O CO2 dP	2	0,561	0,513	0,440	0,499	2	0,557	0,506	0,457	0,499

From the previous preliminary results is possible to assume as follows:

- The configuration that uses the flows produces better results, especially on the V set, meaning that the preliminary assumption was correct and that the quality of the data does not affects strongly the fitting performance.

- The variables removal entails a gain of performance on both the validations set and a slight drop on the training set, meaning that the unnecessary variables produce an over-fitting effect.

- The net decrease in the number of features and the substantial drop in fitting performance characterize the water as a predictor. This is confirmed also by the consecutive test on possible combinations where, in case of its removal, the performance drops.

- The removal of the temperature, pressure differential and carbon dioxide guarantee, especially in combination, a net gain in performance (6.3% on the Vset with the T CO2 dP removal), which allow these three variables to be definitely removed as possible predictors. While for the first two variables the result was expected, in case of the CO2 this result demonstrates that its PC2 negative contribution spotted is either referred to the SO2 or due to other un-justified algebraic effects.

- The best performance on the V set is attained with the further removal of the inlet flow, which demonstrates then to not be a predictor.

5.1.2   Results and discussion
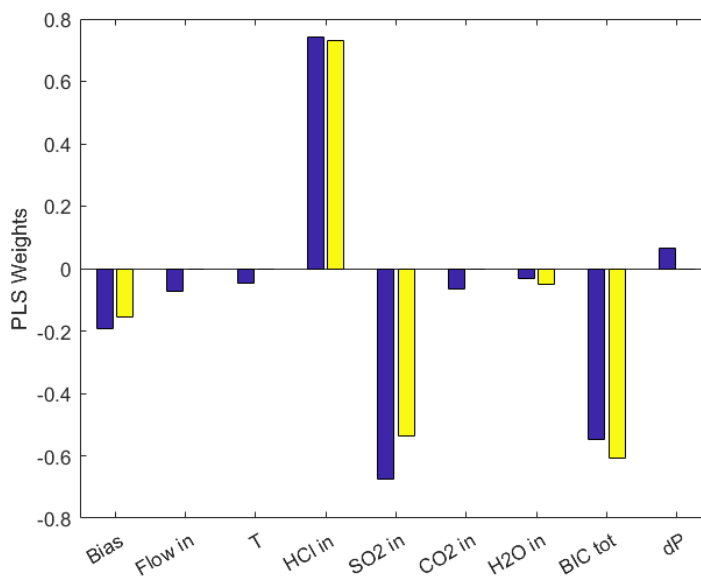
Figure 5.3: Variables weights comparison

The removal of the irrelevant variables produces, together with a net performance gain of the 7.7% on the V set, also a clear rearrangement of the remaining weights contributions. Among the confirmed predictors, in addition to the obvious influence of the inlet HCl and of the bicarbonate the negative contribution of the inlet sulfur dioxide is confirmed. The water shows a slight negative contribution which can be due to

an enhancing effect that the higher relative humidity gives to the reactivity of the solid sorbent. Among the remaining variables, the absence of any influence of the inlet gas flow is probably due to the proved low quality of this measure, which is known to be almost constant. The lack of any effect of the temperature is instead expected according to the fact that, for temperature higher than the one of the bicarbonate decomposition, the effect of this parameter on the reactivity is known to be negligible. In the following chart (Fig.5.4) the model predictions plotted against their real value demonstrate a good fitting performance, especially if compared with the results of the previous available models (see 4.2.1) and if related to the simplicity and computational cheapness of the linear model. The concentrations time series (Fig.A.28 and A.29) shows indeed an overall good agreement of the time trend even though, as it was expected by the previous probability distribution analysis, the trained model miss to predict correctly extremes measures. In particular, as it is confirmed by the clear deviations in the reported parity plot (Fig.5.4), this model tend to overestimate the lower conversions corresponding to lower bicarbonate feeds nearer to the stoichiometric needs ($rs = 1$). This is probably due to intrinsic inability of the linear model to predict the known non-linearity that the performance experiences related to bicarbonate dosage in this region. Unfortunately, depending on the higher reliability which should be required on the highest concentrations peaks, this misfitting problem leads to a substantial problem which would require the upgrade of the model features with a further analysis or even the use of a different modeling approach.
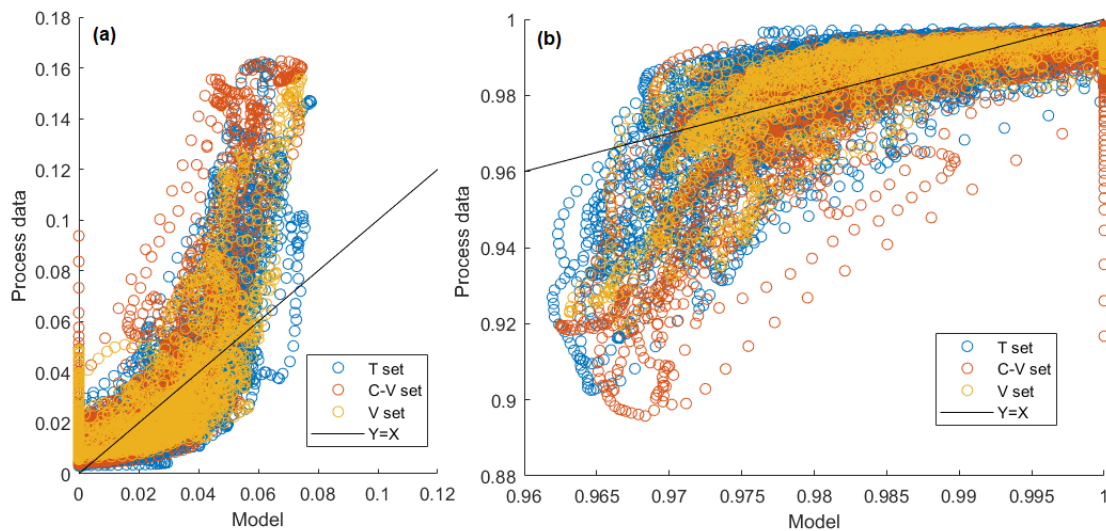
Figure 5.4: Model prediction vs process data: (a) Outlet HCl flow [kmol/h], (b) HCl conversion

5.2   SO2 model

The sulfur dioxide is characterized by ill conditioned measures sets, both at the inlet and even more at the outlet, which depending on their very low average concentrations and low variability, are highly squeezed toward zero. This, in addition to some other not deeper analyzed trends of the inlet SO2 in the CV set, does not allow to confirm the preliminary assumptions done on PC2 population. So, whereas the sulfur dioxide is present in negligible quantities compared to the hydrochloric acid, this study is carried on qualitatively only as a conclusion of the system state predictors analysis.
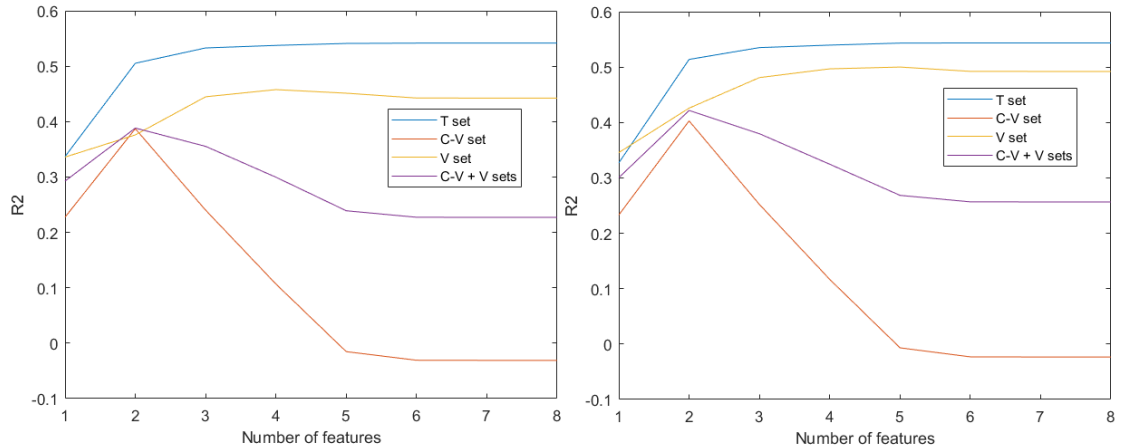
Figure 5.5: Fitting performance vs N°of features: (a) No range correction, (b) With range correction

As can be seen from the previous chart, differently from what used to happen with the hydrochloric acid, the V set gives an higher performance than the CV. The effective curves trends does not change indeed, but is clear instead that the range correction in the right part does not seem to have the same beneficial effect against the over-fitting which was spotted in the previous analysis. The different performances of the two validations sets in the two cases suggest firstly that firstly, given the spotted plateau also without the range correction, the type of population characterizing the V set does not undergo over-fitting and secondarily, that the type of population characterizing the CV set suffer an overestimation over-fitting which cannot be corrected even with the range correction. The quantitative analysis of these plots show a comparable performance with the HCl model but, given the preliminary stated purpose, no further fitting performance analysis will be performed. Whereas the performance on the CV set drops drastically after the two features, the overall performance is privileged and two is kept as optimal parameter of the model.
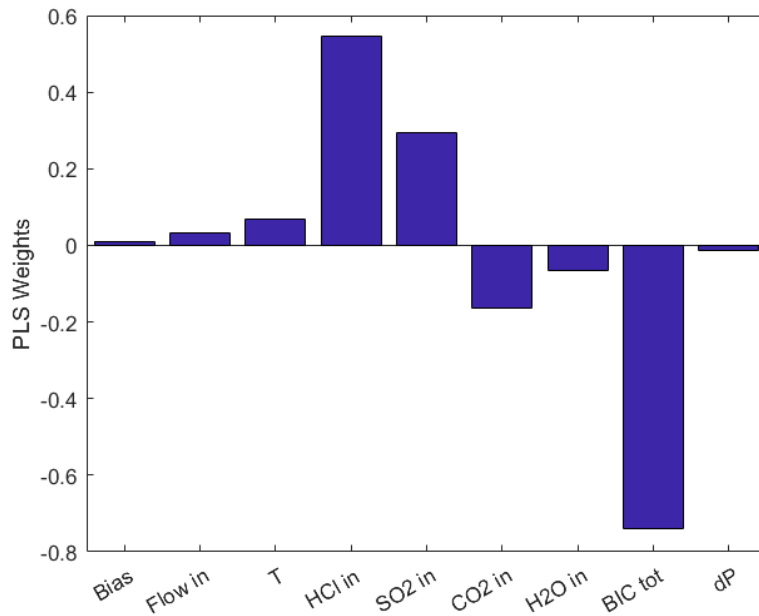
Figure 5.6: Variables weights for the five features PLS model

The low number of features suggests preliminarily that, among the starting variables, the number of true predictors is limited. Indeed, looking at the weights chart (Fig.5.6 the only substantial and dominating contributions are the ones of the two inlet acids and the bicarbonate. The contribution of the carbon dioxide is as well substantial, especially if compare dwith the one spotted for the HCl, and lets understand that the one which was spotted during the principal component analysis was indeed referred to the sulfur dioxide and not to the hydrochloric acid. The large inlet HCl contribution, actually bigger than the of inlet SO2, is easy explained as the natural counterpart of the negative contribution which the SO2 has on the outlet HCl concentration. Indeed, as already explained the HCl has an higher affinity for the sorbent and, if presents in larger quantities, will have then the priority in the bicarbonate consumption with a consecutive increase in the outlet SO2.

Chapter 6

Conclusions

Statistics offer a broad range of basic and advanced techniques which allow to study even large amount of data while reducing the complexity that characterize these systems. The analyzed dataset needs to meet some minimum standard requisites in order to be used and, whereas DCS data are usually characterized by large quantities of errors and missing measures, they need to be pretreated with the proper use of tailored algorithms. In some cases however, the possible presence of large holes of missing data which cannot be re-constructed reliably, obliges a fragmentation of the dataset which is not always congenial to the next analysis and, given that process data can be analyzed reliably only if they are temporarily contextualized, could even make the dataset useless. In addition to this, the most used analysis and modeling techniques assumes preliminarily specific conditions for the probability distribution of the dataset which, if not satisfied, lead to poorer results which are more difficult to be interpreted. This makes highly suggested, in order to allow a reliable and easier study, to act directly on the process operation by carrying out test runs that ensure the most suitable dataset possible. Differently, given the objective difficulties which the implementation of a test run can entail and the unquestionable convenience (or unavoidable necessity) which the use of an already available dataset implies, as in the studied case, the dataset will naturally require further elaborations which, however, do not ensure the best final results.

From this point of view, despite the highlighting of some marginal issues, the available process data have been demonstrated highly suitable for this study. The principal component analysis of the dataset allows in fact to characterize several known and unknown phenomena, occurring before and during the acid gases abatement. Specifically, it has been possible to condense the DTS ruling phenomena in four main variables which, in addition to the direct study of the system, can be also implemented directly for on-line monitoring and control applications which allows a clear view of specific process trends and anomalies derived or implied in productivity or safety issues. The chemical and process expertise plays an essential role for the correct interpretation of these results and has been proved fundamental for the conception of statistical custom applications, such as the characterization of the filter reactivity potential, which allows to obtain peculiar and objective results which are essential for a reliable analysis.

Nevertheless, the algebraic procedure behind PCA shows a strong dependence on the specific assembling of the dataset which in some cases, even with a strong expertise basis, could not allow to obtain absolute results. In these cases the next modeling phase need to assume then, in addition to its intrinsic predictive purpose, also a fundamental role in the final exploration and validations of the relations found in the dataset.

The partial least square regression is highly suitable to fulfill this necessity because it allows to associate its good performance to an excellent interpretability that has proved essential for the definitive assessment of what are the variables which truly affect the system and how much strong is this influence. Moreover, the final designed model represents a clear step forward if compared to the previously available empirical model because it is capable to reproduce future instantaneous trends with an appreciable fitting grade. Unfortunately, the inability of the model to reproduce specific non-linear behaviors, does not allow the final performance to be considered sufficient for a reliable process implementation and this suggests, as a possible future development, to carry out a further analysis in order to upgrade the model or suggest the use of different modeling approaches.

Appendix A

Supporting material

A.1 Case study
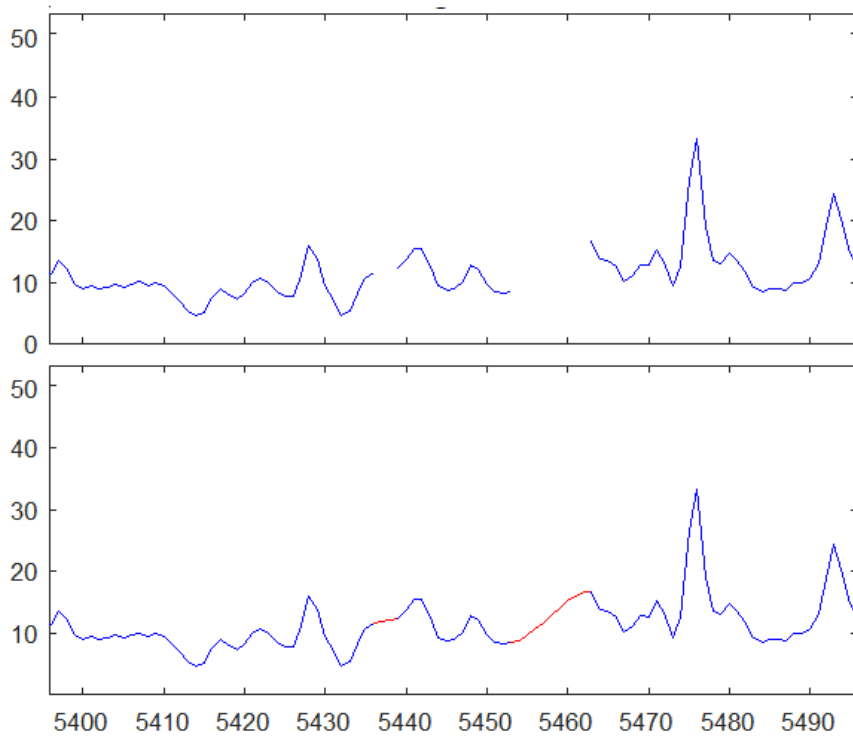
Figure A.1: Reconstruction of a 2 minutes and a 10 minutes windows

	Number of measures	% of the total
Total dataset	488400	/
Missing measures (NaN)	19885	4.1
After dataset selection:		% of the selection
Total dataset	214710	/
Missing measures (NaN)	2530	1.2
Outliers	801	0.4
Total reconstructed data	3331	1.6

Table A.1: Quality of the data and cleaning operation recap
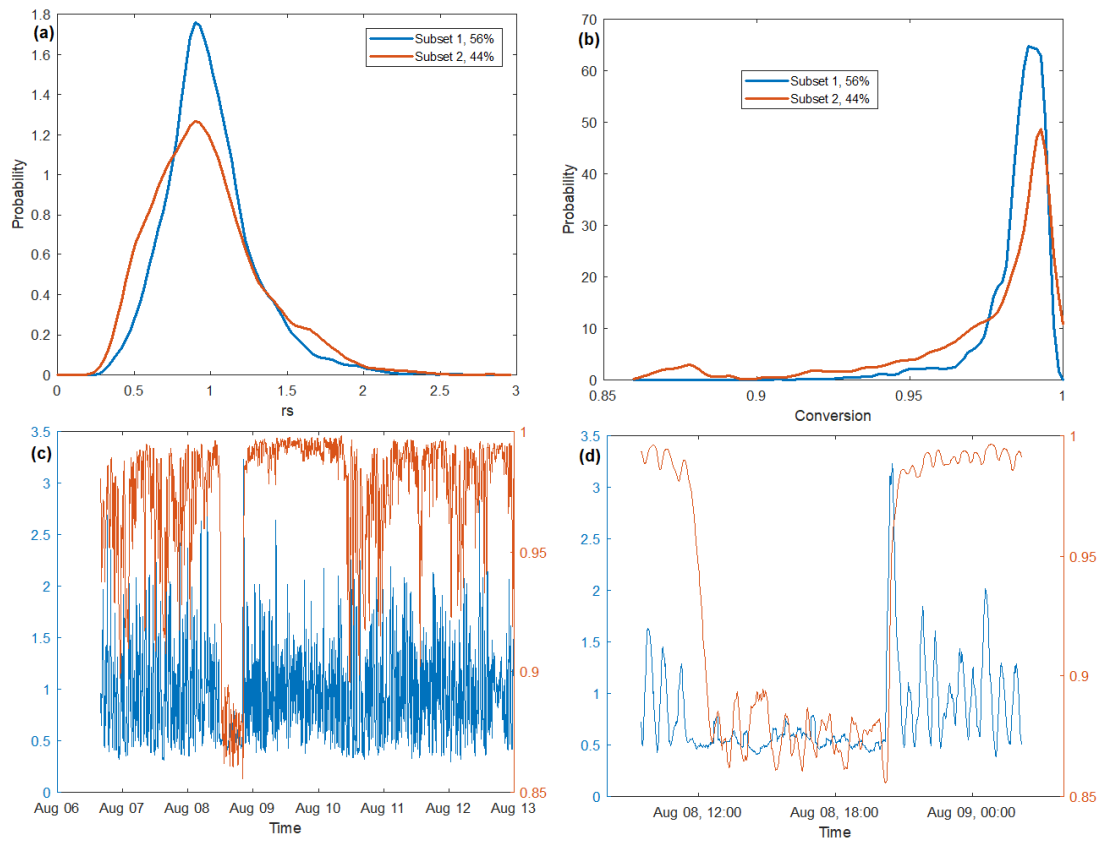
Figure A.2: Probability distributions: (a) *rs* and (b) conversions χ. Time series, *rs* (blue) and conversion χ (orange) : (c) second set and (d) zoom
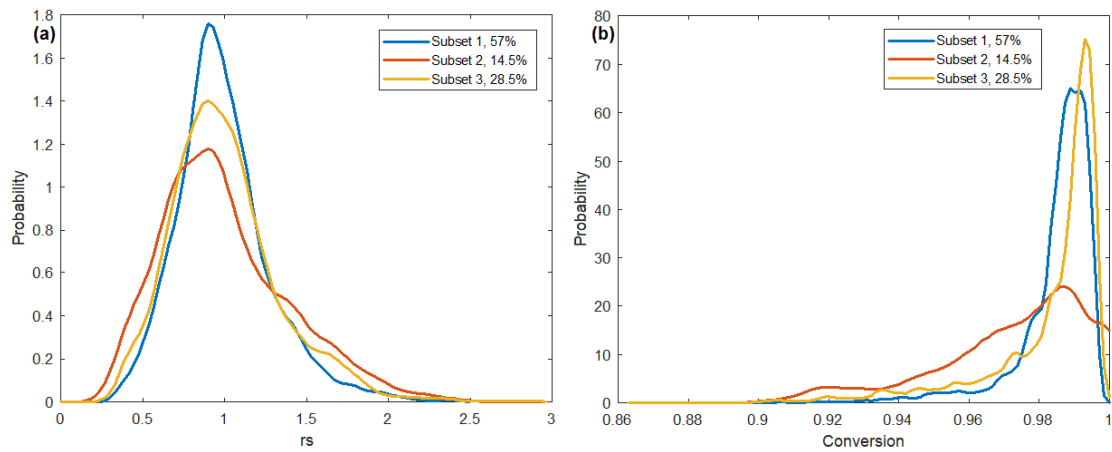
Figure A.3: Probability distributions: (a) *rs* (b) χ

A.2 Data analysis

N°PC	EigenVal	Variance (%)	Total Variance (%)
1	2.33	23.3	23.3
2	1.7	17.0	40.3
3	1.47	14.7	55.0
4	1.27	12.7	67.7
5	0.9	9.0	76.7
6	0.68	6.8	83.5
7	0.64	6.4	89.9
8	0.41	4.1	94.1
9	0.39	3.9	98.0
10	0.2	2.0	100

Table A.2: PCA results
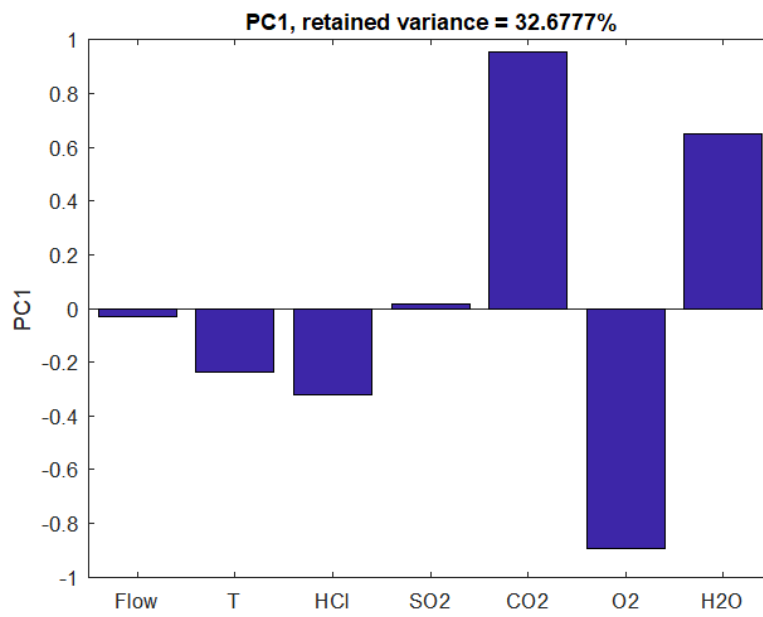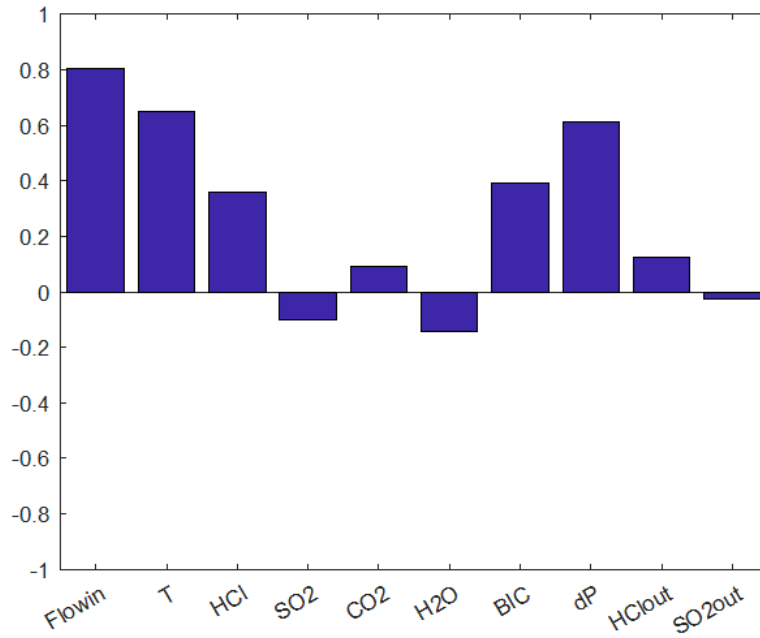
Figure A.5: Inlet stream: PC1 loading contribution plot
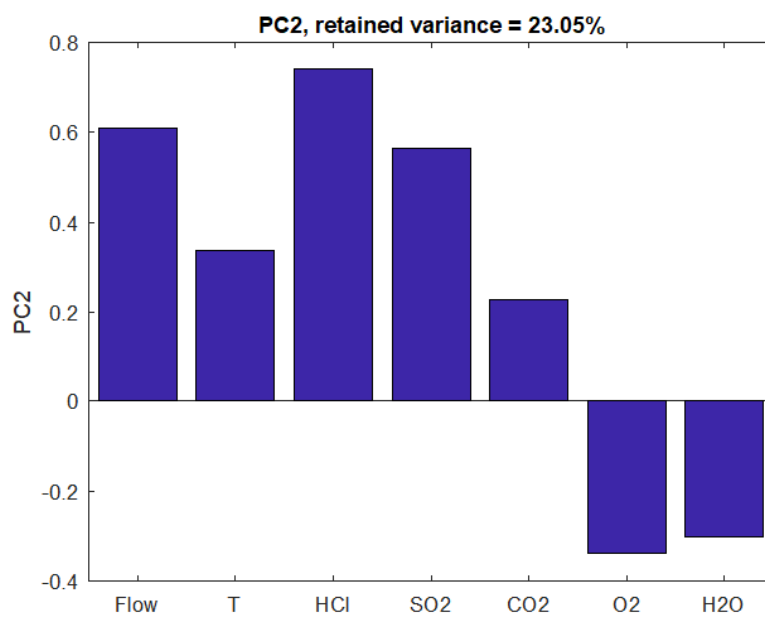
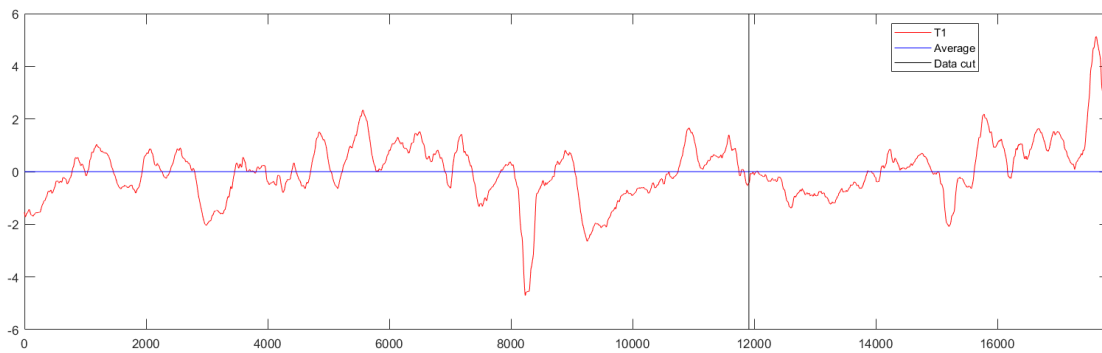Figure A.6: Inlet stream: PC2 loading contribution plot
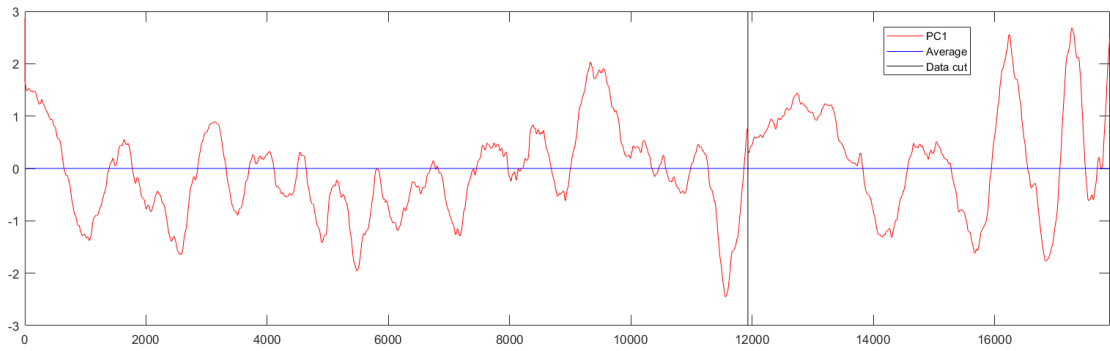
Figure A.7: PC1 time series representation, moving average at 180 minutes
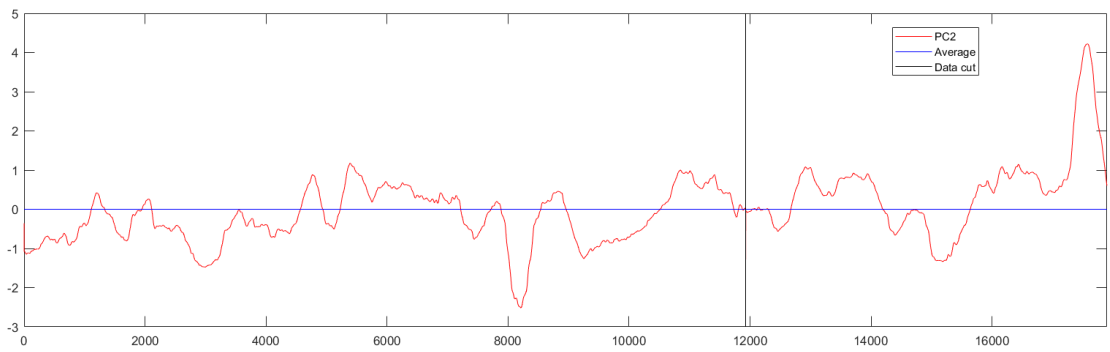
Figure A.8: PC2 time series representation, moving average at 180 minutes
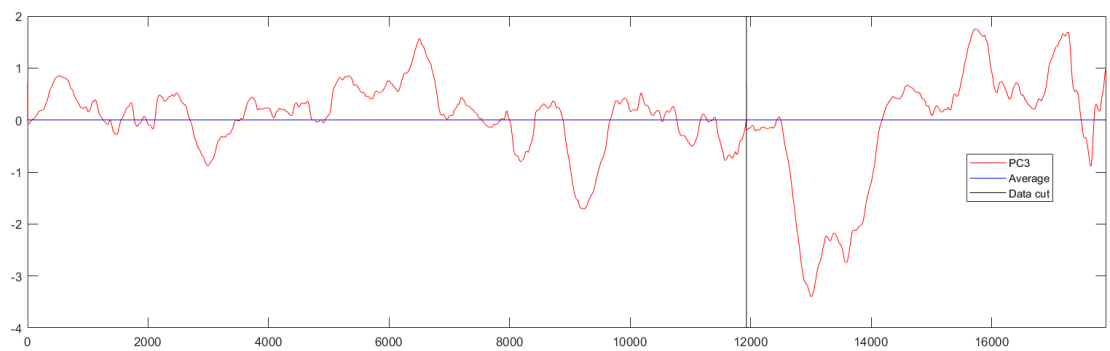
Figure A.9: PC3 time series representation, moving average at 180 minutes
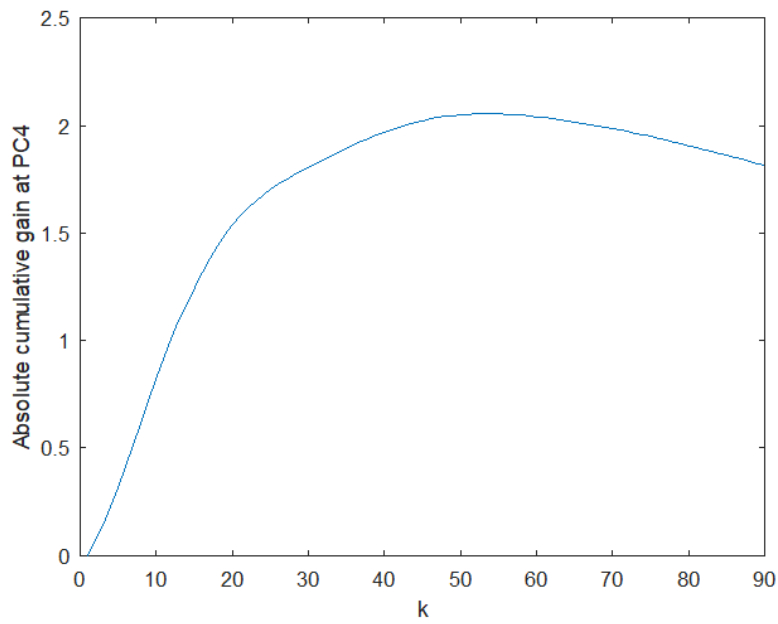
Figure A.10: Overall retained information at PC4, Total variation
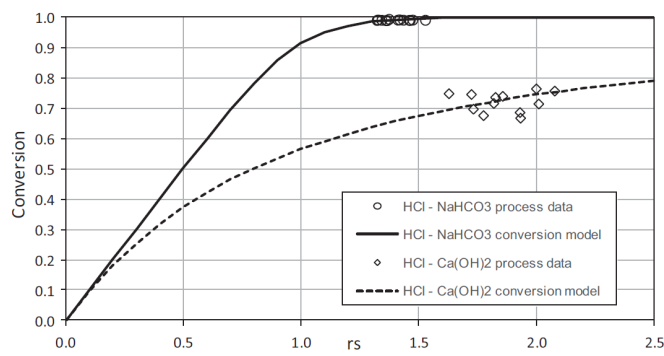
Figure A.11: Model calibration vs. process data (1 day average) [6, Antonioni, 2014]
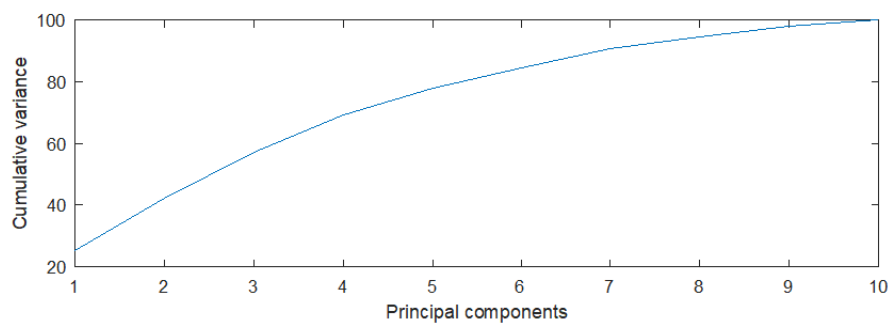
Figure A.12: Total retained variance
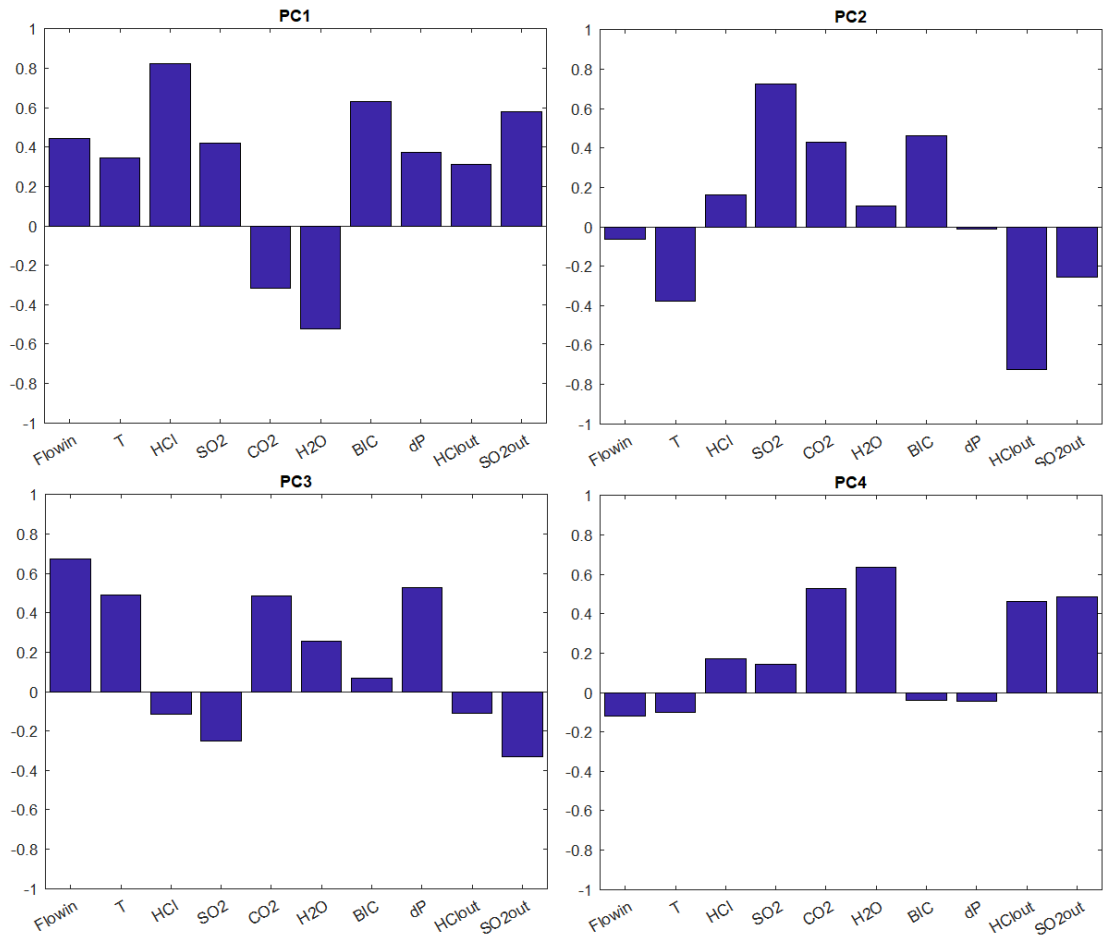
Figure A.13: Loading contribution plots for the first four PCs
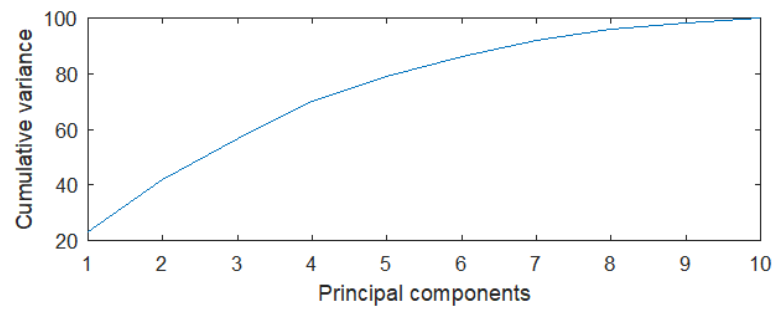
Figure A.14: Total retained variance
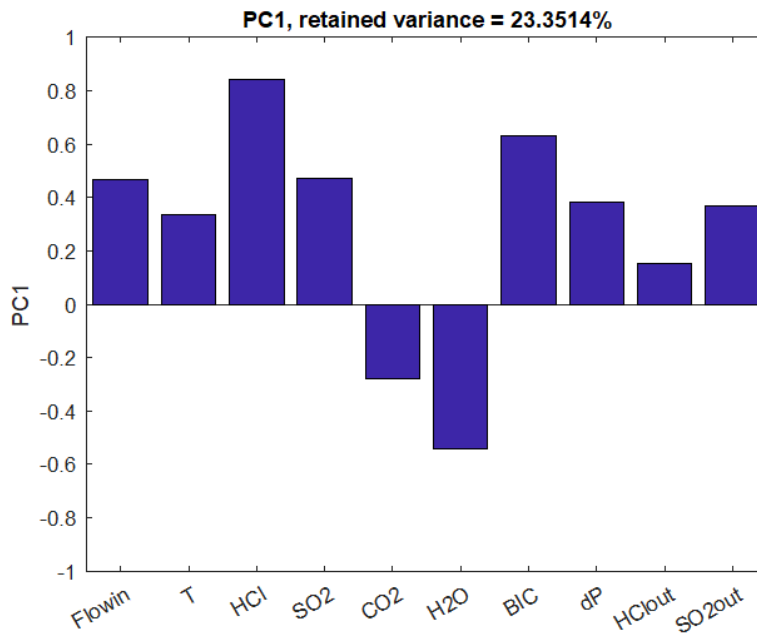
Figure A.15: PC1 mono variate contribution plots
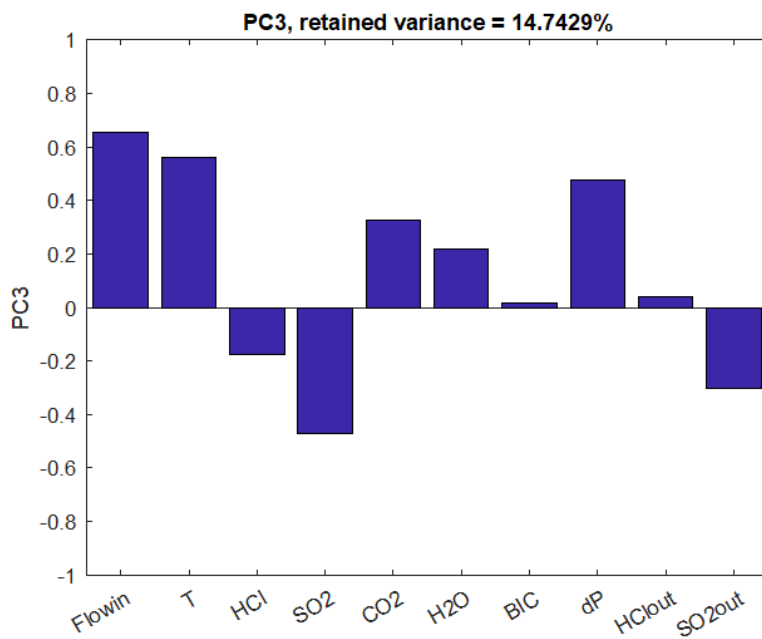
Figure A.16: PC3 mono variate contribution plots
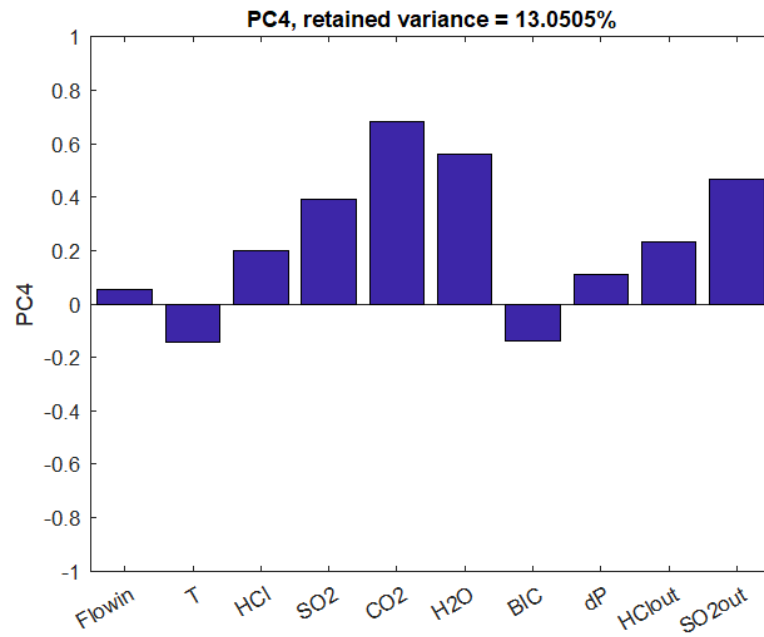
Figure A.17: PC4 mono variate contribution plots
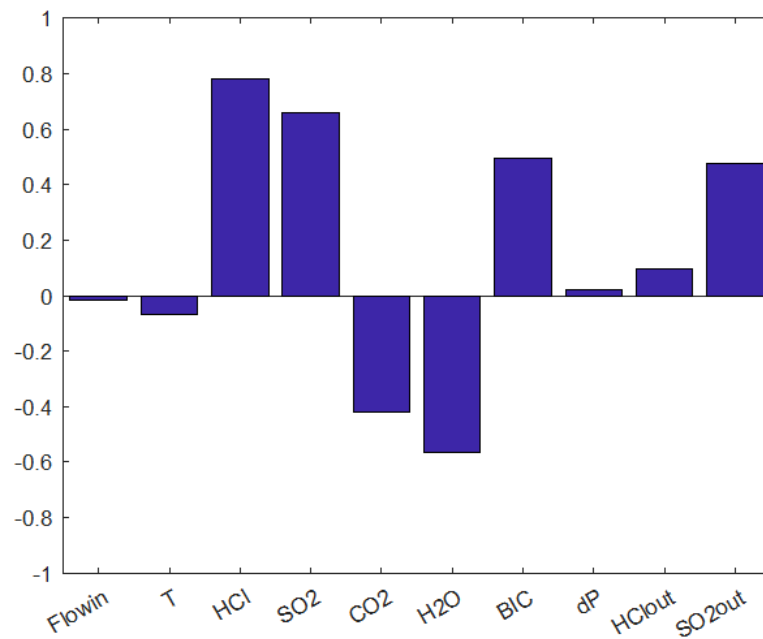
Figure A.18: First trajectory mono variate loading contribution plot

Figure A.19: Second trajectory mono variate loading contribution plot

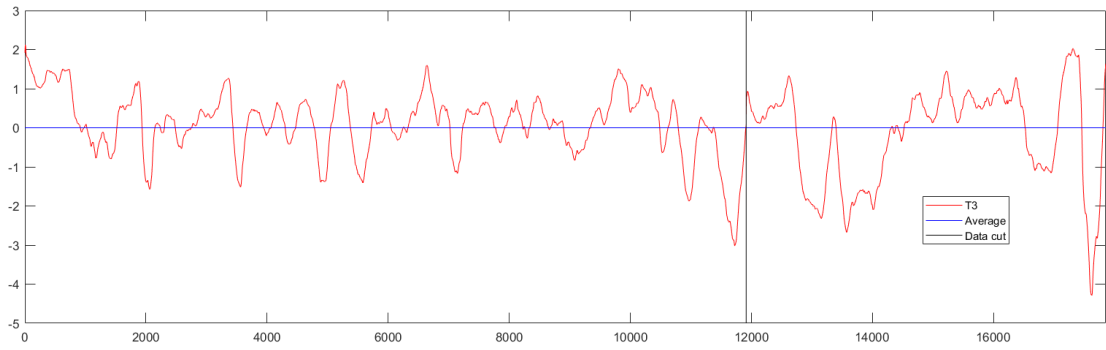Figure A.20: T1 time series representation, moving average at 180 minutes

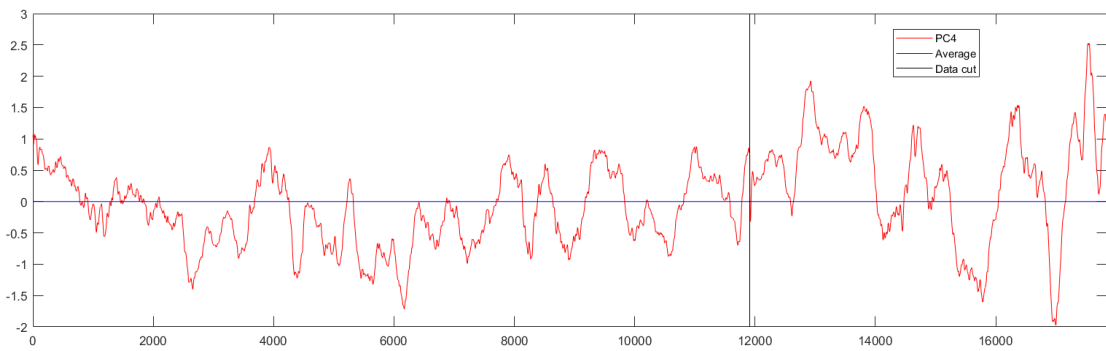Figure A.21: T3 time series representation, moving average at 180 minutes

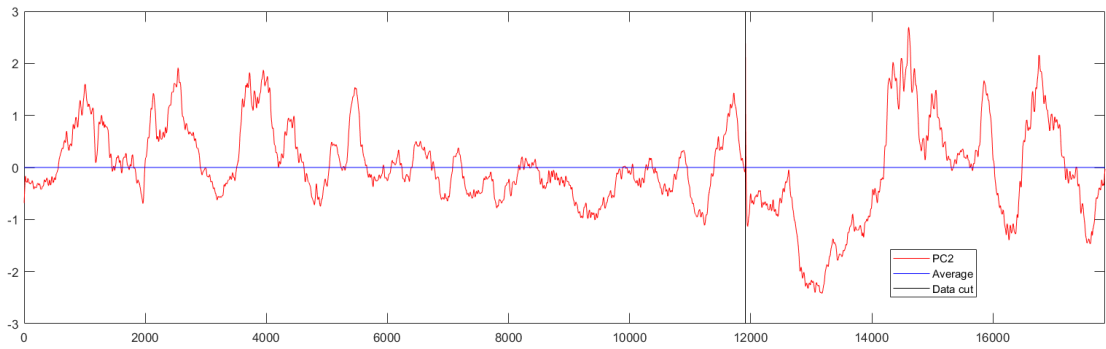Figure A.22: PC4 time series representation, moving average at 180 minutes

Figure A.23: PC2 time series representation, moving average at 180 minutes
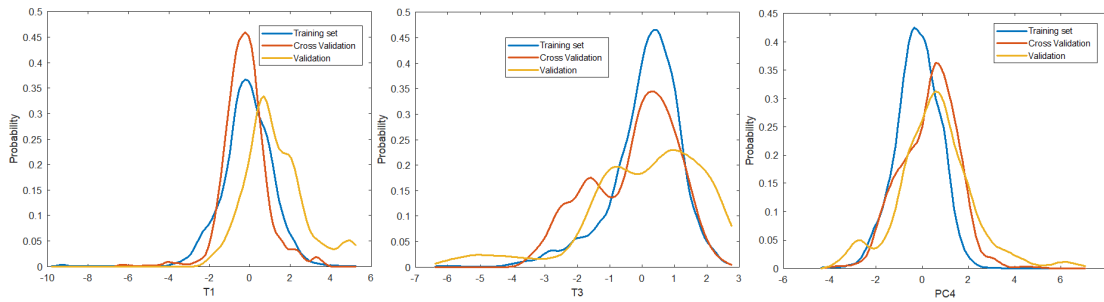
A.3 Modeling

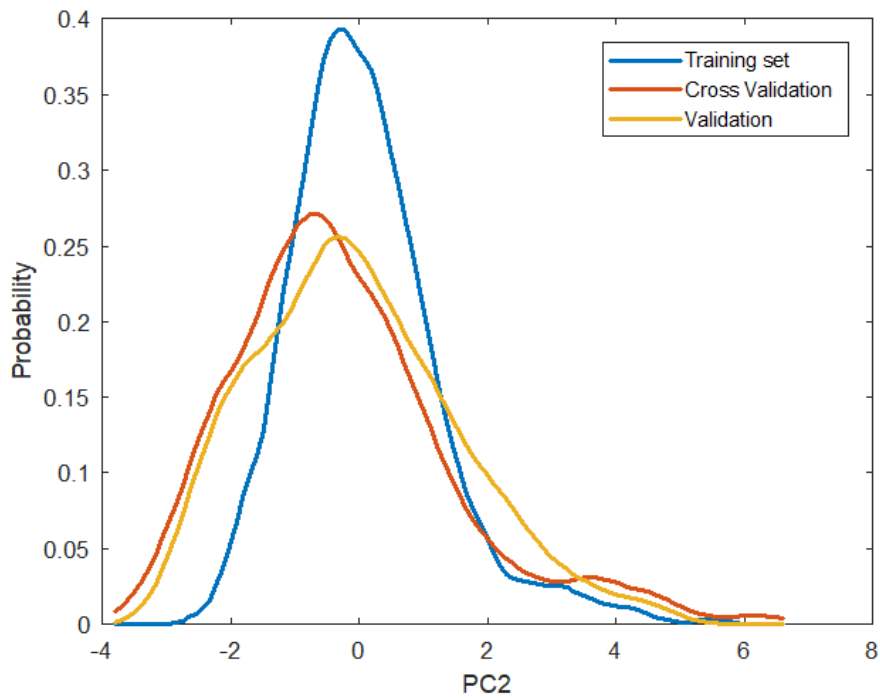Figure A.24: T1 T3 and PC4 probability distributions

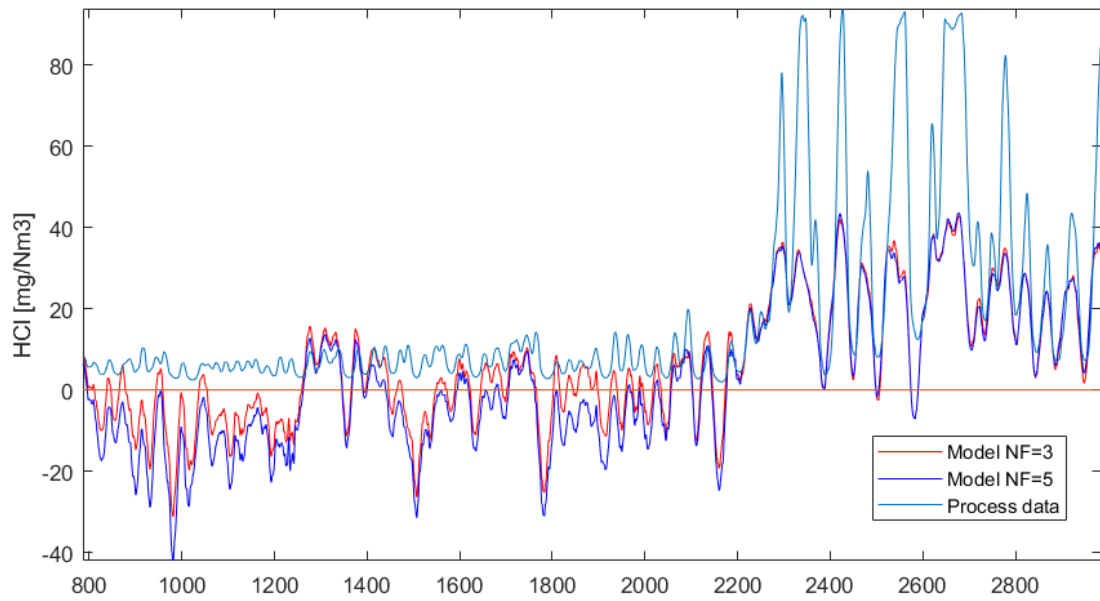Figure A.25: PC2 probability distributions

Figure A.26: Fitting performance comparison on the CV set: three features vs five features
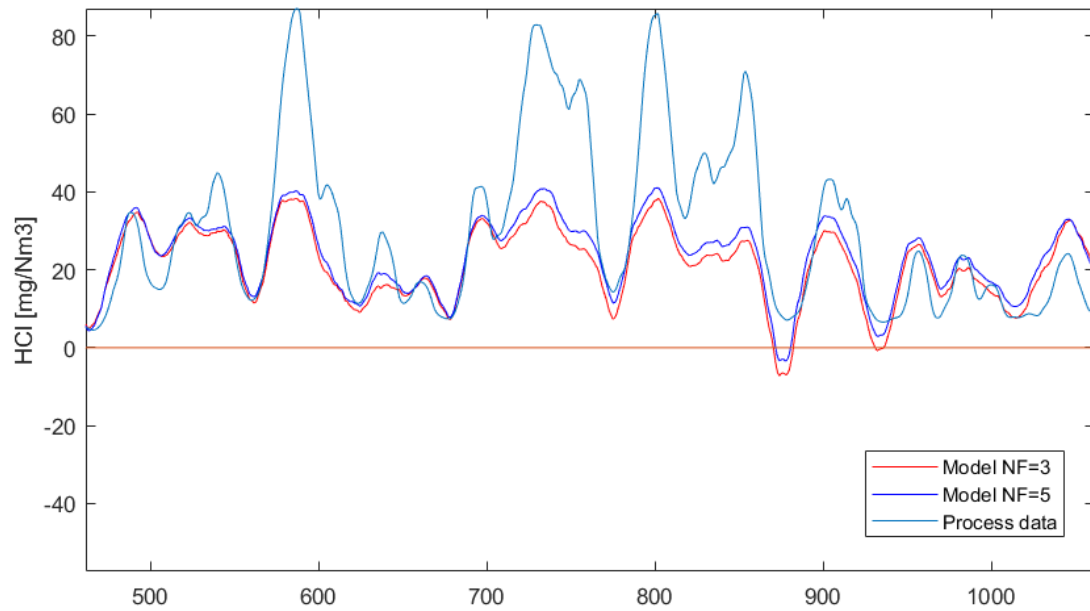

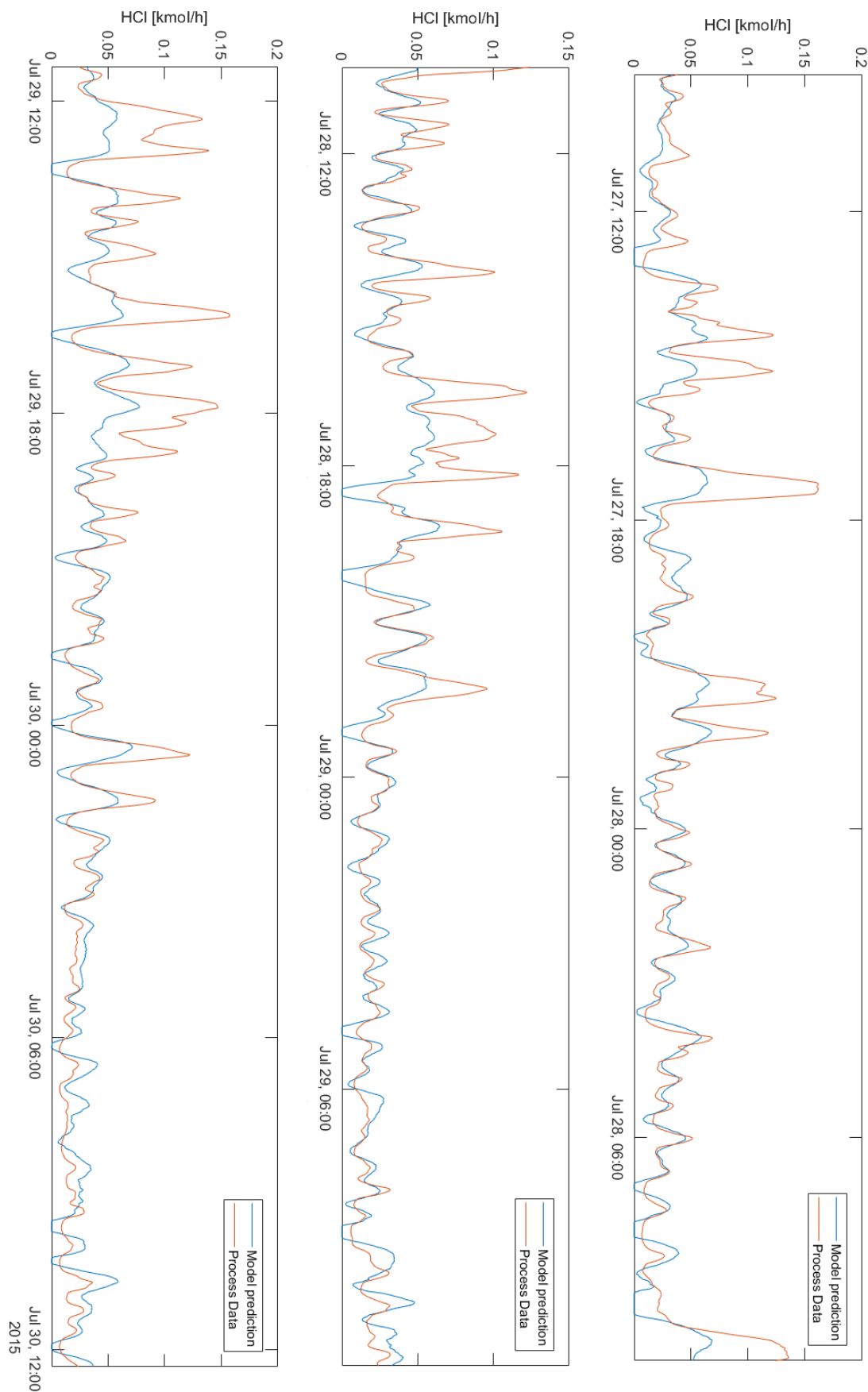Figure A.27: Features tuning: Fitting performance comparison on the V set

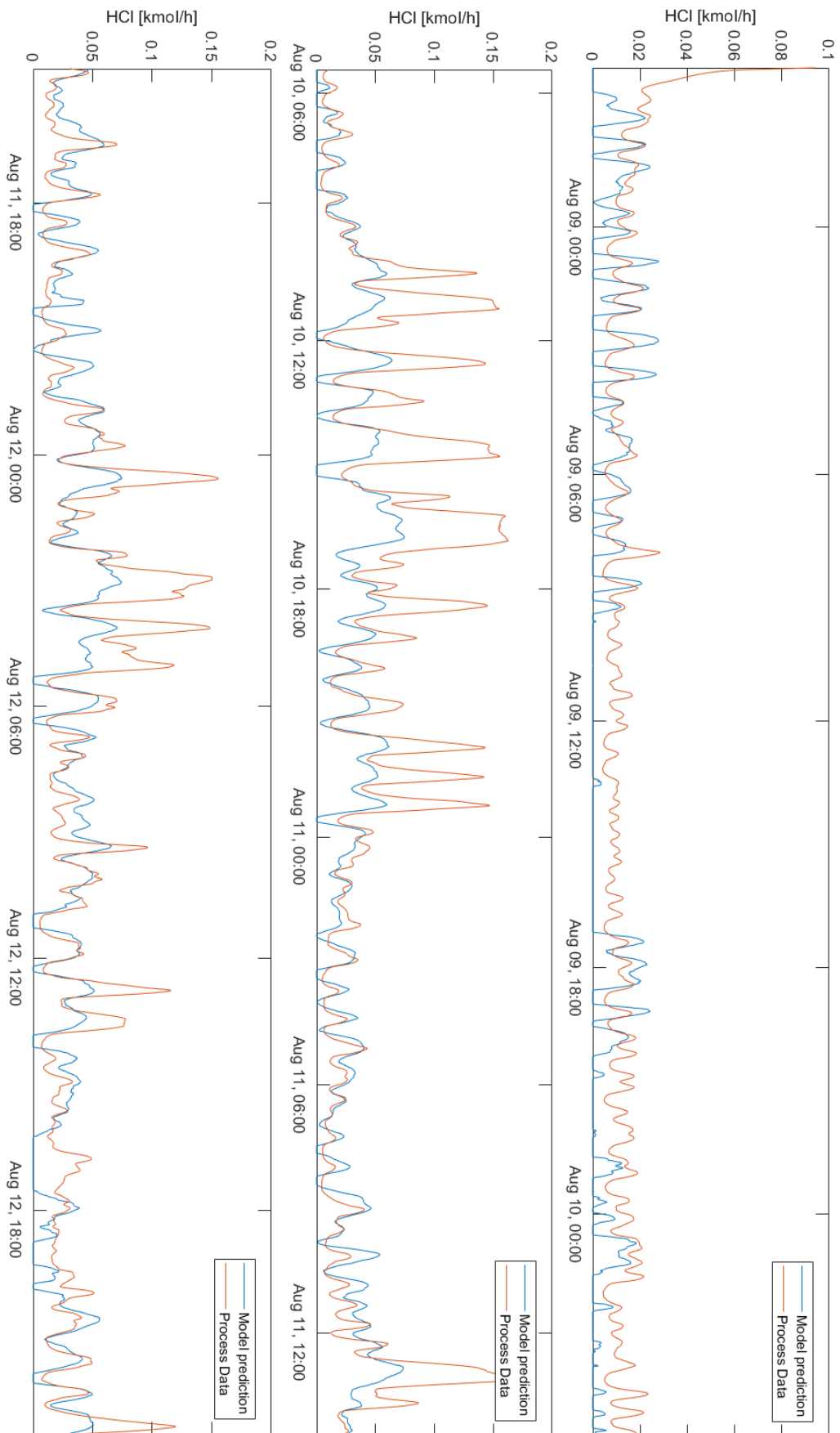Figure A.28: Fitting performance of the final model on the T set: sample from Jul 27 to Jul 30

Figure A.29: Fitting performance of the final model on the CV and V set

Appendix B

Algorithms

Notice that all the used algorithms have been created with Matlab. In addition to the ones listed, all the used and designed functions comes directly from the Matlab library or by the combination of the Matlab algorithms. All the functions designed for visual representation and exploration are not reported.

B.1   Data cleaning and filtering

B.1.1   Outlier detection

```matlab
function X=NO_outliers(X)
[m,n]=size(X);
%    ( 1,2,3      ,4         ,5         ,6,7      ,8         ,9         ,10
    ,11,12,13,14)
alpha=[0,0,10^-5,10^-20,10^-20,0,10^-5,10^-120,10^-50,10^-9,0,0,10^-13,0,0];
    %Tolerances
for i=1:n
    if alpha(i)~=0
        p=pdf(fitdist(X(:,i),'Rayleigh'),X(:,i));
        out=find(p<alpha(i));
        for j=1:m
            if X(j,i)==0
                X(j,i)=NaN;
            end
        end
        for j = out
            X(j,i)=NaN;
        end
    end
end
end
```

B.1.2   Nan removal

```
function X=NO_NaN(X)
% This function close the holes of missing measures in the dataset
[m,n]=size(X);
for j=1:n
    for i=1:m-1
        if isnan(X(i,j))
            k=i+1;
            while isnan(X(k,j))&&k<m
                k=k+1;
            end
            set=i:k;
            x=[i-4,i-3,i-2,i-1,k,k+1,k+2,k+3];
            x=x(x>0&x<m);
            v=X(x,j);
            X(set,j)=interp1(x,v,set,'pchip');
            i=k;
        end
    end
end
end
```

B.2   Empirical model fitting

B.2.1   Empirical model

```
function chi = Antonioni(rs,n)
% This function is an empirical model able to give back the conversion of
% the reactor. Input parameters are two:
%
chi=(rs.^n-rs)./(rs.^n-1);
end
```

B.2.2   Cost function

```
function [J,der] = Cs(rs,n,y)
m=length(rs);
J=(1/(2*m))*sum((Antonioni(rs,n)-y).^2);
der=(1/m).*transpose((rs.^n.*log(rs))./(rs.^n - ones(m,1)) +
    (rs.^n.*log(rs).*(rs - rs.^n))./(rs.^n -
    ones(m,1)).^2*(Antonioni(rs,n)-y);
end
```

B.2.3   Statistical fitting

```matlab
function [n,cost] = Train_emp(rs,chi)
% Create "short hand" for the cost function to be minimized
costFunction = @(t) Cs(rs,t,chi);

% Now, costFunction is a function that takes in only one argument
% Minimize using fmincg
[n,cost] = fminunc(costFunction,1);%, options);

end
```

Bibliography

[1] J. Vehlow,
Air pollution control systems in WtE units: An overview,
Waste Management, Volume 37, 2015, Pages 58-74, ISSN 0956-053X,
Available at: `https://doi.org/10.1016/j.wasman.2014.05.025`
(last accessed: 22/07/2018).

[2] Alessandro Dal Pozzo Et. Al.,
Trattamento a secco dei fumi di combustione prodotti nella termovalorizzazione di RSU, Tecnica, La termotecnica, Termovalorizzatori & Ambiente, Aprile 2016, Pages 53-57.

[3] Alessandro Dal Pozzo Et. Al.,
Comparison of alternative flue gas dry treatment technologies in waste-to-energy processes, Waste Management, Volume 51, 2016, Pages 81-90, ISSN 0956-053X,
Available at: `https://doi.org/10.1016/j.wasman.2016.02.029` (last accessed: 22/07/2018).

[4] *Directive 2010/75/EU of the European Parliament and of the Council of 24 November 2010 on Industrial Emissions (Integrated Pollution Prevention and Control),* OJ L 334/17, Available at: `http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2010:334:0017:0119:en:PDF` (last accessed: 22/07/2018).

[5] Antonioni, G., et al., 2011.
Simulation of a two-stage dry process for the removal of acid gases in a MSWI, Chem. Eng. Trans. 24, 1063–1068,
Available at: `http://dx.doi.org/10.3303/CET1124178` (last accessed: 22/07/2018).

[6] Antonioni, G., et al.,
Modelling and simulation of an existing MSWI flue gas two-stage dry treatment, Process Safety and Environmental Protection (2013),
Available at: `http://dx.doi.org/10.1016/j.psep.2013.02.005` (last accessed: 22/07/2018).

[7] Brivio S., 2007,
Depurazione dei fumi e valorizzazione dei prodotti,
Power Technology 4, 42-44

[8] Barry M. Wise et al,
The process chemometrics approach to process monitoring and fault detection, Z Proc. Cont Vol. 6, No. 6, pp. 329 348, 1996, Eigenvector Research, 830 Wapato Lake Road, Manson, WA 98831, USA.

[9] Rossi F., Manenti F., CMIC "Giulio Natta"Dept.,
Real-time Big Data Reconciliation (BDR), Big Data nell'Ingegneria Chimica: metodi per l'analisi, la convalida e la modellazione,

[10] Guglielmi G., Dal Pozzo A., Antonioni G., Cozzani V., Montanari A.,
Analisi di soluzioni innovative per la riduzione dei costi operativi nel trattamento fumi in sistemi a secco multistadio operanti in impianti di termovalorizzazione di rifiuti: relazione finale, Convenzione di ricerca tra TM.E. S.p.A. – Termomeccanica Ecologia e Dipartimento di Ingegneria Civile, Chimica, Ambientale e dei Materiali Alma Mater Studiorum Università di Bologna

[11] O. Levenspiel,
Chemical Reaction Engineering, 3rd ed. John Wiley & Sons (USA), 1998

[12] Rick Johnston,
Approaches to Debottlenecking and Process Optimization, BioProcess International, May 1, 2012.

[13] F.D. Bähner et. al.,
Analysis and Modelling of an Industrial Pressure Filtration using Process Data, IFAC PapersOnLine 50-1 (2017) 12137–12142

[14] A. Klimkiewicz et. al.,
Multilevel Modeling for Data Mining of Downstream Bio-Industrial Processes, Chemometrics and Intelligent Laboratory Systems 154 (2016) 62–71

[15] M. Khayet et. al.,
Artificial neural network model for desalination by sweeping gas membrane distillation, Desalination 308 (2013) 102–110

[16] John W. Tukey,
The future of data analysis, 1961

[17] James, Gareth (2013).
An Introduction to Statistical Learning: with Applications in R, Springer. p. 176. ISBN 978-1461471370.

[18] Herve Abdi, Lynne J. Williams,
Principal component analysis, WIREs Computational Statistics, Volume 2, July/August 2010 2010 John Wiley & Sons, Inc.