École polytechnique de Louvain

# Calibration strategy of a 1D model for supersonic ejectors

Author: **Lucas Delbecque**
Supervisors: **Yann Bartosiewicz, Miguel Alfonso Mendez**
Advisors:
**Jan Van den Berghe, Antoine Metsue**
Readers: **Philippe Chatelain, Grégoire Winckelmans**
Academic year 2021–2022
Master [120] in Mechanical Engineering

# Acknowledgements

This thesis marks the end of a year in learning which was made possible thanks to many people.

Firstly, I would to express my gratitude to my supervisors: Prof. Yann Bartosiewicz for his support and his availability for the last reviews and Prof. Miguel Alfonso Mendez for his highly valuable advice. Prof. Mendez was involved with an interest all along this work. His good ideas and motivation pushed me to give the best of myself.

I would also like to particularly thank Jan Van den Berghe for his time, his advice and his devotion all along the year. This work would not have been the same without him.

Moreover, I have to thank Antoine Metsue who gave me some helps with explanation when I needed it, all the team of the VKI for their kind welcome.

Finally, I would like to thank all the people who gave me more than academic support. At first my parents Patrick Delbecque and Virginie Van Caeneghem which gave me support all this year and especially my mother for her numerous review even when she was on vacation. I would like also to thank my closest friends who have always been around me; especially Ben Delcoigne, Louis Delait, Antoine Callant, Olivier Graux and finaly Nell Daeleman for her big support during the harder moments of the last steps of this work.

# Contents

# List of Symbols

**Abbreviations**

$CFD$ Computational Fluid Dynamics

$MAE$ Mean Absolute Error

$ML$ Machine Learning

$MSE$ Mean Squared Error

$VKI$ Von Karman Institute

**Greek letters**

$\beta$ Compound flow indicator

$\omega$ Equivalence ratio

$\pi$ Pressure ratio

**Roman letters**

$A$ Cross-sectional area

$\dot{m}$ Mass flow rate

$E$ Error criterion

$J$ Cost function

$M$ Mach number

$p$ Pressure

$T$ Temperature

$V$ Velocity

## Sub- and Superscripts

| | |
|---|---|
| $*$ | Choked condition |
| $\cdot$ | Time derivative |
| $0$ | Stagnation conditions |
| $c$ | Back pressure |
| $c$ | Critical conditions |
| $d$ | Diffuser outlet |
| $is$ | Isentropic |
| $m$ | Mixed flow |
| $out$ | Outlet condition |
| $P$ | Prediction |
| $p$ | Primary flow |
| $s$ | Secondary flow |
| $t$ | Throat section |
| $y$ | Hypothetical throat |

# Chapter 1

# Introduction

## 1.1 Context

An ejector is a device powered by a high pressure fluid, called the primary -or motive -fluid, which entrains another fluid at a lower pressure, called the secondary - or suction - fluid. At the exit, a fluid at a pressure between the high one and the low one comes out.
Ejectors can play the role of a gas and vapor pump to create a vacuum pump but also the role of a mixing pump [2]. They can also play the role of a compressor.
What differentiates an ejector from a pump, a compressor or an aspirator, is the absence of moving part in the way to proceed. An ejector therefore needs low maintenance and operate in a single way.
All of this makes ejectors cheap devices for multiple aspects. Moreover, they are able to play with enormous volumes of fluid compared to their size [2]. This makes the ejectors an attractive alternative compared to what is traditionally used for many areas.
Among them, ejectors can be particularly efficient for the cooling and heating systems when they are included to refrigeration by steam jet cycle, using halocarbon compounds or by solar-powered cycle [3]. Since ejectors are small and light compared to their performances [3], they can also be used in the aerospace and aircraft industry for the thrust augmentation and the jet engine noise reduction [4]. Ejectors are also appealing to make efficient vacuum pumps by training gases and vapor out of a system [2]. Moreover it can be mentioned that they are usefull devices for ecological purposes like desalination applications [5], for smoke extraction and in the recovering of wasted energy in thermal powerplants [6] or for the improvement of the overall efficiency of a fuel cell [6]. All of the above is far from being comprehensive.

## 1.2 Ejector fundamentals

An ejector is composed of four parts. Figure 1.1 represents a classical schematic of an ejector

- Primary nozzle
  This is the place where the primary flow is accelerated until its pressure becomes sufficiently low to entrain the secondary flow. For the following, as this study concerns supersonic ejectors, the primary flow will always be considered supersonic at the exit of the primary nozzle. It is therefore considered that it is choked in the near vicinity of the throat. The nozzle may as well be convergent as convergent-divergent [6].

- Suction chamber
  The suction chamber is considered to be a constant pressure place where the secondary flow is trained [7]. The primary flow being accelerated in the primary nozzle, its static pressure becomes lower than that of the secondary fluid at rest. This produces a movement of the secondary fluid from outside the ejector into the suction chamber. After the chamber, the fluid is accelerated in what is called the secondary nozzle formed by all the ejector. Two modes can be distinguished depending on whether the secondary flow is choked or not. These two modes are called respectively on and off-design

- Mixing chamber
  This area begins just after the primary nozzle with a convergent section. In most cases, a constant section follows after. It is the part of the ejector where two entering flows mix to end up with a single flow. There is therefore a mechanical energy transfer from the primary to the secondary flow. During the mixing phenomena, complex mechanisms occur like shock waves or turbulent mixing [8].

- Diffuser
  It is the last part of the ejector which is a divergent section. If the ejector is well designed, The diffuser decelerates the flow in order to make the fluid recover its static pressure. The pressure at the end of the diffuser is called the back pressure. It is the pressure that is achieved by the flow after the mixing of the high stagnation pressure of the primary flow and the low pressure of the secondary stream.

During this study, it will be considered that the primary nozzle ends-up in the converging section of the ejector. This kind of geometry is called CPM (Constant pressure mixing). The nozzle exit position is an important parameter for the performances of an ejector. Although an other geometry could add some interesting

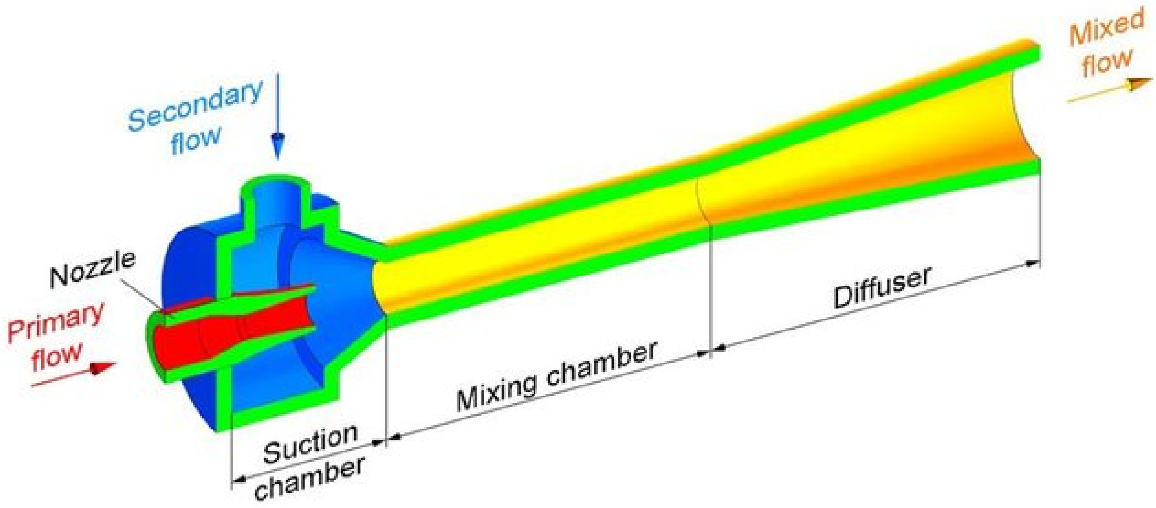


Figure 1.1: Schematic diagram of the supersonic ejector [7].

benefit which is to increase the mass flow rate, the CPM configuration is generally preferred for working with. Indeed, it makes the operation easier by allowing a wider range of back-pressure while remaining stable [6].

An ejector has often to entrain a maximum of secondary flow with a minimum of primary flow. Therefore, the main operating quantity which represents the performance of an ejector is the ratio between the primary and secondary mass flow rate. It is called the entrainment ratio.

$$\omega = \frac{\dot{m}_s}{\dot{m}_p} \tag{1.1}$$

Besides the entrainment ratio, the back pressure may also be a substantial quantity when the purpose is to increase the secondary flow pressure. When using dimensionless quantities, it is more appropriate to speak of compression ratio $\tau$ as performance indicator [6].

$$\tau = \frac{p_b}{p_s} \tag{1.2}$$

From then, a convenient way to represent the characteristics of an ejector is with the curve representing the evolution of the entrainment ratio as the function of the back pressure. It is called the operating curve. Figure 1.2 presents what is expected to be seen in an ideal case for an operating curve.

In the on-design regime, the secondary mass flow is chocked. The primary being chocked as well, the two mass flow rates are both independent of the back pressure (and therefore $\omega$ too) until the value for which the secondary flow is no longer chocked and become subsonic. This pressure is noted $p_b^*$ and is called the critical pressure. Both modes can also be called critical and sub-critical modes instead of on- and off-design. Once the pressure is above its critical value, the secondary flow is subsonic and is therefore no longer independent of the back pressure. As the back pressure increases, the secondary flow and the entrainment ratio decrease. Above a threshold value called the breakdown pressure, the back pressure is so high that the secondary flow is no longer entrained and the primary flow reverses back into the suction chamber. Above the breakdown pressure, the ejector is said to be in malfunction. This case will not be explored in this work.

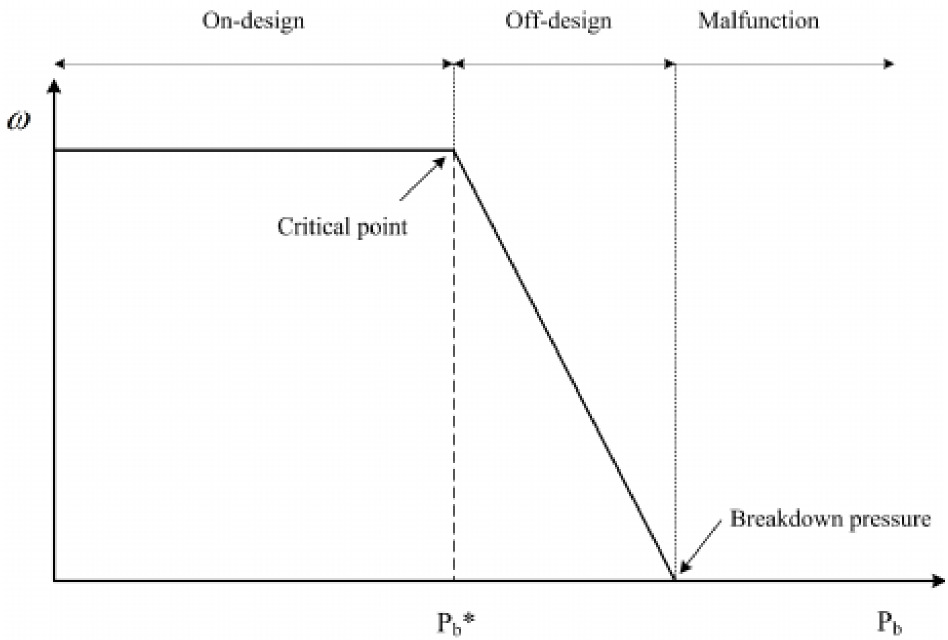


Figure 1.2: Ideal operating curve [9]

## 1.3 Ejector modeling methods

The main purpose of the modeling method for an ejector is the great reduction of the necessary means for the design of an ejector. It is a low cost alternative to experiments. Thanks to the simulation models, ejectors operations can be predicted before their design and this helps to optimize the performances. On the contrary, design without model often results in off-design operations [6].

There are two main ways to model an ejector.
One way is to use thermodynamic and analytical models. Due to the complexity of modeling the flow in an ejector, these models are 1 or 0D which often consider the ejector at steady state. They solve the model with the integral form of the conservation equations. Structures like oblique shock waves and turbulent mixing of the two streams can not be modeled by this kind of method. Instead, these models use efficiencies that have to be tuned to take into account the losses generated by these effects. Nevertheless, thermodynamic models allow to very rapidly obtain the state and operations parameters at key points of the ejectors as well as the overall performance parameters.
On the other hand it is possible to get more details of the flow with CFD software models. This method can use the differential form of the conservation equations and allows therefore to model the spatial distributions of any variables of the ejectors. However, these models are a lot more expensive in computational time compared to thermodynamic models. It could take several days or weeks [6, 10] to obtain one simulation while a thermodynamic model takes a few seconds.

### 1.3.1 Thermodynamic and analytical modeling

Analytical models were the first to appear due to their simplicity. They are based on thermodynamic and compressible flow theory. Hence, this is why they are also called thermodynamic models. The first model, which was created by Keenan et al. [11] was considering critical condition with no friction losses. It assumed constant area-mixing and ideal gas. The first constant pressure-mixing model, which are those that are the most used, were created only a few years later by the same team [12].
Since then, analytical models now take into account isentropic and mixing losses with efficiency coefficients. They assume, for most of them, satisfied chocking conditions (on-design) and constant pressure mixing [3].
The process is based on an iterative procedure where a series of compressible flow equations is solved step-by-step for selected zones. Except for a few exceptions [13], these models are, at most, 1D. This means that the calculated variables are assumed to be uniformly distributed in the radial direction. A detailed explanation of the Chen et al [14] model improved by Metsue et al. [1] and the resulting equations will be drawn in chapter 2. This improved model enlarges the operating regime to off-design application for both single and two-phaseflows.

The most challenging issue is to model the mixing phenomena. The first step was proposed by Munday and Bagster [7, 15] with the principle of two separated flows; the primary and secondary flows which each had their own velocity and pressure into the mixing chamber until they mix after the hypothetical throat, represented

as section y on figure 1.3. It is at this section that the secondary flow becomes choked for on-design application.
Two explanations can be used to describe the way the flow becomes choked at the hypothetical throat: The Fabri-choking [16] and the compound-choking [17]. The Fabri-choking simply assumes that a flow composed of multiple different flows is chocked if every flow constituting it would be choked if it was alone. It is the most used chocking mechanism for ejectors in the literature. The compound-choking criterion is an old theory but it has only been applied quite recently because the more recent models tend to prefer it [1]. It assumes that the flow is choked when a sum composed of heat capacities, section and mach number of each fluid become inferior to 0. By this way, the compound-choking theory assumes that, when considering only two flows, if the mach number of the first flow is greater than one, the flow becomes choked when the second flow has still a mach number below one.

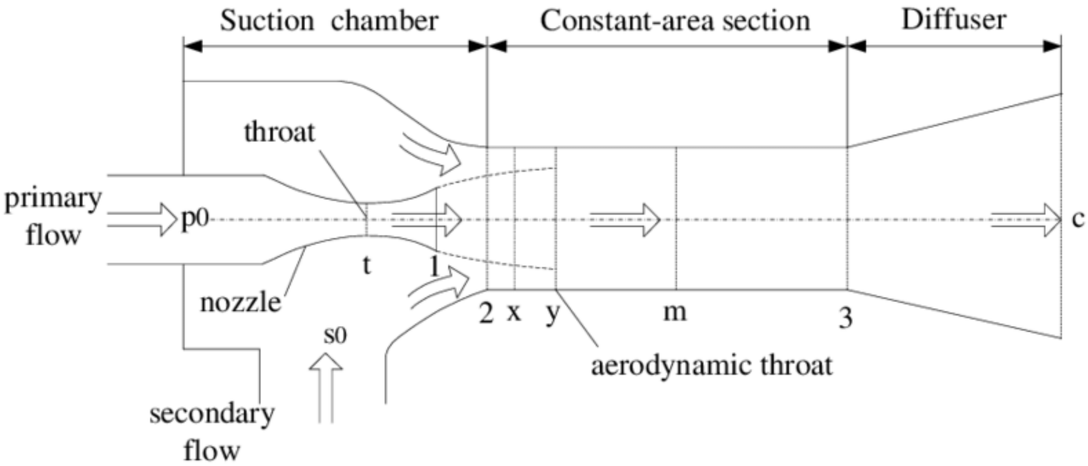


Figure 1.3: Schematic of ejector layout [18]

### 1.3.2 Numerical ejector modeling

As said earlier, the thermodynamic models remain quite limited due to what they are built on at the beginning. Indeed, even if these models are very useful for the preliminary design and operational information, local information can not be obtained by the integral form of the conservation equations. This includes internal friction and stream mixing that are needed to calculate efficiencies. It also means that assumptions are required such as uniform radial and axial distributions and the existence of unique normal shock. However, nothing confirms the accuracy of these assumptions [6]. CFD models thereby are a way to deal with these limitations by considering the local form of the conservation equations and therefore allow the

2D-3D modeling.

All fluids behave according to the Navier-Stokes equations which represent respectively the conservation of mass, momentum and energy. For most of applications an analytical solution of these equations can not be found hence the need for models to characterise the flow.

In matters of accuracy and reliability, CFD are, from far, the best tools to model an ejector. They are able to model the flow in 2 or 3D and can be established without restrictions on initial conditions and impositions of correction factors. Moreover, these models can take into account and compute accurately complex phenomena that occurs within an ejector like mixing, shock waves, phase change [7]. CFD allowed to have a better understanding of the ejector and do the research progress about ejectors. However, despite these many qualities, thermodynamic models are still used nowadays, because CFD remain more complex and expensive in time as well as in infrastructure costs. Moreover, these models are sensitive to the numerical setting and discretization.

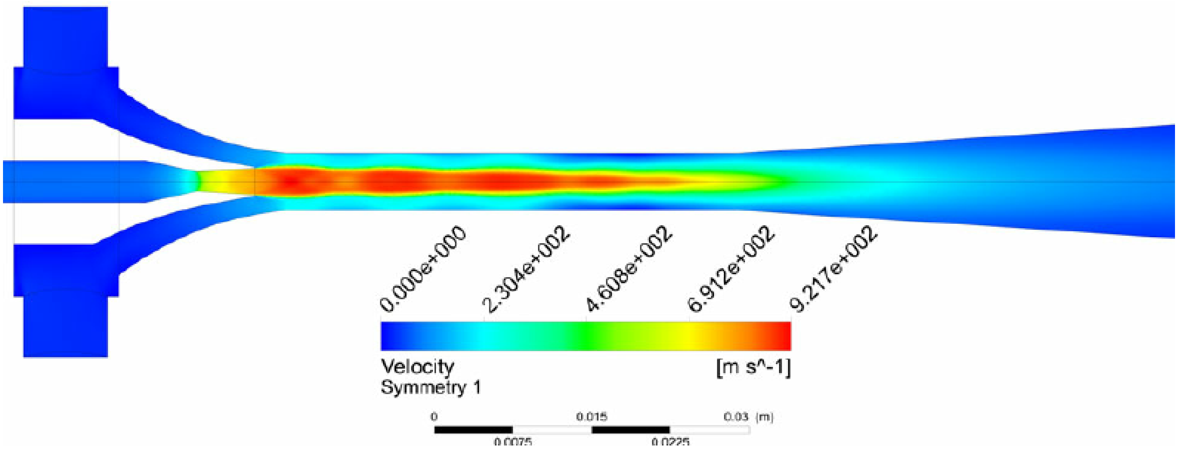


Figure 1.4: CFD model of steam-air ejector for turbo-ejector power system [19]

The CFD for ejector appeared long after the analytical models (in the 1990's). The first tests of Computational Fluid Dynamics on ejectors lacked experimental validations but allowed to make already a good visualization of their local flows. For example the work Kim et al. [20] and Neve [21] gives an idea of the these first works. With constant progress of computational power and the new developments in data-processing technologies, one can now be more comprehensive for many ejector studies. A paramount choice for the modelisation was the turbulence model.

CFD compared to data has been widely used for the evaluation and selection of turbulence models [8]. Moreover, they gave a better idea of the location of the hypothetical throat [6]. Lamberts et al. [22] highlighted a similarity with the Fabri-chocking for the secondary flow chocking phenomenology of the secondary stream. Other numerous works with CFD were focused on the stream mixing and they established a link between local flow and global parameters [7]. CFD are useful to design and adapt geometry and operation parameters [6].
One last application where CFD are helpful is to perform sensitivity analysis for a wide range of cases. As did by Brezgin et al. [23].

## 1.4 Scope of this study

In thermodynamic models, the value of the efficiencies varies from an ejector to an other, different geometry involving different losses. It is expected that these efficiencies vary with pressures temperatures and sections but it is not necessarily the case. In these models, efficiencies must thereby be calibrated.

In this study the thermodynamic model used is that from Metsue et al. [1]. It contains 5 efficiencies to calibrate. Chapter 2 contains a review of the previous similar thermodynamic models followed by a brief explanation of the one used in this work.

The main purpose of this thesis is to calibrate the thermodynamic model by the means of a machine learning model. Such a model would be able to find the five efficiencies from the other input parameters of the thermodynamic model allowing it to return the corresponding mass flows. A flowchart of this process is illustrated in figure 1.5. It would be trained by sets of data containing the input parameters as well as the outputs which are the mass flows.

Before starting to create a machine learning model, it must be ensured that all the conditions to make such a model work are fulfilled. At first, the way that the model would be trained is with efficiencies found by fitting the outputs of the model with the actual values of the data for the same inputs. A schematic of this process is depicted in figure 1.6. In an other way, it could be possible to merge the two steps of the training into one single step as illustrated in figure 1.7. Although this way to proceed would be ideal, this is more complicated to put in place because of the need of a back propagation along the thermodynamic model. It will therefore not be applied in this thesis.

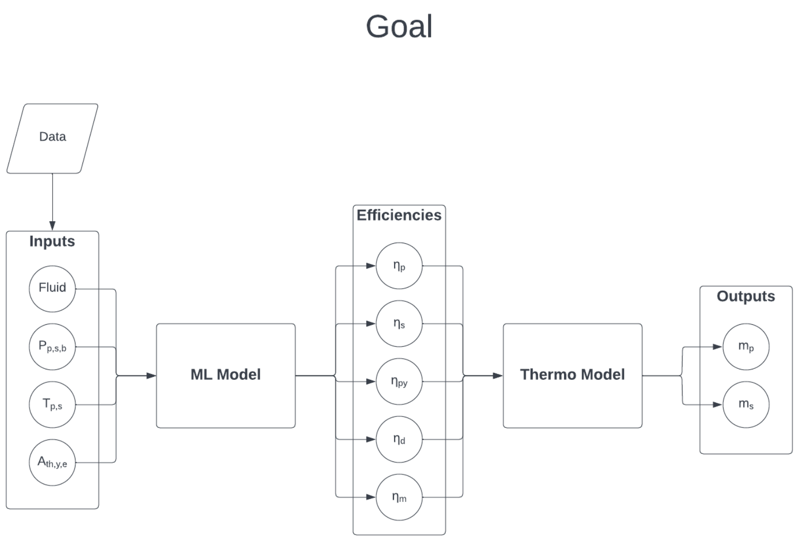


Figure 1.5: Flowchart of the goal to achieve with the machine learning model

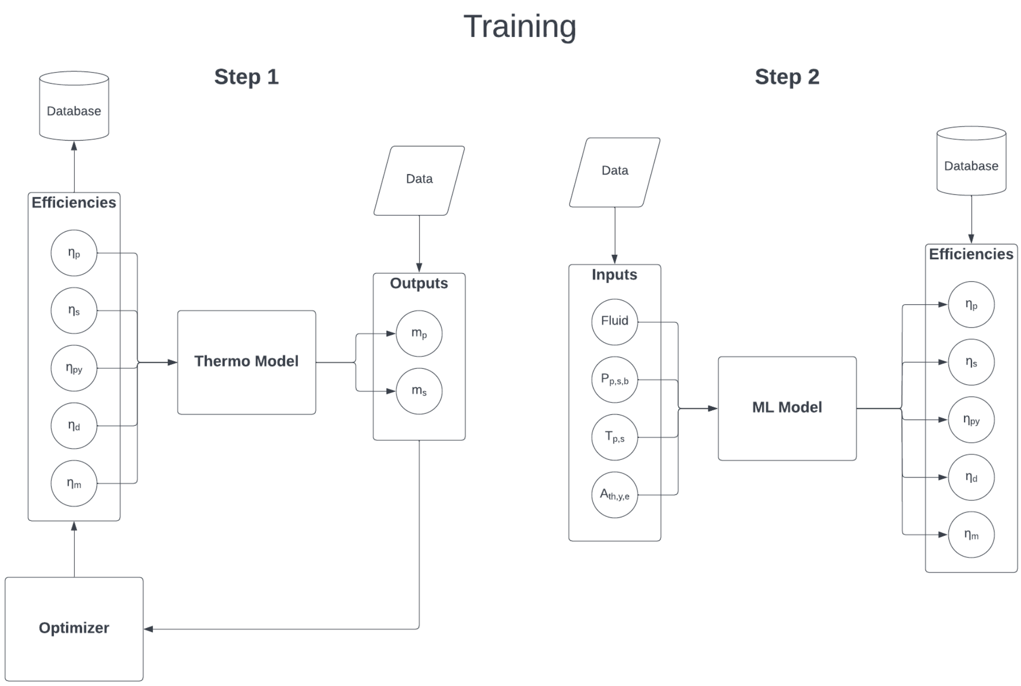


Figure 1.6: Flowchart of the way to train to machine learning model

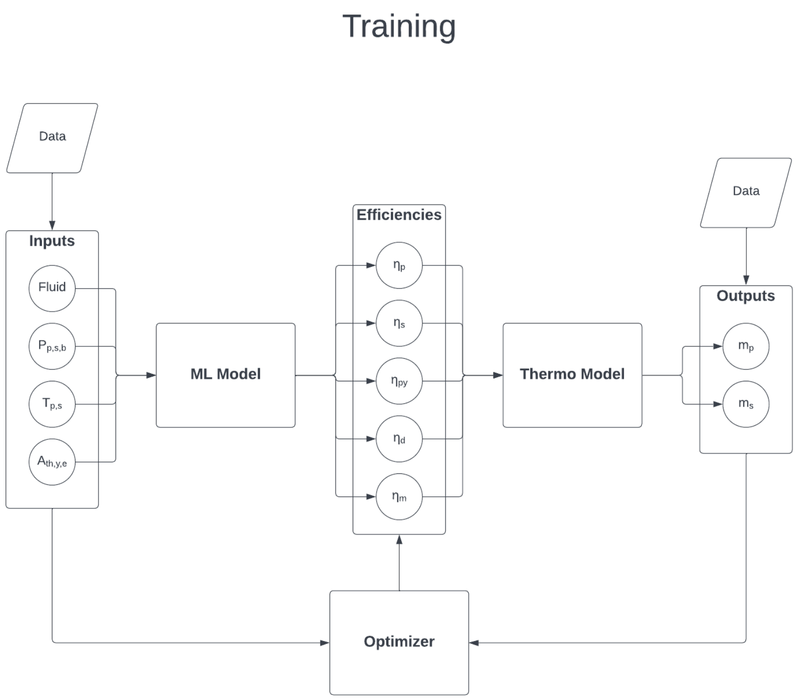


Figure 1.7: Flowchart of another way to train the machine learning model

The machine learning model will not work if, for the same input, multiple outputs can exist. This leads to the first question that will be studied in this thesis : **Is there more than one combination of the five efficiencies to predict the correct mass flow rate with fixed pressures temperatures and sections ?** This question will be the main topic of chapter 3.

If the answer to this question is yes, then, it must no longer be the case. There is indeed one single set of five efficiencies that is working during the operating of an ejector. By searching for solutions for this issue follow the questions : **Are there too many efficiencies defined in the model ? Is each of them relevant?** This question will be studied in the first part of chapter 4.

On the other hand, it may be wise to ask if all the inputs and outputs together are enough to characterise sufficiently well the physics of the ejector. Maybe more efficiencies should remains in the thermodynamic model and the outputs are not sufficiently numerous. The second part of chapter 4 will be dedicated to this topic while trying to answer the following question : **Are the mass flow rates alone, accompanied by the efficiencies, sufficient to characterise all the states on the ejector?**

Finally, by analysing a data-set point per point, the variation of the parameters in the vicinity of the points is not taken into account. Yet, the way the neighbouring of a point varies according to the back pressure in its parameters space can have an impact on the calibration of efficiencies. This leads to the question **what impact could taking into account the space in the vicinity of a point have on the calibration ?** This last question will be discussed in chapter 5.
The conclusions drawn from this thesis, as well as the perspectives for future works, are gathered in chapter 6.

# Chapter 2

# Thermodynamic model

As stated earlier, this study will focus on the efficiencies calibration of a thermodynamic model. Before starting with the research, an explanation of the model has to be made to make sure the reader has the key to understand what will follow.
Therefore, this chapter aims to explain in more details the conception of the thermodynamic model that will be calibrated during all this study. It will be divided into two parts.
First, a review of similar models that appeared before the actual one and from which it can be inspired will be discussed. A brief review of the assumptions the actual model is based on will be made at the beginning.
Afterwards, a more detailed explanation will be made about the algorithm that composes the model with a special focus on the five efficiencies, as stated, what they means and how they are defined.

This chapter does not aim to explain in detail how the model works. For more information, the reader is invited to read the paper of Metsue et al. [1].

## 2.1 Thermodynamic model review

The model used in this work is a new state-of-the-art thermodynamic model appeared in 2021 and created by Metsue et al. [1]. In this one, most of the assumptions are taken from the model of Chen et al. [24]. As said these are :

- the flow is one dimensional, steady and adiabatic, The inlet and outlet velocities are supposed to be negligible.
- Real gas properties are considered.
- Friction losses are characterised with five efficiencies. There is one more, compared to Chen et al. [24] in order to take into account the losses on the

diffuser. The definition of these five efficiencies will be discussed in detail in the next section.

- Hypothetical throat assumption is used so that the mixing does not start before this section [25].

However, instead of the constant pressure mixing which is generally used as a mixing assumption, this model has introduced the mass conservation equation and, this time, the hypothetical throat is described according to the compound-choking criterion. The model has been firstly validated for the use of a single phase ejector and, afterwards, for two phases ejectors as well [26].

The first model which will be discussed in this chapter is the model of Huang et al. [27]. This model dates from 1998 when CFD and computer power did not allow simulations as precise as now. It is a 1D-analysis that can only take into account critical mode operation (on-design regime). It assumes steady flow and ideal gas theory. Moreover, Huang et al. [27] used Bagster et al. [15] theory for its model which means that it assumed that the mixing occurs at the constant section area with constant pressure after the hypothetical throat. The schematic diagram of ejector performance used for this model is shown in figure 2.1. The following models presented in this section will also be based on a diagram of this type. Moreover, the ejector is considered to be adiabatic and the losses are represented with 2 isentropic efficiencies and two other arbitrary losses coefficients.
The two efficiencies are noted $\eta_p$ and $\eta_s$. They represent the losses in the primary and secondary nozzles, respectively, by taking into account the differences of enthalpy between the process without frictional losses and the experimental process. In addition, an arbitrary coefficient noted $\phi_p$ takes into consideration the losses of the primary flow from section 1-1 to section y-y of figure 2.1. Finally, there is a last coefficient noted $\phi_m$ that represents the frictional losses during the mixing process. Because they depend on too much parameters like machining, center-line alignment, interior surface polishing, material used and suction port configuration etc. [27], These efficiencies have to be found experimentally. The 11 ejectors used during the study of Huang et al. [27] have suggested that $\eta_p$, $\eta_s$ and $\phi_p$ can be taken as relatively constant but that $\phi_m$ depends on the sections. This last dependence can roughly be represented with the following law:

$$\phi_m = 1.037 - 0.02857\frac{A_3}{A_t} \tag{2.1}$$

Based in particular on the work of Huang et al. [27], Chen et al. [14] created a new 1D model in 2013 followed by another more precise one in 2017 [24]. The main improvement of the first model is the extension of predictions to sub-critical regime (off-design). For the assumptions, compared to the model of [27]:

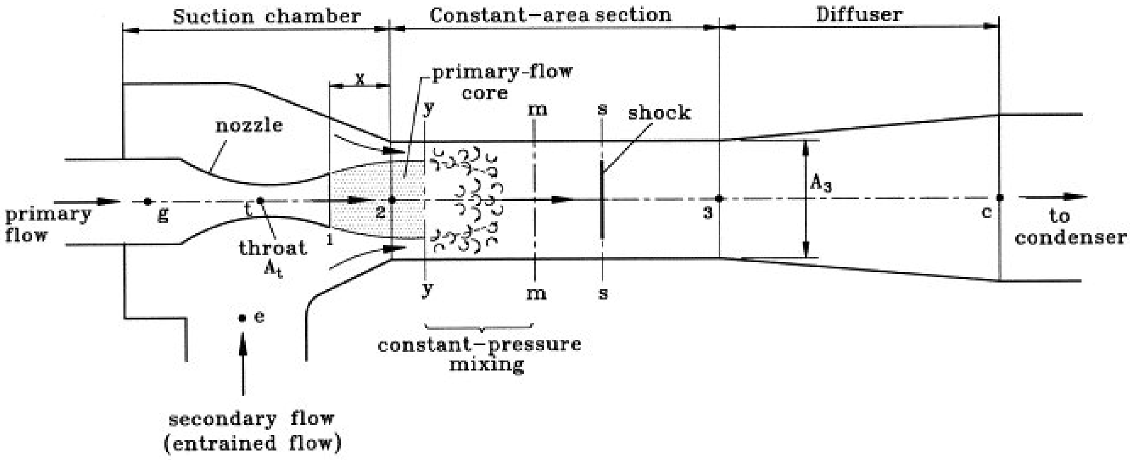


Figure 2.1: Schematic diagram of ejector performance [27].

- The model still considers Bagster et al. [15] theory by assuming hypothetical throat even for off-design regime where there is no choked flow but a maximum velocity that is achieved at this section.
- The mixing is still considered to be done in all cases after the hypothetical throat section with uniform pressure.
- Ideal gas assumption is considered for the working.
- The flow is steady and one-dimensional, the inner wall of the ejector is adiabatic and the kinetic energy at inlets, as well as the exit of diffuser, is negligible.

As for the losses coefficients, they are still 4 of them and they are used in the same way as for the model of Huang et al [27]. Using the data-set from Huang et al. [27], Chen et al. [14] model has shown roughly the same accuracy as the latter in the on-design regime. For the off-design operation, data from Hemidi et al. [28] has been used and the theoretical results coincide fairly well compared to the model. The second theoretical analysis made by Chen et al. [24] has allowed to consider real gas instead of ideal gas assumption. This has been done with the help of the REFPROP [29] library. This last proposed model has given better accuracy in off-design mode compared to the model using ideal gas law.

All this makes the foundation for the conception of the model studied in this work. The main difference between the model of Metsue et al. [1] and the last model of Chen et al. [24] resides in the use of the compound-choking criterion. Although the compound-choking criterion has been known since the 60', the Fabri-choking criterion had been preferred, until now, to define the flow in ejectors. It has nevertheless recently been shown that the compound-choking theory may be more adequate to predict the behaviour of the double choked ejectors. Indeed, by definition of a choked flow, the mass flow rate has to be maximized which is not the case in the ejector with the Fabri-choking criterion [16]. In addition, the mixing in the ejector does not occurs with constant pressure. This last consideration leads to another way to implement the computation of the state after mixing.

## 2.2 The model

### 2.2.1 Ejector solving algorithm

As stated in the previous section, the model is mainly inspired from Chen et al. [14, 24] research. The improvements are the use of the compound-choking criterion instead of the Fabri-choking criterion in the first place but also the process of maximization of the primary mass flow rate instead of imposing M=1 at the throat. With this latest assumption, the M=1 section is located just after the nozzle throat. In this way, the primary definition of a choked flow is respected.
Given that its purpose is to study supersonic ejectors, the model assumes choked primary flow. This means that the primary mass flow rate is constant with respect to the downstream quantities. Thereby, all the variables of the primary nozzle are calculated in the first place starting from the stagnation pressure and temperature. The losses in this nozzle are characterised with the primary efficiency which is supposed to be known. The schematic diagram of the upper half ejector for this model is shown in figure 2.2.

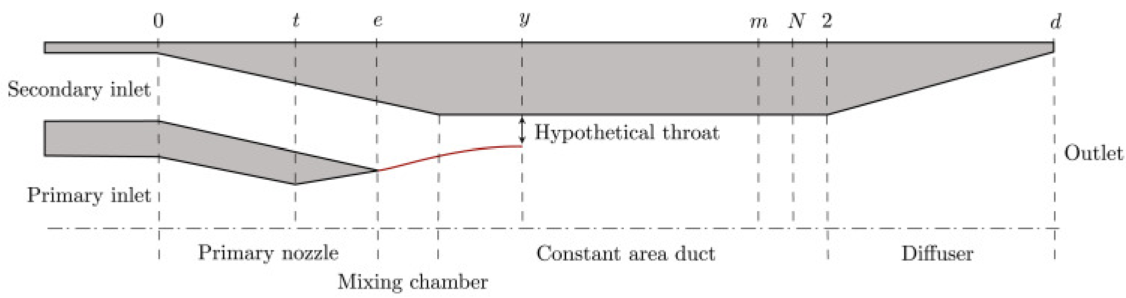


Figure 2.2: Schematic diagram of the ejector model (top half) [1]

Once the state of the primary flow is determined, the primary mass flow rate can be found which already gives the first output for the model. The thermodynamic state and mass flow rate of the secondary flow for on-design condition can be found afterwards with the compound chocking criterion [30]:

$$\beta = p_y(\frac{A_{p,y}}{\rho_{p,y}V_{p,y}^2}(1 - M_{p,y}^2) + \frac{A_{s,y}}{\rho_{s,y}V_{s,y}^2}(1 - M_{s,y}^2)) = 0 \tag{2.2}$$

This criterion is respected at the hypothetical throat. The state of the primary flow as well as the pressure need therefore to be known at this section in order to compute the state variables of the secondary flow. To do so, the value of $p_y$ is at first estimated. From this estimation, the secondary flow state can be calculated. The error with the chocking criterion is then measured to better estimate $p_y$ and, by iteration, a solution is found.

Once the secondary flow for on-design operation is found, the purpose of the model is to know if the current operation is on- or off-design. To do so, the critical back pressure is calculated. Thereby, if the back pressure, which is a parameter of the function, is lower than the critical pressure, the ejector regime is on-design. In this case, the secondary state has already been calculated. The entrainment ratio can be obtained and it is therefore not required to solve the flow downstream of section y. However, if the back pressure is higher than the critical back pressure, the regime is off-design. In this case, the information travels from the end of the diffuser. The secondary flow can thereby not be solved without taking into consideration what happens downstream of the hypothetical throat section.

In order to find the critical back pressure, the state just after mixing is found. This state corresponds to section m on the diagram. For doing this, the 3 conservation equations in addition to the equation of state linking density pressure and enthalpy are used.

If the ejector is at off-design operation, the state at the outlet of the diffuser must be calculated. The diffuser being divergent, the flow has to be subsonic to recover the outlet pressure. However, the flow can still be supersonic after the mixing. In this case, the model assumes a normal shock at section N to prevent the diffuser from being used as a divergent nozzle. Even if it is not the case in figure 2.2, sections m, N and 2 are assumed to be at the same location. Section m is the state just before and section 2 the state just after the normal shock. Finally, the state at the exit of the diffuser is obtained with the conservation equation. The exit state of the ejector being determined, the secondary flow state can now be computed as well as the entrainment ratio.

The complete flowcharts extracted from the paper of Metsue et al. [1] are displayed in Appendix B.

### 2.2.2 The efficiencies

The efficiencies being of prime interest during this work, a special attention is payed to them in this subsection.
As said earlier, the model of Metsue et el. [1] is composed of 5 efficiencies. Each of them represents losses in a part of the ejector.

- The so called primary efficiency, noted $\eta_p$, is a classical isentropic efficiency. It was already used in the model of Huang et al. [27] and after in the model of Chen et al. [14]. It represents the losses occurring in the primary nozzle of the ejector, as said between section 0 and section e of figure 2.2. In the algorithm, it appears with the following formula :

  $$h_p = h_{p,0} - \eta_p(h_{p,0} - h_{p,is}) \tag{2.3}$$

  Given that this efficiency only depends on the thermodynamic state of the primary flow, it should not depend on downstream quantities.
  According to the calibration that has been made in the different models, it is taken with values between 0.95 and 0.977 and is considered as constant.

- The efficiency noted $\eta_{py}$ is also present since the model of Huang et al. [27]. According to the models, both flows are considered with independent states until section y where the mixing begins [15]. $\eta_{py}$ takes into account the entropy losses between section e and section y for the primary flow. In the process, it appears with the following formula:

  $$h_{p,y} = h_{p,e} - \eta_{p,y}(h_{p,e} - h_{p,y,is}) \tag{2.4}$$

  The value can oscillates between 0.88 and 1.0 in the different models using it. As seen from the work of Metsue et al. [1], the operating curve is less sensitive with this efficiency compared to the others. Its value can rise when considering that the primary nozzle flow is almost isentropic which can be the case for real ejectors.

- The secondary efficiency, as well as the two others, is an isentropic efficiency. It is noted $\eta_s$ and represents all the frictional losses for the secondary flow before mixing (between section 0 and y). It appears with the following formula of the algorithm:

  $$h_{s,y} = h_{s,0} - \eta_s(h_{s,0} - h_{s,y,is}) \tag{2.5}$$

  In on-design operation, this efficiency should not depends on what happens downstream. Its values taken in the different models are included between

0.75 and 0.89. Those values are lower because models consider more losses compared to primary efficiencies due to the less convenient geometry of the suction chamber.

- The mixing efficiency is quite different from the others. First, it is not an isentropic efficiency, moreover, it is used for the first time with this definition in the model of Metsue et al. [1]. The model no longer assuming a constant pressure for the mixing, this efficiency takes into account the losses occurring during the mixing due to the pressure change. It appears therefore in the momentum conservation equation:

$$V_m = \eta_m \frac{\dot{m}_p V_{p,y} + \dot{m}_s V_{s,y} + (p_y - p_m) A_y}{\dot{m}_p + \dot{m}_s} \tag{2.6}$$

  With this definition, there is no longer the need to be dependant of the exit pressure for this efficiency as opposed to the definition of Chen et al. [14, 24].

- The diffuser efficiency is again a classical isentropic efficiency. It is of interest for off-design applications, which explains its absence in the model of Huang et al. [27]. It characterises the losses of the mixed flow between section 2 and section d with the following definition:

$$\eta_d = \frac{h_{d,is} - h_2}{h_d - h_2} \tag{2.7}$$

These efficiencies are parameters of the model. They are therefore supposed to be known. Given that they depend on too many physical quantities that can not be obtained in the vast majority of cases, they need to be calibrated with experimental data. For now, the dependencies of these efficiencies according to pressures, temperatures, mass flows and dimensions of the ejector are not well known. Thereby, the major part of this work will be dedicated to a way to calibrate those efficiencies by the knowledge of the other parameters. The final purpose would be to obtain immediately a calibrated model when given to it any combination of parameters. All this thanks to a machine learning model and a large amount of data. The following chapters will study the feasibility of such a model and also inspect the different links between efficiencies.

# Chapter 3

# Efficiencies fitting

One of the big issue with a machine learning model is the fact that one set of input parameters has to correspond with one set of outputs which is not necessarily the case for the thermodynamic model.
In this case, the input parameters of the model are the physical quantities, as said: the pressures and temperatures, the geometrical quantities i.e. the sections, and the five efficiencies. The outputs are the primary mass flow, the secondary mass flow and the entrainment ratio.
For now, nothing proves that there could not be several inputs that correspond physically to multiple sets of inputs.
The question of this chapter is to find out if different combinations of efficiencies $\eta_p$, $\eta_s$, $\eta_{py}$, $\eta_d$ and $\eta_m$ are possible to obtain the same entrainment ratio with given parameters. To do so, the model will be calibrated with the help of a gradient descent optimizer starting from different first guesses. This process could allow to see different solutions appear when the optimizer stops in local minima.
In a first approach, the focus is set on an entire operating curve, , which means that the efficiencies are supposed to be independent from the back pressure. In this way, it will be observed whether one or more sets of efficiencies can exist to fit a whole operating curve.
Afterwards, the aim will be to find out whether different sets of efficiencies are possible but for the individual mass flow rates. This time, it is considered that the efficiencies are no longer independent from the back pressure and the inlet conditions because the calibration leads to a unique set of efficiencies for each point.

## 3.1 Calibration of the entrainment ratio for an entire operating curve

For this part of the study, Data from Hemidi et al. [31] has been used. The calibration has been carried out on a 4 bars primary pressure operating curve composed of 10 points. The parameters and measurements are listed in table 3.1.

| $P_{p,0}(Pa)$ | $T_p(K)$ | $P_{s,0}(Pa)$ | $T_s(K)$ | $P_b(Pa)$ | $\omega(-)$ |
|---|---|---|---|---|---|
| 400000 | 295 | 100000 | 295 | 102629 | 0.75885 |
| 400000 | 295 | 100000 | 295 | 112741 | 0.75717 |
| 400000 | 295 | 100000 | 295 | 122573 | 0.75765 |
| 400000 | 295 | 100000 | 295 | 127591 | 0.75670 |
| 400000 | 295 | 100000 | 295 | 132329 | 0.73348 |
| 400000 | 295 | 100000 | 295 | 134825 | 0.70596 |
| 400000 | 295 | 100000 | 295 | 137066 | 0.65810 |
| 400000 | 295 | 100000 | 295 | 141906 | 0.53748 |
| 400000 | 295 | 100000 | 295 | 146363 | 0.42955 |
| 400000 | 295 | 100000 | 295 | 151254 | 0.31468 |
| 400000 | 295 | 100000 | 295 | 155839 | 0.15936 |
| 400000 | 295 | 100000 | 295 | 158666 | 0.05310 |

Table 3.1: Operating conditions and measurements of Hemidi et al. [31]

### 3.1.1 Optimizer and methodology

In order to avoid too long computation time, the optimization has been separated in two parts. First, the on-design part of the curve is fitted by the calibration of the three first efficiencies i.e. $\eta_p$, $\eta_s$ and $\eta_{py}$. The two remaining efficiencies having no impact on the on-design part of the curve [32], there is no need to include them for the fitting of this part.

Afterwards, once the 5 on-design points are well calibrated, the efficiencies $\eta_d$ and $\eta_m$ are used to fit the whole curve including the remaining off-design points. In this way, the two last efficiencies will fit the remaining points without impacting the first calibration.

This separation adds complexity to the problem because the on-design points are calibrated twice. However, it allows to be more precise in the research of solutions by limiting the computation time. The computation time for an optimizer being exponential with its number of parameters, carrying out two separate optimizations with 2 and 3 parameters is several times shorter than once with 5 parameters.

The cost function used to quantify the accuracy of a solution is the Mean squared error (MSE) which takes the sum of the square of the difference between model and data points composing the curve. In mathematical form, it gives :

$$\min \sum_{i=1}^{n} ||\omega_i^p - \omega_i||^2 \tag{3.1}$$

Where $\omega_i^p$ is the entrainment ratio returned by the model for the ith point on the n points for which the calibration is done. From the optimizer, the only way to do vary $\omega_i^p$ is by changing the efficiencies.

The minimization is done with the optimizer using L-BFGS-B method from scipy.optimise package [33]. The L-BFGS algorithm is an extension of the quasi-newton Broyden-Fletcher-Goldfarb-Shanno (BFGS) algorithm [34] for which the L means "limited memory". This updated version uses a lighter approximation of the Hessian matrix which makes it more suited for solving optimization problem with numerous variables [35]. The L-BFGS-B is an extension of the L-BFGS which allows to add simple bounds on the optimized variables by using ideas from the trust region methods [36].

the optimizer parameters have to be bounded to avoid error code by achieving the limitations of the thermodynamic model with an impossible value for an efficiency. This last extension brings the main feature for what is needed in order to calibrate correctly the efficiencies. Indeed, the model can not handle any impossible value for an efficiency that might be proposed by the optimizer.
Firstly, by its definition, an efficiency can not be bigger than one. This gives the upper bounds. Secondly, the lower bounds need to be all above 0, a negative efficiency being meaningless. Nevertheless, even if none of the data points belong to the back-flow regime, by playing with efficiencies the optimizer could explore this region. The thermodynamic model of Metsue [32] is not able to handle cases for which the primary flow is reverses back into secondary nozzle (backflow). this situation must be avoided. The problem happens often when the mixing efficiency starts to decrease. For solving it, the choice to forbid the optimizer to use a value lower than 0.9 has been made. This decision should not cause losses of solutions, because its value is normally close to one due to the way it is define (see chapter 2). In fact, this efficiency acts in a momentum balance equation. It represents unknown forces between sections py, sy and m. Given the generally high momentum, it can be expected that the momentum equation is almost exact and therefore close to one. This problem is not encountered with the four other efficiencies for any of their values. However, the primary flow being assumed to be choked by the model, if a too low value for the primary efficiency is taken

by the optimizer, the thermodynamic model would not find a choked solution for the primary nozzle and would cause the optimization procedure to stop. It is reasonable to expect that this efficiency will take values from 0.9 to 1.0. By trial and error the optimization process behaved until values under 0.7 which again should not cause losses of physical solutions. A value of 0.7 as bound for the primary efficiency seems therefore to be reasonable. All this keeping in mind that non-choked primary flows are not in the interest of this study. Thereby the five intervals for the variables that will be entered in the inputs of the optimizer are :

$$\eta_p \in [0.70, 1.0] \quad \eta_s, \eta_{py}, \eta_d \in [0.0, 1.0] \quad \eta_m \in [0.9, 1.0]$$

The first part of the optimization is then to fit the on-design part with the $\eta_p$, $\eta_s$ and $\eta_{py}$. The main purpose being to find if different sets of parameters can exist for a single curve, the starting vector should be selected so as to find a maximum of local minima in the solution space according to the parameters. In this way, the optimizer is launched each time from a different starting point. Thus, for this part, the starting vector is uniformly distributed between the bounds of each of the three first efficiencies $\eta_p$, $\eta_s$ and $\eta_{py}$. In more details, the range for each efficiency is divided with a 5 length vector thus forming a 5x5x5 grid. The optimizer is therefore launched 125 times. On this scheme, X represents the starting vector and i j k the indices of the three linspace vectors. Because of the fact that it is very unlikely to obtain efficiencies smaller than 0.5, for this part of this study, the choice to change the lower bounds of $\eta_s$, $\eta_{py}$ and $\eta_d$ from 0.0 to 0.5 has been made. Thanks to this, the space of efficiencies above 0.5 is more precisely explored while keeping computation time sufficiently low.

For the second part of the calibration, the two remaining efficiencies $\eta_d$ and $\eta_m$ are used to calibrate the whole curve from all the sets of three efficiencies found with the previous part (see figure 3.1). The cost function is the same with the difference that the sum is done on all the 10 points of the curve. The bounds are the same and sometimes adapted with the set of efficiencies if the mixing of efficiencies need higher bounds for a certain set of the three first efficiencies.

### 3.1.2 Results and comments

The first part of the methodology has given 63 different results on a total of 125 different launchings in a uniformly distributed space of parameters for the starting vector. This result already means that, when calibrating only for the on-design part of the curve, there are multiple solutions. Moreover one can not talk about few solutions. There is indeed a different result for more than the half of the launchings.

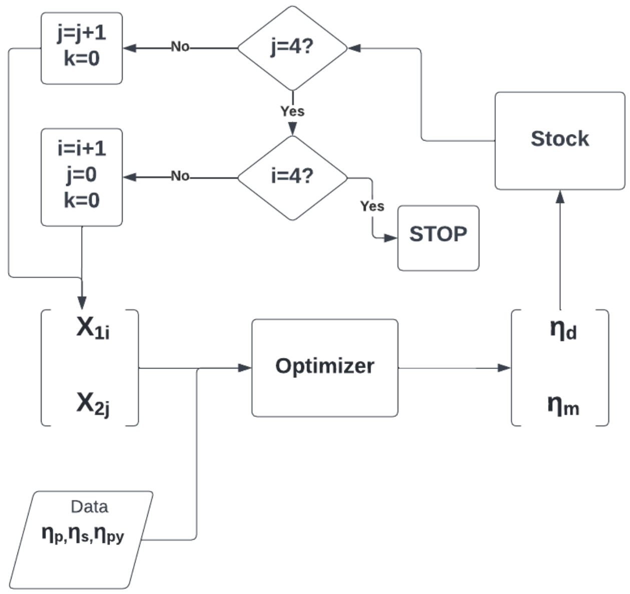


Figure 3.1: Flowchart of the optimization process (second step)

Nevertheless, these results do not take into account the mass flows and the off-design part of the curve. It is therefore expected that less solutions will appear taking into account the last points with the two remaining efficiencies. This first result suggests however that the parameters do not allow to find the entrainment ratio, by a single way, at least for the on-design part. It should however not be forgotten that the only criterion in the cost function is here the entrainment ratio. It is therefore expected that multiple solutions could exists. A same entrainment ratio can be retrieved with piratically no losses for the two flows or with big losses too. The resulted efficiencies spread indeed in all their authorised domain for this part of the calibration.

After that, starting from these 63 different results, the whole curve has now been fitted with the two remaining efficiencies, keeping the previous three coefficients constant. Eight different combinations resulted in a MSE below $10^-3$ and are listed in table 3.2. The curves are illustrated on figure 3.2.

From this last result, two assumptions can be made. At first, when looking at the pair of points 0 1, 4,5 and 6,7. It is highly probable that the two solutions come from a low sensitivity of several efficiencies into certain regions that keeps the solution below the maximum permitted MSE. Indeed, it can be seen on the figure that the difference between curves 0 and 1 is quite high but they both are in the interval of tolerance. After that the points 4,5 and 6,7 have almost

| i | $\eta_p$ | $\eta_s$ | $\eta_{py}$ | $\eta_d$ | $\eta_m$ |
|---|---|---|---|---|---|
| 0 | 0.9123 | 0.6744 | 0.8349 | 0.6000 | 0.9750 |
| 1 | 0.9074 | 0.6836 | 0.9922 | 0.6923 | 0.9949 |
| 2 | 0.7971 | 0.5912 | 0.5000 | 0.8740 | 0.9975 |
| 3 | 0.8534 | 0.6322 | 0.6690 | 0.7697 | 0.9701 |
| 4 | 0.9379 | 0.7000 | 0.9994 | 0.6000 | 0.9773 |
| 5 | 0.9379 | 0.6990 | 0.9868 | 0.6010 | 0.9732 |
| 6 | 0.9485 | 0.7047 | 0.9916 | 0.5554 | 0.9742 |
| 7 | 0.9380 | 0.7000 | 0.9994 | 0.6525 | 0.9580 |

Table 3.2: Values of the 5 efficiencies for the 8 curves that fit all the Data points with MSE<$10^{-3}$

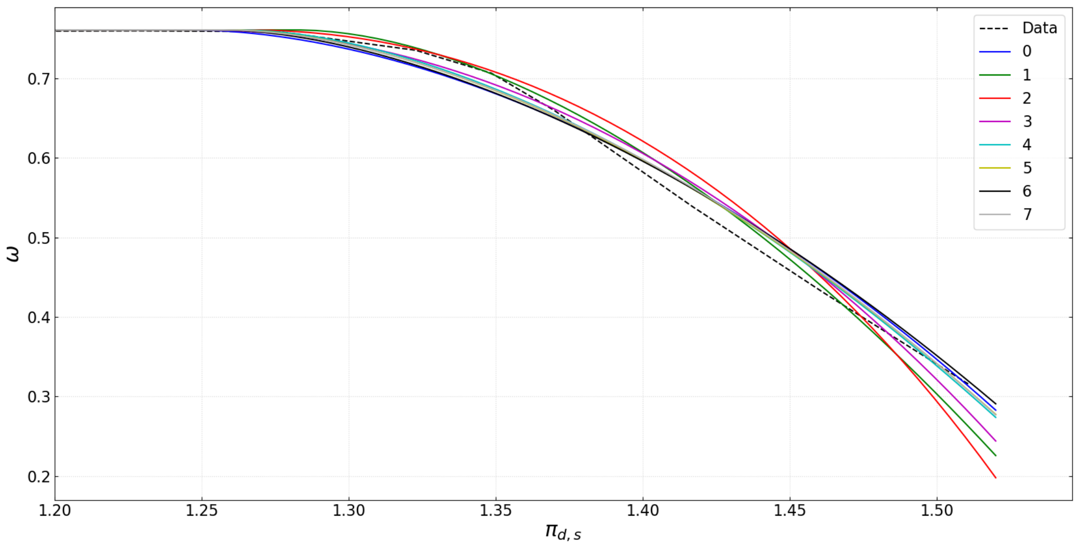


Figure 3.2: Operating curves computed by the model vs experiment of Hemidi et. al [31]. $\omega$ is the entrainment ratio, $\pi_{d,s}$ is the pressure ratio between the outlet pressure and the secondary one.

the same efficiencies. Sets 4 and 5 have such a small difference that they can be considered as the same set. The differences between the two curves are visible anyway. This shows that the efficiencies remain sensitive with $\omega$. The other pairs indicate probably the presence of a flat minimum especially for the curves 6 and 7 which have almost the same curves and 10 percent different $\eta_d$. On the other hand the difference between most of the points can also be explained by some efficiency that counteract the effect of the others by the change of one efficiency when an other operating in the same region decreases or increases. For example, the difuser and mixing efficiencies causes similar effects on an operating curve [1] and between curves 0 and 7, there is a decrease of $\eta_m$ whereas an augmentation of $\eta_d$ can be observed. This is also visible between sets 6 and 7. this already suggests that there are too many efficiencies in the model to be calibrated only on the entrainment ratio. Based on the global information from the mass flow rates, the losses can not be localized in the mixing duct or in the diffuser. So, the two can not be both calibrated and one should be fixed. This latter assumption will be studied more precisely in chapter 4.

However, when looking at the large difference of curve 2 with each of the efficiencies compared to the other curves, it is unlikely that it comes from low sensitivity of the efficiencies. Indeed, this solution suggests that an other minimum in the space of the cost function exists and that there are indeed multiple possible solutions for the calibration of the entrainment ratio along an entire operating curve.

## 3.2 Minimization point per point with complete outputs

### 3.2.1 Methodology

Given that the internal flow field can be largely different depending on the operating conditions, different efficiencies might be needed to reproduce an experimental curve with the model. Consequently, the idea of considering not the entire curve but a single operating condition point is proposed in this section.
Moreover, the entrainment ratio was considered for the calibration without taking into account the individual mass flow rates. Given that more solutions should appear when considering a single point and based on the findings in the previous section, it is deemed better to consider both mass flow rates in order to limit the possible solutions. This also allow to have a more determined calibration problem even if losses in the mixing pipe from the diffuser can still not be distinguished. The data-sets containing mass-flows with entrainment ratio are nevertheless fewer in the literature.

The process is separated in two steps:
First, the calibration is done for the entrainment ratio using exactly the same process as the previous section with the difference that it will be done point per point instead of with a set of points. Since the calibration is more than 3 times faster with a single point, the linspace vector has been doubled in length which means that the optimization is launched 1000 times for this step.
After that, once all the solutions for the efficiencies with correct entrainment

ratio are obtained, the mass flows obtained with those results are computed and compared with the real ones. The sets with sufficiently good mass flows are then selected.
For the first step, the same error criterion for selecting a result or not has been taken. As said : $MSE < 10^3$. For the second step, all the results having a mass-flow in a gap of $\pm 0.5\%$ are kept.
For this part of the study, data from Mazzelli et al. [37] is used. More precisely, the point for which the process is performed corresponds to an on-design point with the following specifications :

| $P_{p,0}(Pa)$ | $T_p(K)$ | $P_{s,0}(Pa)$ | $T_s(K)$ | $P_b(Pa)$ | $m_p(kg/s)$ | $m_s(kg/s)$ |
|---|---|---|---|---|---|---|
| 500000 | 295 | 97400 | 295 | 120000 | 0.328 | 0.166 |

The choice has been made for an on-design point because this regime is easier to model and this is the desired regime to achieve for an ejector.
The geometry of the ejector used for the experiment is drawn on Appendix A.1.

### 3.2.2 Results and comments

The error criterion for the selection of solutions during the first step turned out to be nonrestrictive. Indeed, the optimizer has no trouble to find a solution that is almost zero for a single point with the five efficiencies. Therefore, all the different solutions have been taken for this part.
Once the entrainment ratio was known to be predicted accurately, only the stocked sets of efficiencies for which the primary mass flow rate was 0.5% around 0.328 has been kept which reduced the number of solutions to 13. This means that the individual mass flow rates are absolutely needed and not only the entrainment ratio. Otherwise, the same ratio can be reached with low or high efficiencies and thus low or high flow rates. The sets with their associate primary mass flows are shown on tab 3.3.

Except for $\eta_m$, it can be observed that all efficiencies admit a deviation. It remains quite small for $\eta_p$ that can vary until 2%. It is small enough to consider it as uniquely defined. This means that $\eta_p$ can be fully defined from the primary mass flow rate which makes sense when looking at the equations for a choked primary nozzle [1]. This last result can simplify the calibration strategy and will be experimented on chapter 4. However, one can no longer talk about small variation for $\eta_s$, $\eta_{py}$ and $\eta_d$. The gap between the smallest and largest efficiency is indeed of 6.4% for the secondary efficiency, 23.5% for $\eta_{py}$ and 50.9% for the diffuser efficiency. Moreover it can be seen with $\eta_s$ and $\eta_{py}$ that there are two clusters of values. They

| i | $\eta_p$ | $\eta_s$ | $\eta_{py}$ | $\eta_d$ | $\eta_m$ | $m_p$ | $\omega$ |
|---|---|---|---|---|---|---|---|
| 0 | 0.930 | 1.0 | 0.467 | 0.439 | 1.0 | 0.327 | |
| 1 | 0.936 | 1.0 | 0.482 | 0.707 | 1.0 | 0.328 | 0.506 |
| 2 | 0.939 | 0.999 | 0.489 | 0.713 | 1.0 | 0.329 | 0.506 |
| 3 | 0.935 | 0.945 | 0.679 | 0.255 | 1.0 | 0.328 | 0.506 |
| 4 | 0.929 | 0.937 | 0.691 | 0.278 | 1.0 | 0.327 | 0.506 |
| 5 | 0.929 | 0.956 | 0.623 | 0.367 | 1.0 | 0.327 | 0.506 |
| 6 | 0.932 | 1.0 | 0.473 | 0.277 | 1.0 | 0.328 | 0.506 |
| 7 | 0.939 | 1.0 | 0.488 | 0.281 | 1.0 | 0.329 | 0.506 |
| 8 | 0.936 | 0.998 | 0.487 | 0.557 | 1.0 | 0.329 | 0.506 |
| 9 | 0.932 | 1.0 | 0.471 | 0.523 | 1.0 | 0.328 | 0.506 |
| 10 | 0.932 | 1.0 | 0.472 | 0.527 | 1.0 | 0.328 | 0.506 |
| 11 | 0.927 | 1.0 | 0.459 | 0.746 | 1.0 | 0.327 | 0.506 |
| 12 | 0.920 | 0.936 | 0.496 | 0.237 | 1.0 | 0.327 | 0.506 |

Table 3.3: List of the 13 solutions for the five efficiencies that match the entrainment ratio and the primary mass flows with a precision of 0.5%.

are located around 0.94 and 1.0 for $\eta_s$ and near 0.47 and 0.67 for $\eta_{py}$. This latter result proves the presence of local minima in the cost function instead of a flat minimum for these efficiencies. This means that more than one solution is possible for this point for the model. This also and more importantly means that there is another indetermined couple of efficiencies between $\eta_s$ and $\eta py$. They are inversely related: if one is high, the other is low like $\eta_d$ and $\eta_m$ in the previous section. As for $\eta_d$, the possible solutions are that spread that the question of the independence for this point with this efficiency deserves to be asked. In fact it is logical that $\eta_d$ has no influence on an on-design point because its only effect is to shift the critical point. It should therefore not be included in the parameters of the optimizer for this regime.

## 3.3 Conclusion

This chapter aims to answering the question if multiple combinations of efficiencies could be possible first when calibrating on an entire operating curve and secondly with a single operating point for the calibration of the mass flow and the entrainment ratio. It was found that multiple solutions can fit the curves and the operating point. In both cases, the differences between the sets have strongly suggested that there are indeed multiple possible solutions. Moreover, it appears that couple of efficiencies exist between $\eta_s$ and $\eta_{py}$ and between $\eta_d$ and $\eta_m$. This later result indicates that among these pairs, one should be fixed because they are related. By this way, the model could retrieve a number of efficiencies so as to calibrate it with a single solution.

Even if the question seems to be answered, what was studied in this chapter is far from being representative of the behaviour of the efficiencies in all the parameter space. Indeed the operating curve only represents a single operating condition. For the calibration with a single point, the off-design regime is not taken into account. Moreover, these efficiencies could be sensitive to the ejector geometry too. However, these single conditions allow to bring to light the fact that the calibration solution is not unique. the chosen conditions being not extraordinary, it can be expected that this phenomenon occurs for other regimes and even for most of them as seen with the numerous solutions for this one.

To conclude, with all these results, it can be assumed that multiple solutions exist for a single operating condition. This situation is not desirable for launching a machine learning model and has therefore to be resolved. This last issue will be the main topic of chapter 4.
However, the conclusion drawn from these results should not be taken as general given the specific operating conditions used for the process. Indeed, in this chapter, it was shown what was necessary to prove the multiplicity of solutions with a single example.
To go further, it could have been interesting to enlarge this process to many other conditions. One could propose to try this process on other operating points for example in off-design regime or simply with other pressures, mass-flows or temperatures. In this way, high dependencies or independence of efficiencies according to parameters in a certain region could have been discovered. For example, it is not unlikely that $\eta_d$ is independent from the operating conditions in the on-design regime. Even if this could bring more information about the behaviour of the model with efficiencies, this would exceed the goals of this chapter. Nevertheless, a more precise inspection of how the efficiencies behave with each other and with the other parameters will be carried out in the next chapter by a different approach. By this way, the next chapter will also investigate which of the efficiencies can be left out of the calibration in order to find a unique solution.

# Chapter 4

# Global Calibration

As mentioned in the previous chapter, the model, as it is, can not fit in a single way the value of the mass flows. This means that the five efficiencies are too many to characterise, in a unique and physical way, the actual outputs of the thermodynamic model. The system of the calibration problem is overdetermined. Therefore, one solution could be to modify the thermodynamic model itself with the purpose of solving this calibration problem.

In this chapter, An approach to simplify the model will be proposed in order to satisfy the criterion of having one set of efficiencies for one set of outputs.
First, the idea of removing an efficiency of the parameters or merging two efficiencies to give a single one will be tested. In this way, the model is simplified by reducing the number of parameters.
In the second part of this chapter, the way of adding the critical pressure in the output will be explored. Indeed, as reducing the number of parameters could be a solution to the problem, adding a new criterion to minimize can be a solution too. Moreover, it allows to be more accurate for the representation of the physics of the calculated points.

## 4.1 Setting and fusion of efficiencies

### 4.1.1 Methodology

For all this chapter, the experimental results used to calibrate the model are those from experiment that was conducted at the VKI on the DMU 03 ejector. It should be notice that the primary nozzle of the ejector is divergent. The diameters of the primary throat and exit of the primary nozzle are therefore the same and are equal to $D_e = D_{th} = 28$ $[mm]$. The diameter of the constant area section, as

for him, is equal to $D_y = 52$ $[mm]$. The data-set contains 134 operating points. There are, among them, several points with zero entrainment ratio and others belonging to 6 different operating curves. The optimizer that will be used is the same as that of the previous section. Namely the scipy optimizer with L-BFGS-B algorithm [33, 34].
Before starting any change with the efficiencies, it is necessary to determine a way to observe which efficiencies are the most changing and if there are correlations between them. By this way it will be possible to fix the value of an efficiency if it is observed that this one can often take a wide range of values for the same solution or if its impact is quite low. With this kind of method, it may also be possible to fix an efficiency if a correlation between efficiencies is observed. This would indicate that they have a similar impact on the mass flows and one of the two may be therefore useless.
The way of proceeding will be inspired from the Monte-Carlo method [38] by calibrating a point a large number of times from different initial guesses. An effective means to observe correlations and changing is by the calculation of the mean value and standard deviation of each efficiency for the point calibrated a large number of times.

To do so the following methodology will be used for each point of the data-set. First, a random value is assigned for each of the efficiencies between the following bounds :

$$\eta_p \in [0.25, 1.0] \quad \eta_s, \eta_{py}, \eta_d \in [0.0, 1.0] \quad \eta_m \in [0.9, 1.0]$$

The bounds for $\eta_p$ and $\eta_m$ are chosen such that they avoid a crash of the code which may be due to nonphysical flow or other cases like back-flow that are not handled by the model.
Once the efficiencies have been obtained randomly, they are included as the starting vector of the optimizer parameters. In this way, if different solutions exist for a point, the optimizer should, over the iterations, find them. Moreover, if a solution found by the optimizer has a value of the cost function above a threshold value depending on the number of efficiencies to calibrate, the computed efficiencies are not stocked and the iteration starts again. Nevertheless, a number of maximum re-iteration is taken into account to avoid infinite loop. The value of the threshold value could be increased after when there will be less efficiencies to calibrate. Finally, after 50 iterations, the means and standard deviation are computed from the obtained efficiencies. The amount of 50 was chosen in order to have a sufficiently numbered sample without taking too much computational time. A schematic of the process is drawn on figure 4.1.

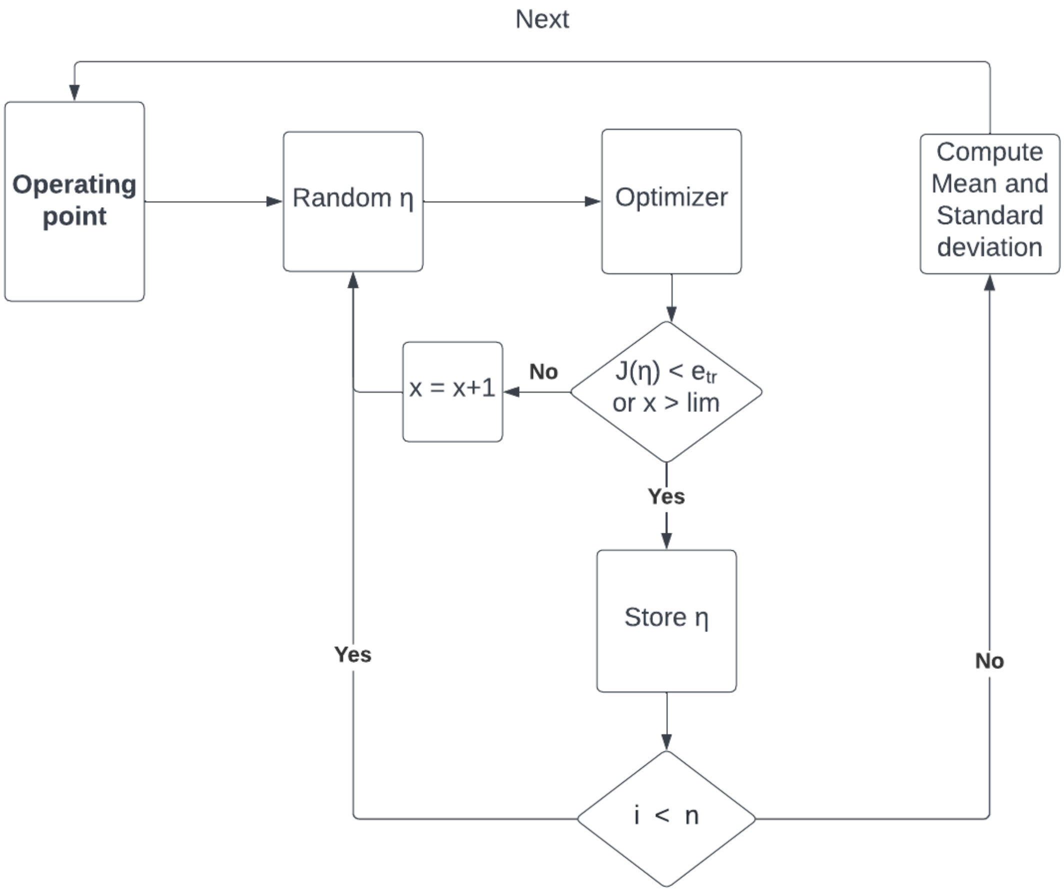


Figure 4.1: Diagram of the algorithm to obtain mean and standard deviation of efficiencies

The precision for a solution of the optimizer being determined by the value of the cost function, this one has therefore to be correctly defined according to what should be minimized. In this case, the desired effect is to obtain efficiencies such that the model fits the experimental points. The criterion to minimize is then the difference between the outputs of the model and the corresponding obtained experimental values. In other words, the goal is to minimize the error on both mass flow rates. The cost can thus be expressed as:

$$J(\eta_p;\eta_s;\eta_{py};\eta_m;\eta_d) = ||m_p^p - m_p||_2^2 + ||m_s^p - m_s||_2^2 \tag{4.1}$$

Where $m_p^p$ and $m_s^p$ are the values of the primary and secondary mass flows predicted by the model and $m_p$ and $m_s$ are the experimental ones. The norm is taken according to the L2 norm definition.

## 4.1.2 Results and comments

### 4.1.2.1 Calibration with five efficiencies

The first result to notice, before seeing the global result, is that for several points without entrainment, therefore with $\omega = 0$, solutions with a zero standard deviation are observed. This means that, for these operating conditions, the solution is well determined and the set of efficiencies is at each time of the following form :

$$\eta_p, \quad \eta_s = 0.0, \quad \eta_{py} = 1.0, \quad \eta_d = 1.0, \quad \eta_m = 1.0$$

With $\eta_p$ that can take different values depending on the value of the primary mass flow. From this, it can already be concluded that: if calibration for these points has to be done, one can not fix the value of the primary efficiency for these points. Moreover, if the value of one or the other efficiencies is fixed in what follows, it must be 0.0 for the secondary efficiency, which makes no physical sense, or 1.0 for the other ones.

The means and standard deviations for all points are represented on figure 4.2 and 4.3 respectively. These figures represent the evolution of the means and standard deviations for each parameter according to each other. These parameters include, among other, the state variables which are inputs of the thermodynamic model expressed in dimensionless ways, which are the two pressure ratios $\pi_{s,p}$ and $\pi_{d,p}$, between the secondary pressure and the outlet pressure compared to the primary one. After that, the correlation with the entrainment ratio $\omega$ is also expressed. Obviously, the five efficiencies are present on the right side of the tabular.

What appears on the upper left part of the figures concerning correlations between inputs with each other and inputs with output is not in the concerns of this chapter. It can nevertheless be noticed that there is no point in the left up part of the scatter plot between omega and the first pressure ratio. This shows the need to have bigger pressure ratio to achieve higher entrainment ratio and the compromise to take between this two performance indicators.
On figure 4.2. There is no clear trend observable between efficiencies and other variables. Nevertheless, a correlation seems to appear on the plot between $\eta_d$ against $\eta_m$. According to what the bounds allow, the values that can appear are wide for all the efficiencies except for $\eta_p$ which keeps high values as expected. $\eta_d$ and $\eta_{py}$ remain generally above 0.5 but $\eta_s$ and $\eta_m$ are completely scattered. The diagrams on the diagonal of the figure indicate the probability that a certain range of values will appear.

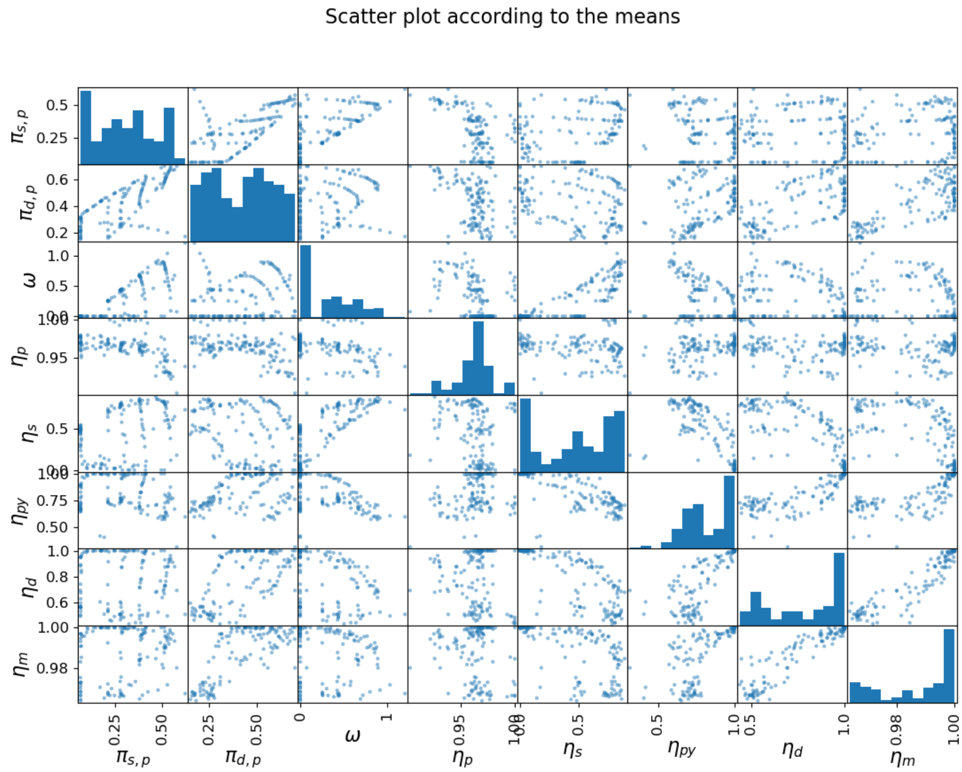


Figure 4.2: Correlations between the inputs of the model, the entrainment ratio and the means of the five efficiencies

On figure 4.3, the first observation to notice is the very high values of standard deviations that can take $\eta_s$, $\eta_{py}$ and $\eta_d$. Moreover, even $\eta_m$, which is constrained between 0.9 and 1.0, has standard deviation that is, for most of the operating points, between 0.01 and 0.02. Yet, the standard deviation of $\eta_p$ is more acceptable. The calibration of this efficiency seems moreover quite independent from the other. This is certainly because of the fewer variables that are involved in its definition2.

Recalling the question of the introduction, it appears now that, based on the latter result the parameters of the model are too numerous. It is therefore expected that the mass flows found by the model will fit well (even to well) the mass flows of the data-set. Figure 4.4 and 4.5 draw the differences, for primary and secondary mass flows respectively, between the value of the data-set (ordinate) and computed by the model. The difference is plotted for each of the 50 iterations of each point of the data-set. The black line represents the value for which the model would perfectly match the data. The dotted lines represent a ten percent error of the model compared to the data.

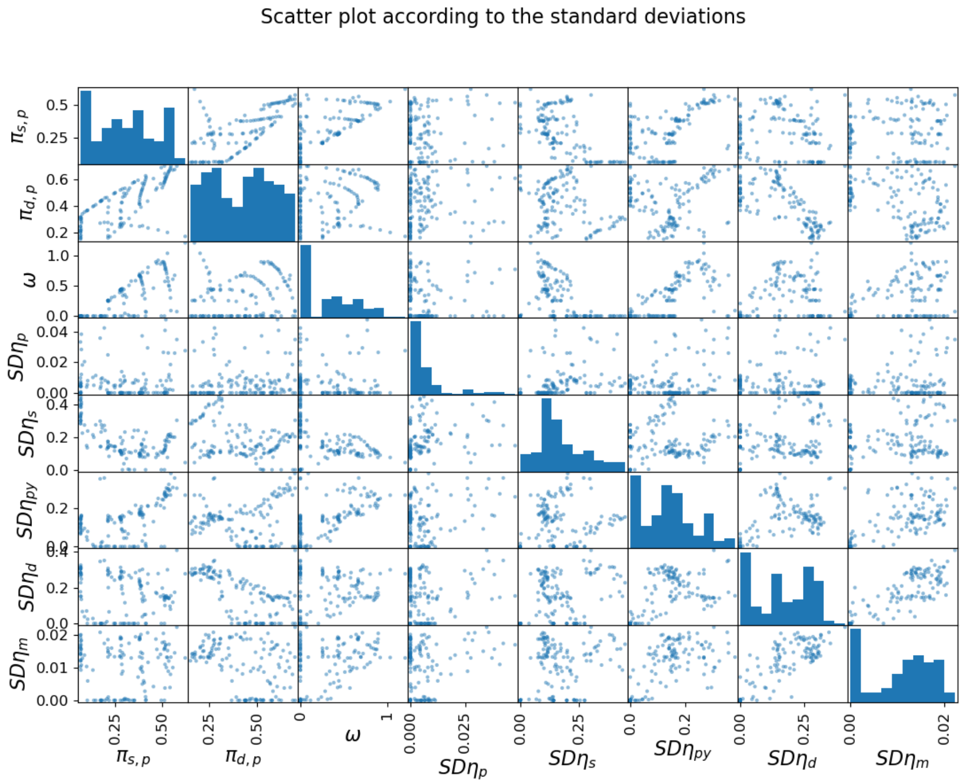


Figure 4.3: Correlations between the inputs of the model, the entrainment ratio and the standard deviations of the five efficiencies obtained by optimization point per point (see section 4.1.1)

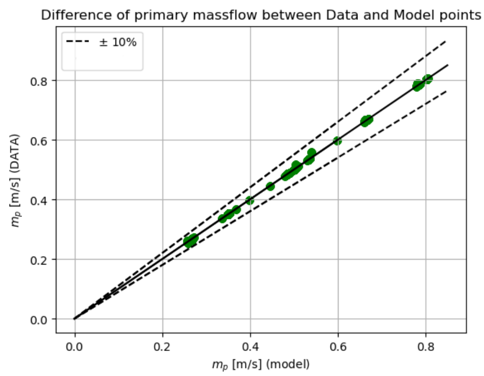


Figure 4.4: Error on the primary mass flow

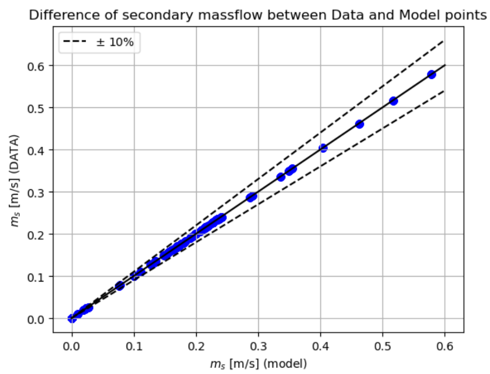


Figure 4.5: Error on the secondary mass flow

As expected, there is a very good fit when the optimizer uses the five efficiencies. It can be noticed that the primary mass flows are a little less well estimated compared to the nearly perfect fit for the secondary flow. However, in both plots, these very precise results suggest that the solutions are overfitted. Overfitting is dangerous because it makes a solution so suited to a problem that trying to use it by extension for another problem will lead to very bad results. In fact, the points are well estimated but not what is nearby them. It is absolutely to be avoided to make the machine learning methods usable.

#### 4.1.2.2 Calibration with 4 and 3 efficiencies

From the results of the previous subsection, the question is now to choose a way to reduce the number of parameters for the optimizer.
For this purpose, one of the efficiencies has to be fixed. As seen in the previous section, this efficiency can not be the primary one and the other must be fixed to a certain value. Therefore, the value of the secondary efficiency cannot be fixed because this would impose a zero secondary flow no matter the other conditions. The question is now to choose which efficiency will be set to as equal one.
It appears that $\eta_{py}$ has a lower impact on the outputs [32] than $\eta_d$ or $\eta_m$, it is physically less important than them. Moreover no correlations with inputs or any other variable clearly appear. On top of that, the mean values of $\eta_{py}$ are quite high. It should therefore not be a problem to set the value of this efficiency to 1.0 considering its high standard deviation as well.

Figure 4.6 and 4.7 show the scatter plots of the means and standard deviations with $\eta_{py}$ fixed to one. The values of the standard deviations have generally reduced. However, they are still unacceptably high. Especially with $\eta_s$ and $\eta_p$ for which the standard deviation can often achieve more than 0.2. This means that the parameters are still too numerous and the setting of $\eta_{py}$ is not sufficient to obtain a single set of efficiency for each point.

The errors with four efficiencies are visible on figure 4.8 and 4.9 for the primary and secondary mass flows respectively. As it can be seen, errors are already bigger for some points. They sometimes even exceed 10% but remain even so generally small. Nonetheless, the secondary mass flow near zero is not well determined. This means that $\eta_{py}$ has an important role for low secondary mass flow since there was no such observation with five efficiencies.

The global reduction of standard deviation lets appear clearer correlations between

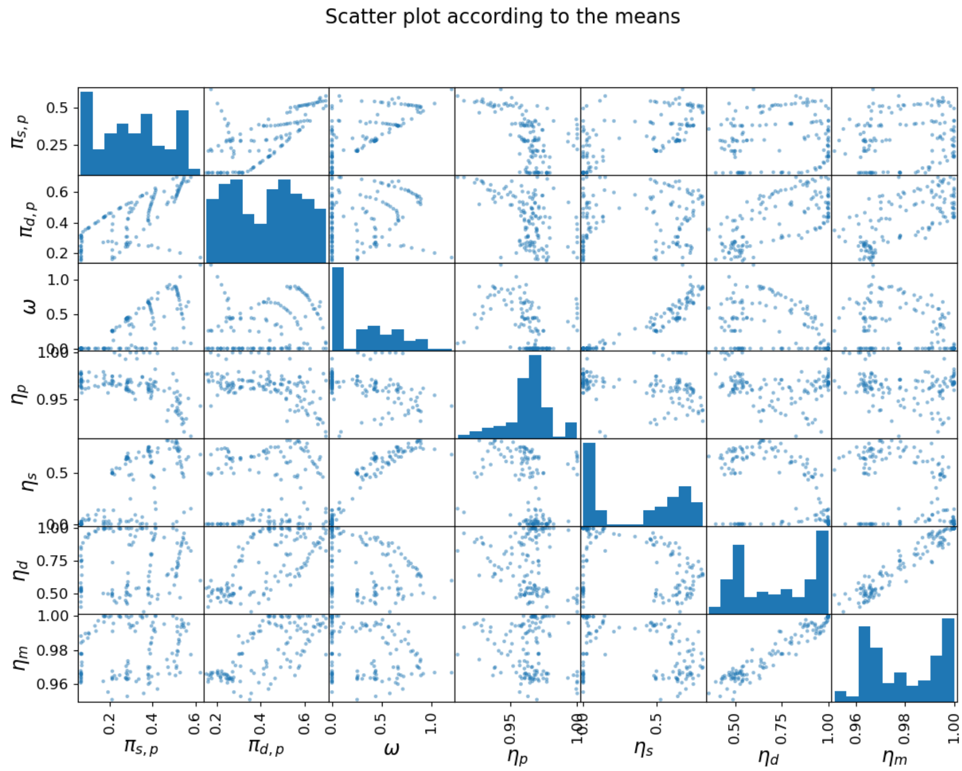


Figure 4.6: Correlations between the inputs of the model, the entrainment ratio and the means of the four efficiencies

some variables. First, it becomes clearer that a correlation between $\eta_d$ and $\eta_m$ exists. As seen from the shape of the correlation, it would be a linear positive correlation. Thanks to this, one efficiency could be able to be expressed in terms of the other one. It also means that, if one of both is defined as having a value equal to 1.0, the remaining should be able to take the role of both. In this way, one more parameter of the optimizer can be removed. Because $\eta_d$ has the biggest standard deviation and deleting $\eta_{py}$ has not reduced it significantly, it will be defined as being equal to one.

With more efficiencies being equal to 1.0, the bounds of $\eta_m$ can now be extended to:

$$\eta_m \in [0.8, 1.0]$$

The results of the scatter plots with 3 efficiencies as parameter of the optimizer are displayed on figure 4.10 and 4.11. As can be seen on figure 4.11, the standard

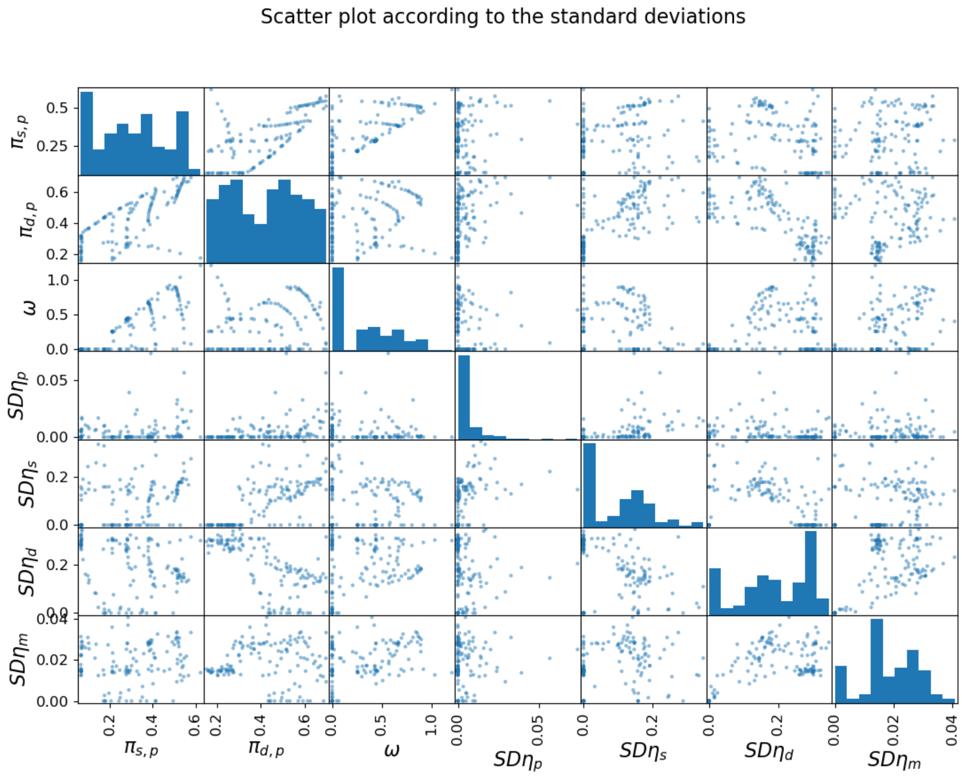


Figure 4.7: Correlations between the inputs of the model, the entrainment ratio and the standard deviations of the four efficiencies

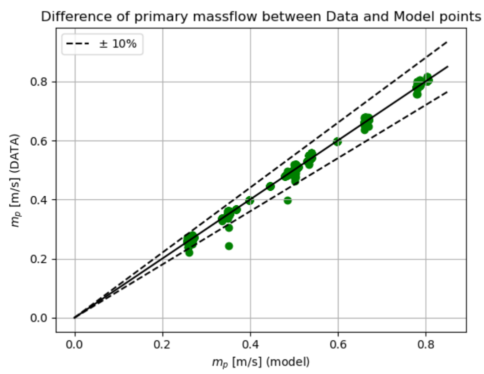


Figure 4.8: Error on the primary mass flow

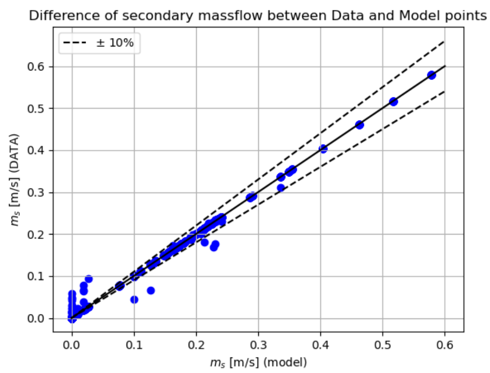


Figure 4.9: Error on the secondary mass flow

deviations, although quite reduced for $\eta_s$, are still not significantly reduced. On the contrary, they are sometimes higher with 3 efficiencies than with 4 for $\eta_p$ and $\eta_m$.

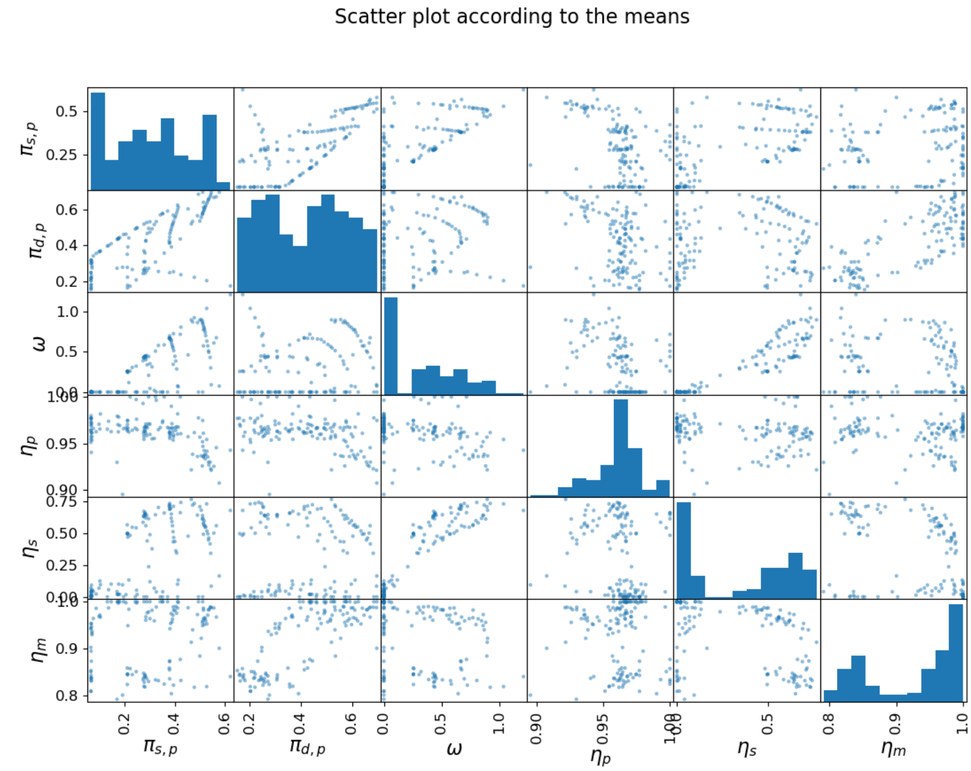


Figure 4.10: Correlations between the inputs of the model, the entrainment ratio and the means of the three efficiencies

On the means scatter plot of figure 4.10, new values of means between 0.8 and 0.9 are appearing for $\eta_m$ due to its wider range. Moreover, some correlation between $\eta_s$ and $\omega$ is apparent. This correlation will be explored in more detail in what follows.

As can be seen on the error representation in figure 4.12 and 4.13. The error on the secondary mass flow rate has not significantly changed. In contrast, the primary mass flow has lost a lot in precision. Without the diffuser efficiency, the solution found by the optimizer with the model often overestimate the primary mass flow which is quite surprising given that $\eta_p$ should be sufficient.

With this last experiment, it seems that the enlargement of the mixing efficiency bounds has compensated, in a way, the deletion of the diffuser efficiency by allowing more solutions with its new range. However this new range has maintained quite low error for secondary mass flow but it did not avoid the increasing primary mass flow error. Nevertheless, the standard deviation remains too high. This means that this approach is not sufficient to obtain a single solution for each operating

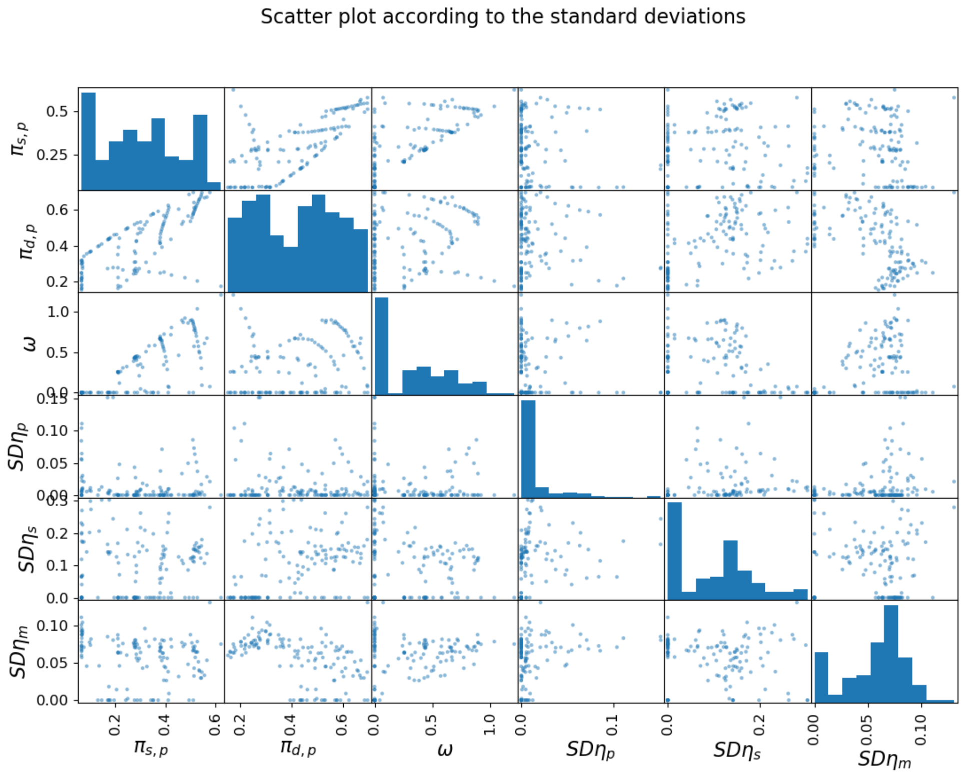


Figure 4.11: Correlations between the inputs of the model, the entrainment ratio and the standard deviations of the three efficiencies

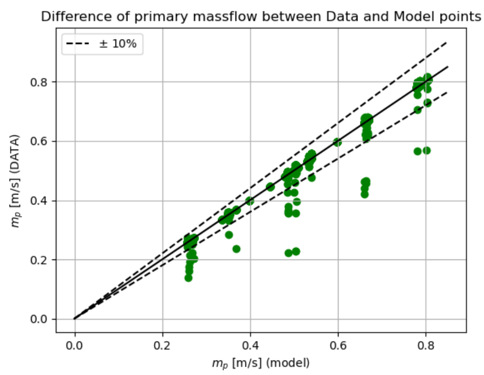


Figure 4.12: Error on the primary mass flow

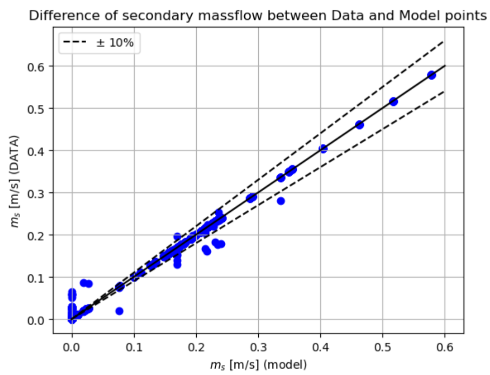


Figure 4.13: Error on the secondary mass flow

condition.
In what follows, a slightly different way to proceed will be explored.

#### 4.1.2.3 Calibration with two efficiencies

In the above, with the optimization with five parameters. It was first mentioned that for some points for which $\omega = 0$, $\eta_p$ was the sole efficiency to determine the primary mass flow. This seems consistent given the parameters determining the primary mass flow, as can be seen in the model of Metsue et al. [1]. However, it does not mean that the primary efficiency cannot impact the secondary mass flow. In addition to that, some correlation has been observed between $\eta_s$ and $\omega$. With $\eta_p$ and $m_p$ determined, it would mean that only $\eta_s$ could be necessary to find the remaining output which is the secondary mass flow rate. Moreover $\eta_s$ should not impact the primary mass flow rate and therefore $\eta_p$.

In what follows, the way to proceed will therefore be to find the primary mass flow by the calibration of $\eta_p$ and nothing else. After that, the best fitting value of $\eta_p$ for each point will be taken back and the secondary mass flow rate will be calibrated with only the secondary efficiency. When this process is expressed as an optimization problem, it gives :

$$\min_{\eta_p} J_1(\eta_p) = ||m_p^p - m_p||^2 \tag{4.2}$$

$$\min_{\eta_s} J_2(\eta_s; \eta_p) = ||m_s^p - m_s||^2 \tag{4.3}$$

As shown in figure 4.15, this method brought down the standard deviations. It becomes negligible for all the operating conditions with $\eta_p$ and with most of them for $\eta_s$. Several points have nevertheless still large standard deviation with the secondary efficiency. Even if there are only a few such points, this means that two values of $\eta_s$ are allowed to give the same secondary mass flow rate. This calls for reflection on the question of whether the mass flow rates are sufficient to characterise the entire flow on the ejector for a given set of inputs. This subject will be the main concern of the following section.

In the mean scatter plot (figure 4.14), the values of the primary efficiency are, for most of the points, between 0.95 and 0.97. These values decrease a little when pressure ratios take higher values. Higher compression ratio often corresponds to off-design application. Physically speaking, the values of $\eta_p$ correspond well to what is expected. However, with regard to the secondary efficiency, there are, at first, a bit less correlations with the entrainment ratio. On the other hand, The value of $\eta_s$ takes very often the value 0.0. A zero efficiency is physically meaningless. This raises another question that will also be discussed in the next section. This zero points are each time corresponding to zero entrainment ratio

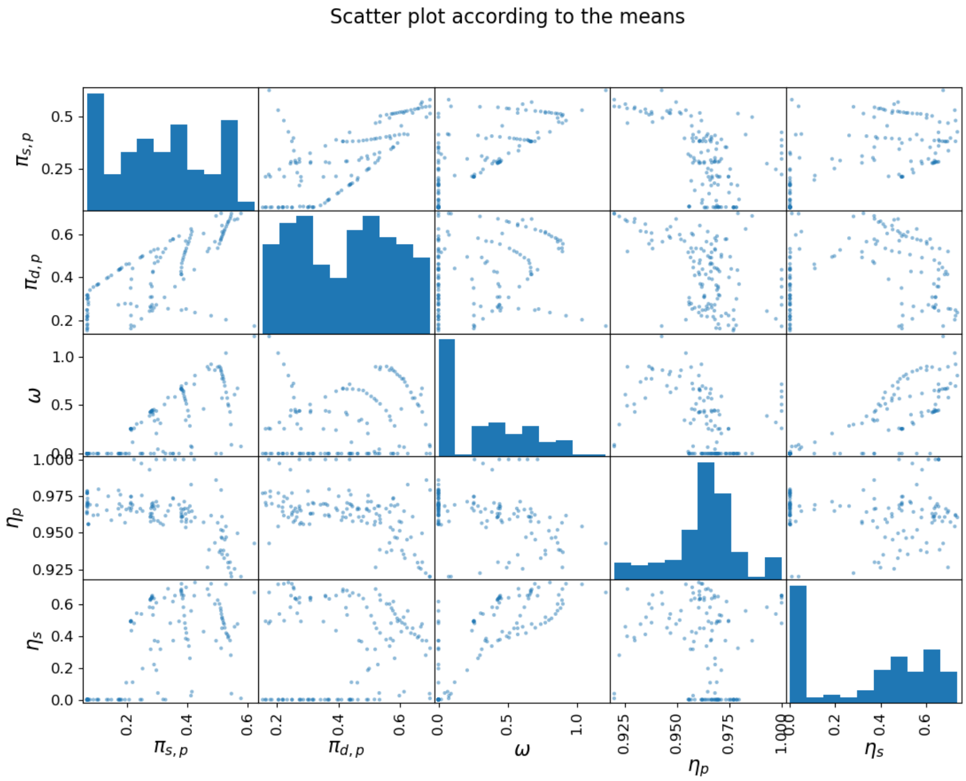


Figure 4.14: Correlations between the inputs of the model, the entrainment ratio and the means of the two efficiencies

points but there are also $\omega = 0$ points for which the secondary efficiency is different from zeros. Nevertheless, it should be remembered that the thermodynamic model may cope limitations with the nil value of the entrainment ratio. In view of the very low values of $\eta_s$, it can be assumed that the losses that the three other deleted efficiencies should take into account are taken into account by the $\eta_s$. Finally an extension of the model would likely cause improved calibration results for omega close to zero.

The accuracy of the model with the optimizer to find the primary mass flow is depicted on figure 4.16. In addition to causing almost no standard deviation, the process to find $m_s$ with only $\eta_p$ is remarkably effective. Moreover this accuracy proves again that the secondary efficiency has no influence on the primary mass flow. As for the secondary efficiency (figure 4.17), even if the accuracy is not as good as for the first one, the process demonstrates more accuracy compared to 3 or 4 efficiencies. There are still a little more errors with some zero entrainment ratio points. The better accuracy appear probably because the calibration is split in two

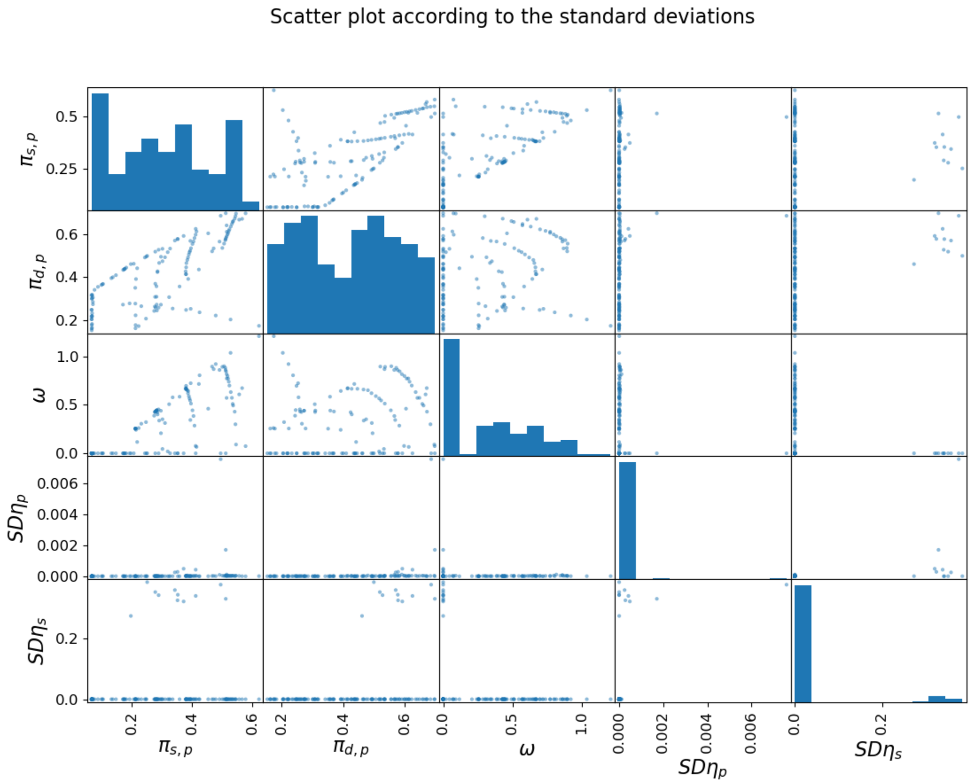


Figure 4.15: Correlations between the inputs of the model, the entrainment ratio and the standard deviations of the two efficiencies

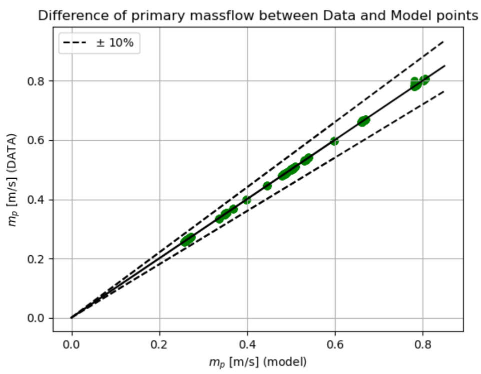


Figure 4.16: Error on the primary mass flow

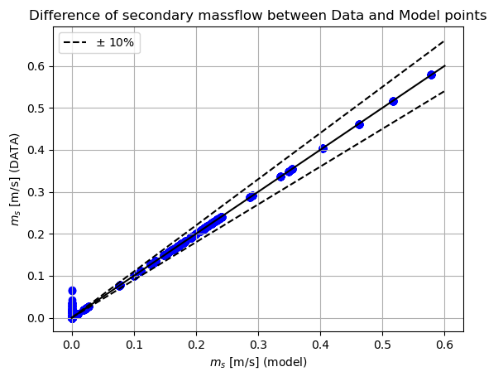


Figure 4.17: Error on the secondary mass flow

easier optimization problems in a lower dimensional space

To conclude, there is no doubt that the process with two efficiencies as input parameter of the optimizer is the best. However, it is probably not sufficient to launch a machine learning model. Indeed, there are still issues and questions concerning the results obtained with the secondary efficiency. The fact that only two efficiencies are sufficient to find the output is also questionable. The following section tries to give some answers to this.

## 4.2 Addition of outputs

According to what has been discovered in the previous section, a question worth considering :
For given inlet and outlet total pressures, inlet static temperature, geometry of the ejector and the 5 efficiencies, are the two mass flows sufficient to characterise an operating point of an air ejector ?
When analysing the impact of each efficiency on an operating curve, it appears that the diffuser and mixing efficiencies have no impact on the mass flows in the on-design regime [1], which makes sense given that both flows are chocked in this regime. This means that these two efficiencies are irrelevant in the on-design regime if the outputs of the model are only the mass flows. Moreover, in the previous section results, it appears that, for low value of $\omega$, the optimizer compensates by a low value of $\eta_s$. Although that result remains physically correct, it should not be a low efficiency that makes the entrainment ratio decrease but the increasing outlet pressure. The meaning of the solutions of the previous section therefore appears clearly. For each point in the on-design regime, only the primary and secondary efficiencies can be needed to find the mass flow rates. Therefore, for each set of mass flow rate, an on-design solution is found with the two first efficiencies. From this, it can be assumed that another off-design solution exists for these points that cannot be differentiated with the mass flow rate as sole outputs. As illustrated on figure 4.18 and considering constant primary mass flow, many sets of efficiencies corresponding to one set of mass flows can be found for points belonging to many operating curves. One point is on-design, and the others are off-design. This means that, although the latter results from the previous section globally give the correct results according to mass flows rate, they can be completely wrong locally or even at a different operating regime.

### 4.2.1 Critical pressure as new output

Given that all points of the data-set are not on-design, an additional variable can be added in the outputs of the model. This output has to be an indicator of the operating regime.

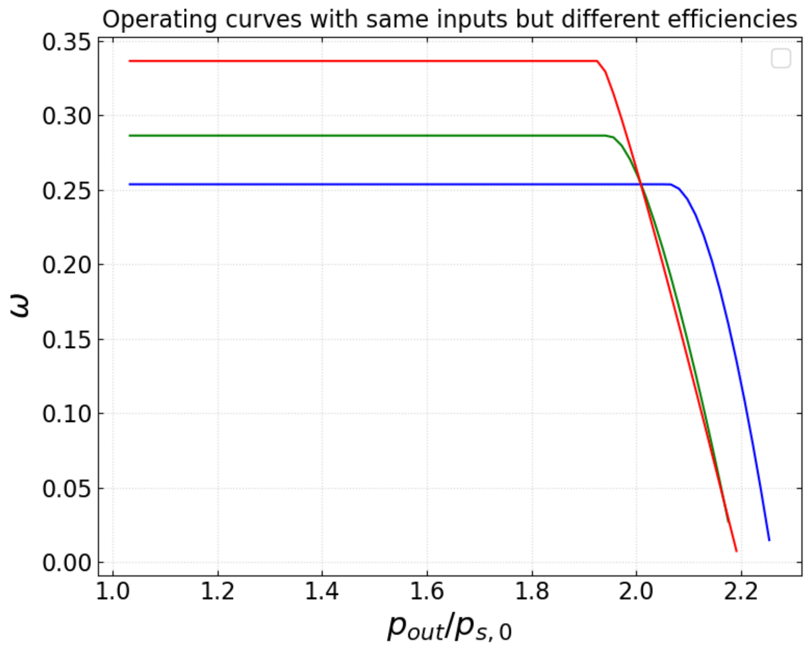


Figure 4.18: Operating curves with same inputs but various efficiencies, the crossing point has the same mass flow rates but different operating regimes

Considering the information the data and the model can provide, including the critical pressure as output is a choice allowing to obtain more information about the whole operating curve to which the point belongs(at least if it is on- or off-design) This solution keeps simple calculation for the model and with the using of data.

As the critical pressure is already computed into the model of Mestue et al. [1], it does not brings any problem to add it in the outputs. However, the purpose is to be able to add the difference between the critical pressure computed by the model and the real one in the minimisation cost function of the optimizer. This leads to an issue which is that the critical pressure is, for most cases, not provided with the experimental data-sets.

In this case, the data from VKI allow nevertheless to obtain an estimation of the critical pressure. The set contains 6 operating curves composed of 63 points. It is therefore workable to extract a critical pressure for each of the 6 points by selecting manually the point from which the curve seems to decrease for more than 10% according to the back pressure. The figure 4.19 illustrates the process for one of those curves.

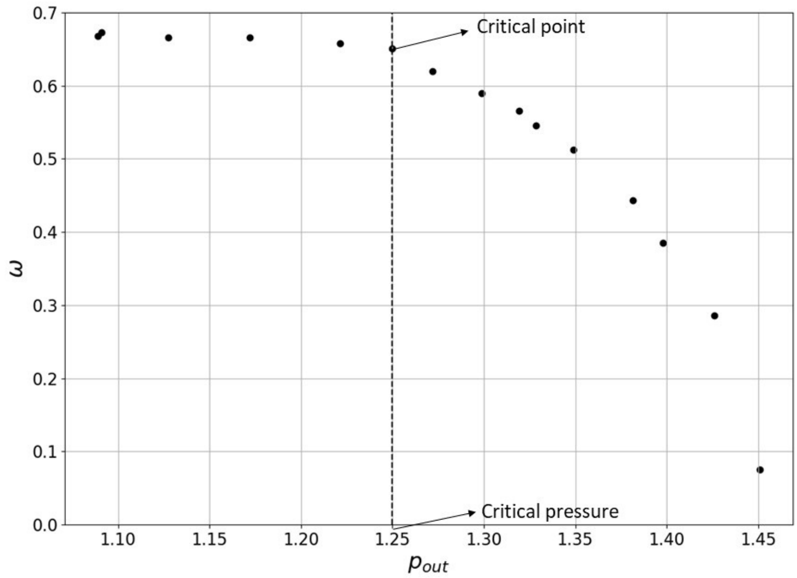


Figure 4.19: Data points belonging to the same operating curve with the sixth being the critical point

Thereby, the additional criterion to minimize for each points will be written as follows :

$$E_c = ||p_c^p - p_c|| \tag{4.4}$$

Where $p_c^p$ is the critical pressure computed by the model and $p_c$ the real one extracted from the data.
Three ways will be explored for the calibration with this new criterion.
First, as for the last case of the previous section, the calibration with two efficiencies will be tested. It will verify if the new criterion is relevant and if the calibration with two efficiencies is still feasible.
The calibration of the primary flow with the primary efficiency was sufficiently efficient to be kept it as it was. The secondary efficiency will therefore be used for the calibration of the secondary mass flow rate and the critical pressure. The minimization problem system can be written as follow:

$$\min_{\eta_p} J_1(\eta_p) = ||m_p^p - m_p|| \tag{4.5}$$

$$\min_{\eta_s} J_2(\eta_s; \eta_p) = ||m_s^p - m_s||/m_s + ||p_c^p - p_c||/10^5 \tag{4.6}$$

With $J_1$ and $J_2$ the two cost functions. The pressure criterion is divided by 1 bar and the secondary mass flow rate by the mass flow from the data-set in order to non-dimensionalyse and keep the same order of magnitude between both criterions. Dividing by 0 can no longer cause an indeterminacy with this data-set given that there is no longer $m_s = 0$ point within. Nevertheless, the small $m_s$ points have more weight in the cost function compared to the critical pressure.
The second minimization choice will take into account one more efficiency. Since the main effect of the diffuser efficiency is the variation of the critical pressure without changing significantly the other parameters [32], it should be helpful to include it in the calibration. The calibration of the primary flow with the primary efficiency is kept. The minimization problem system is therefore:

$$\min_{\eta_p} J_1(\eta_p) = ||m_p^p - m_p|| \tag{4.7}$$

$$\min_{\eta_s,\eta_d,\eta_p} J_2(\eta_s;\eta_d;\eta_p) = ||m_s^p - m_s||/m_s + ||p_c^p - p_c||/10^5 \tag{4.8}$$

In the third approach, the three criterions will be minimized with the three efficiencies together. This will allow to observe if the primary efficiency can add precision to the calibration of the critical pressure by losing a part of it for the calibration of the primary mass flow rate. The minimization problem system thus becomes:

$$\min_{\eta_p,\eta_s,\eta_d} J(\eta_p;\eta_s;\eta_d) = ||m_p^p - m_p||/m_p + ||m_s^p - m_s||/m_s + ||p_c^p - p_c||/10^5 \tag{4.9}$$

The following subsection presents the results for each of the three approaches.

### 4.2.2 Results and comments

When only the points belonging to operating curves remain, the Data-set is reduce to 63 points and there are no longer zero entrainment ratio points.

#### 4.2.2.1 Calibration with efficiencies calibrated separately

Figures 4.20 and 4.21 show the scatter plots for the means and standard deviation with two efficiencies calibrated this time with the critical pressure as output. As it could have been expected, the standard deviations for the two efficiencies are again reduced. Moreover, they take such low values that they could be assumed to be numerical oscillations of the optimizer. On the other hand, except for a slight diminution of correlations, no significant changes are visible for the means compared to figure 4.14 for which critical pressure was not considered.

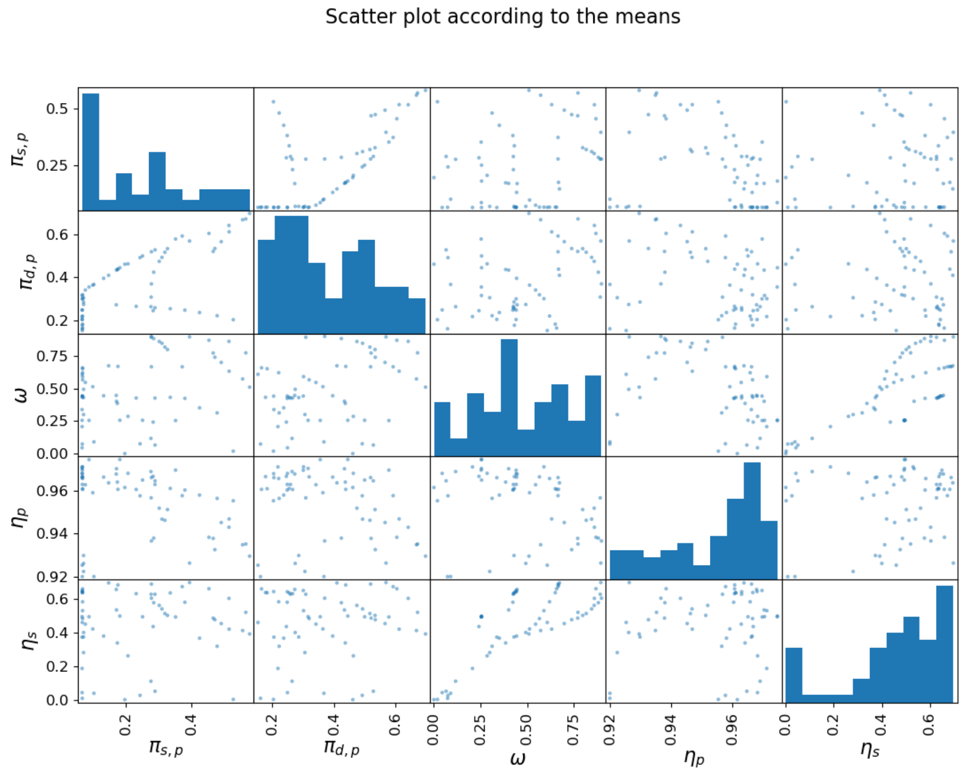


Figure 4.20: Correlations between the inputs of the model, the entrainment ratio and the means of the two efficiencies

On the other side, some results are more interesting to observe on the accuracy plots. The figures 4.22, 4.23 and 4.24 illustrate the error between the experimental value from the data-set and the calibrated value for the primary mass flow rate, the secondary one and the critical pressure respectively. For the mass flows, the calibration is almost exact. The error is as small as for the case of figures 4.16 and 4.17 with the difference that the less accurate points have just been deleted from the data-set because they were those for which $\omega = 0$. Figure 4.25 gives a better illustration of what represents such an error on critical pressure by comparing it with operating curves. On those whole curves, the inlet conditions and efficiencies from the chosen calibrated point are kept.
The critical pressure however is less accurate. The more the model deals with higher critical pressures, the more it tends to overestimate them. As seen from the difference of accuracy between the pressure and the mass flow rates, one could propose to put more importance on the minimization criterion of the pressure compared to the mass flow rate in order to balance the errors. Yet, this has given exactly the same results with 10 times more weight for the critical pressure. On

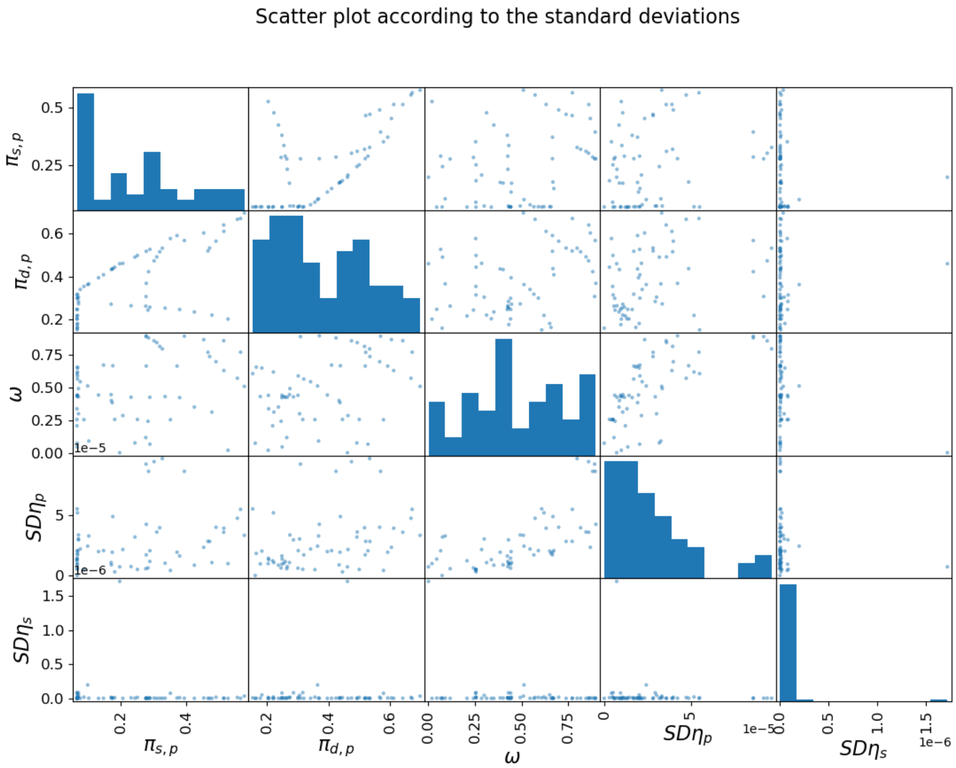


Figure 4.21: Correlations between the inputs of the model, the entrainment ratio and the standard deviations of the two efficiencies

the other hand, the critical pressure, because of the relatively poor accurate way it is obtained, should not require very high accuracy. However, the scatter plot of the means (figure 4.22) does not give enough correlations to allow to launch a machine learning model with this method.

By the results of the mean scatter-plot and the accuracy for the mass flow rate compared with the results without considering critical pressure 4.1.2.3, it can be assumed that the optimizer behaves almost identically as if the outputs were not containing the critical pressure. In this way, the critical pressure is not really optimized. Indeed, the value of the standard deviation in figure 4.15 was already too small to let to the optimizer enough room for maneuver. The only additional choice the optimizer could make was to select better the efficiency for the several points which still have high standard deviation for $\eta_s$ on figure 4.15. It could indeed be expected that two calibration parameters would not be sufficient to calibrate 3 different outputs.

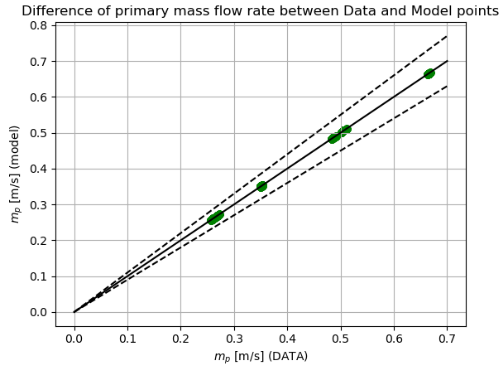


Figure 4.22: Error on the primary mass flow

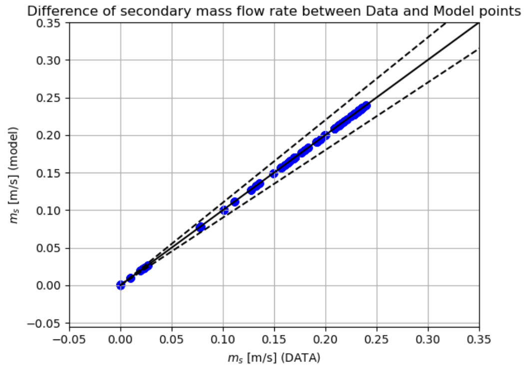


Figure 4.23: Error on the secondary mass flow

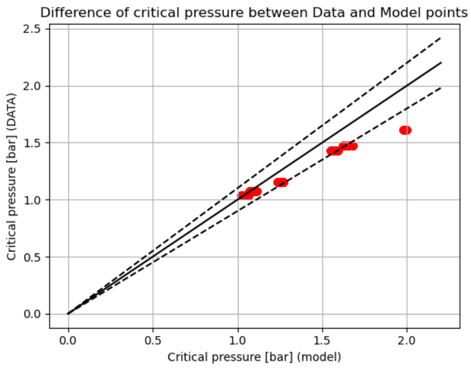


Figure 4.24: Error on the critical pressure

#### 4.2.2.2 Calibration with three efficiencies with primary mass flow rate calibrated separately

The main result to observe on figure 4.26 and 4.27 is the raising of the standard deviation for $\eta_s$ as well as the numerous points for which the deviation is also high for the diffuser efficiency. Obviously, because the same process is used to find it, all the results concerning the primary efficiency are the same as just above. As for the means, there is no trend or correlation visible enough to draw new conclusions. One could try the calibration with $\eta_{py}$ or $\eta_m$ but this does not give more satisfying results. This has led to the idea of the minimization including the three efficiencies together.

Figures 4.35, 4.36 and 4.37 depict the accuracy of the calibration using 3 efficiencies with $\eta_p$ found by the calibration of only the primary mass flow rate. As can be seen,

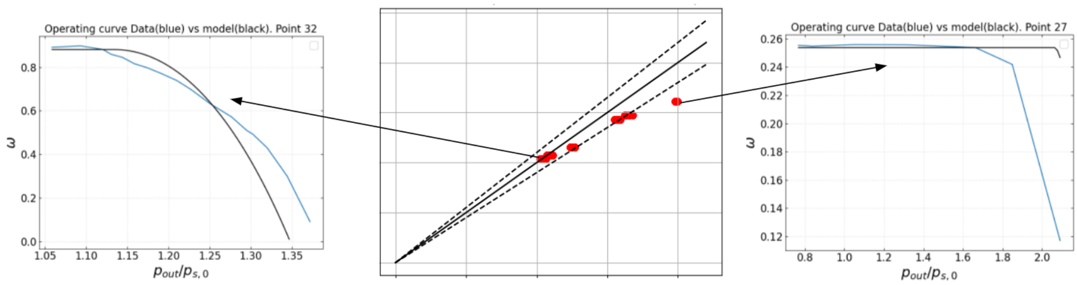


Figure 4.25: Error on the critical pressure represented by 2 operating curves for 1 point. The black one(smooth) is the model and the blue(broken) is made up with data-points

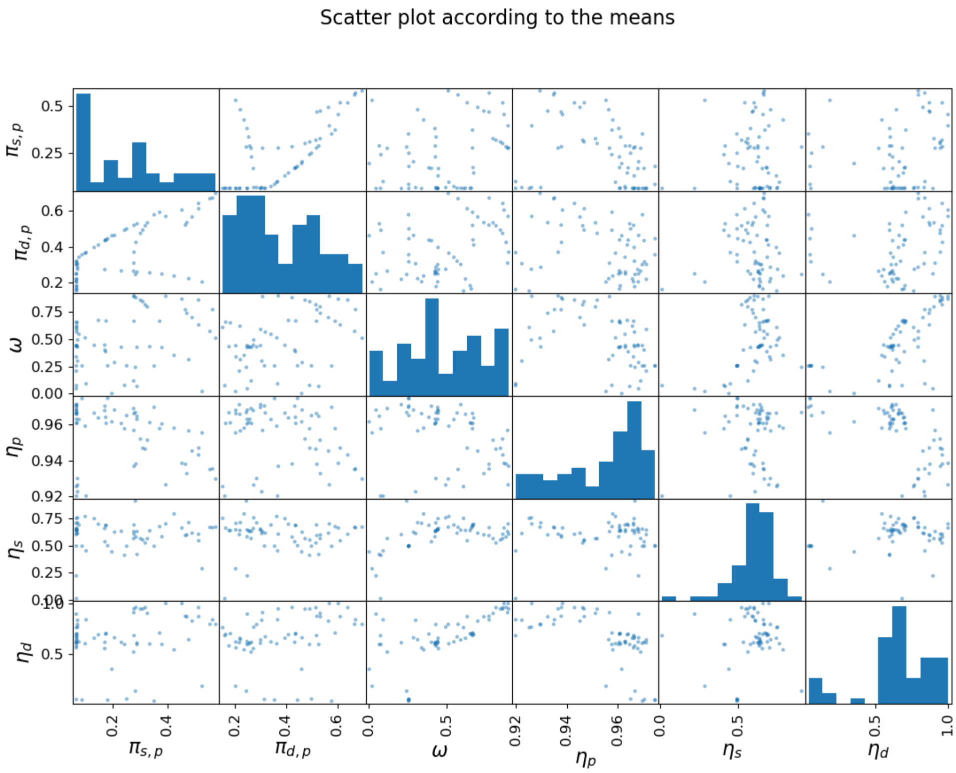


Figure 4.26: Correlations between the inputs of the model, the entrainment ratio and the means of the three efficiencies

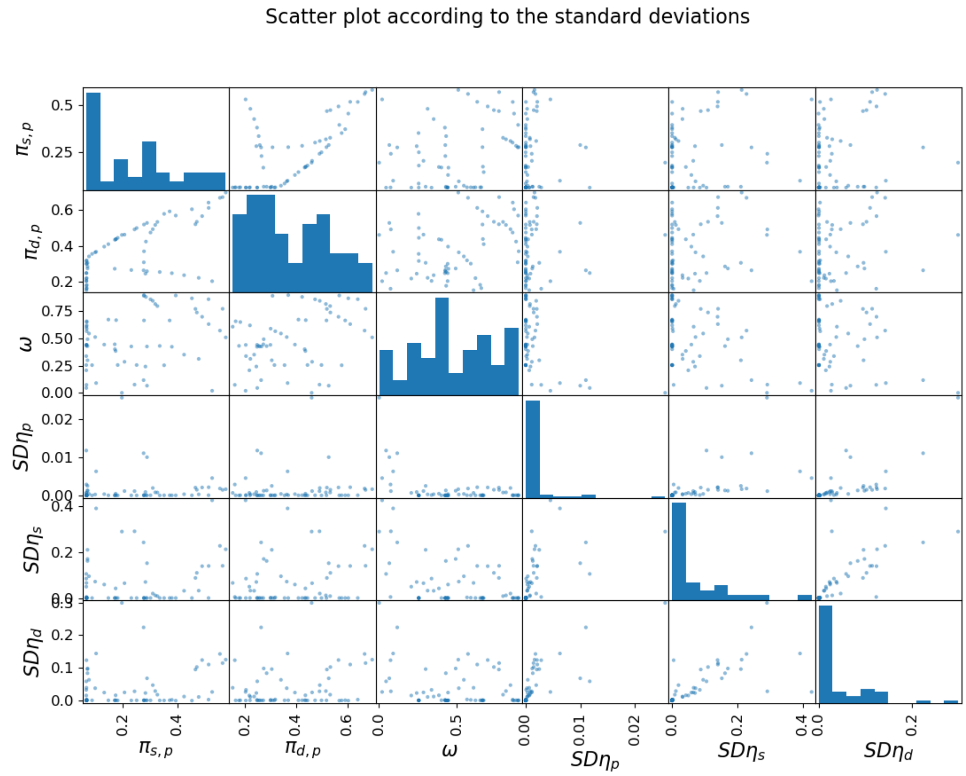


Figure 4.27: Correlations between the inputs of the model, the entrainment ratio and the standard deviations of the three efficiencies

the accuracy of the primary and secondary mass flows remain very satisfying. On the other hand, the higher critical pressure are better estimated than just above. However, the model completely overestimates the critical pressure for one entire operating curve. On the other hand, the point for which the model has returned a zero critical pressure is assumed to be an outlier. These results demonstrate that the diffuser efficiency may have a significant influence on the critical pressure. However, it can not handle all the cases as seen from the overestimate set of points.

#### 4.2.2.3 Calibration with three efficiencies calibrated together

As for the means and standard deviation scatter plot on figures 4.31 and 4.32, the results are not much different as just above with the exception that the standard deviations have slightly decreased. They remain however too high for a lot of points. Nevertheless, after inspecting the characteristic of the high standard deviation points, it appears that all these points are in off-design regime. This can be seen

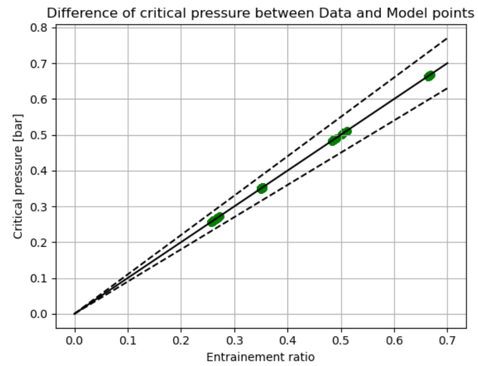


Figure 4.28: Error on the primary mass flow

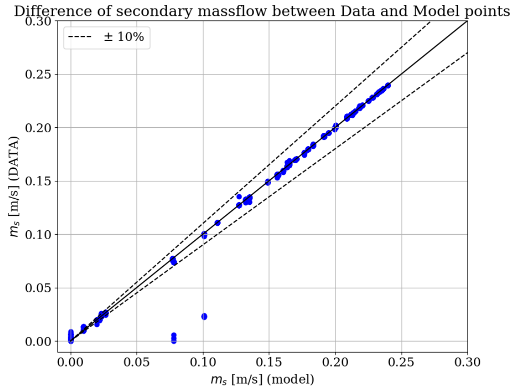


Figure 4.29: Error on the secondary mass flow

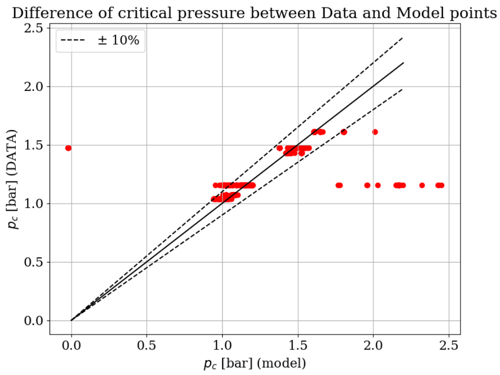


Figure 4.30: Error on the critical pressure

on figures 4.33 and 4.34 which shows the standard deviation for on- and off-design points separately. The on-design (green points) points have negligible standard deviations for $\eta_p$ and $\eta_s$ and it is very small for $\eta_d$. This means that these points are effectively well defined with three efficiencies. On the other hand, the number of possible sets of efficiencies raises for off-design points. This could come from the fact that, even with similar critical pressure, a point may belong to different operating curves in the off-design part.

However the means plot does not make appear clear correlations between variables. This may be due to the higher standard deviations but the fact is that it does not encourage creation of a machine learning process to find efficiencies. Without clear

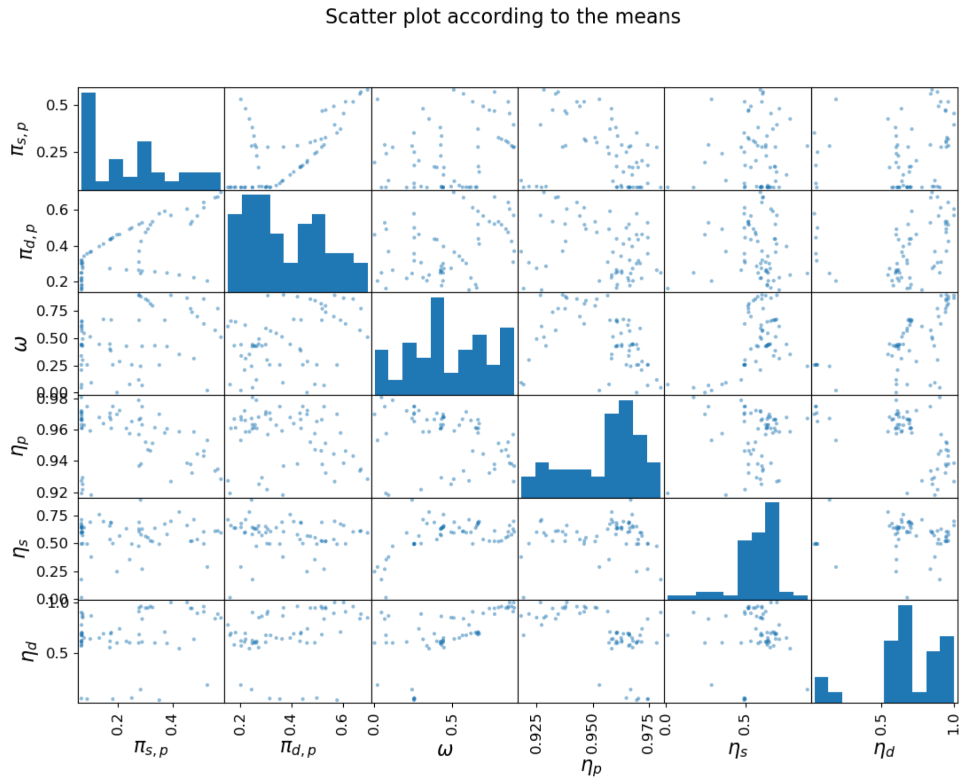


Figure 4.31: Correlations between the inputs of the model, the entrainment ratio and the means of the three efficiencies

trends between inputs and efficiencies, a neural network can not estimate well an output. However, this does not mean that the parameters of the scatter plot are indeed correlated. Is is quite possible that such correlations do not exist at all. Nevertheless, this does not prove that the training of a machine learning model is not possible. Trends between data could appear in a way or an other by playing with the way the calibration is made.

The error plots of figures 4.35, 4.36 and 4.37 are, for their part, quite different when the three efficiencies are calibrated together. There are a little more error on the primary efficiency but it remains very accurate. It is however less the case fore the secondary mass flow which has loss a lot of accuracy especially for several points. The critical pressure is, for its part, calibrated with more accuracy. Moreover, the overestimated curve and outlier point are no longer present. It seems that, by putting all the efficiencies together for the calibration of all the outputs with critical pressure, the optimizer have better balance the error of calibration of the output by keeping globally the same results for the scatter plots.

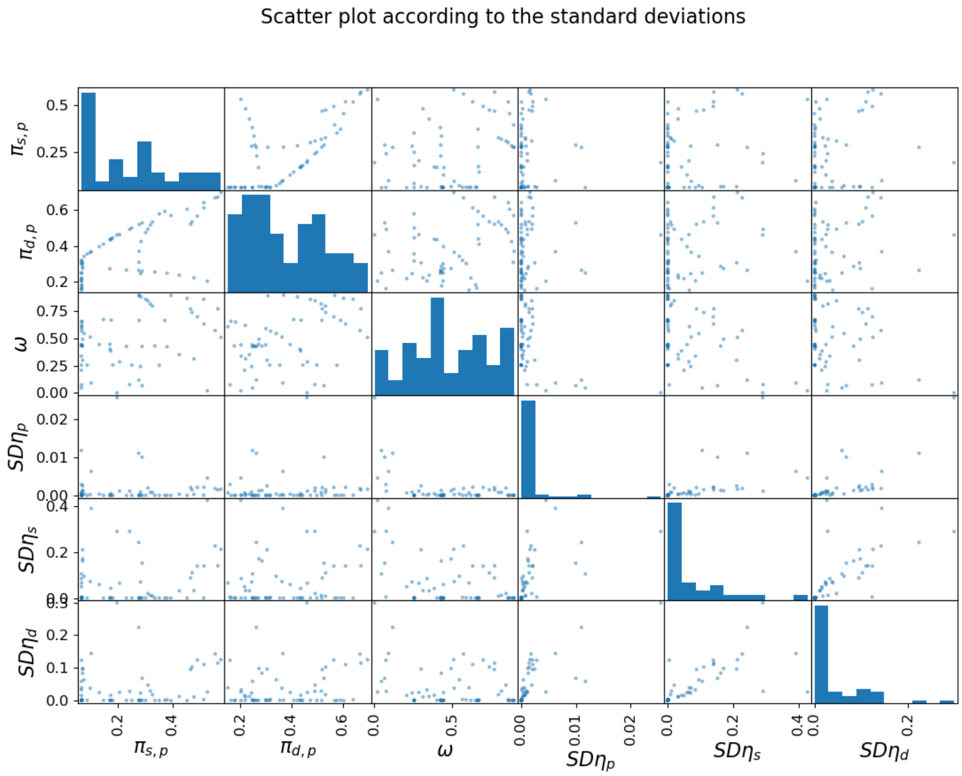


Figure 4.32: Correlations between the inputs of the model, the entrainment ratio and the standard deviations of the three efficiencies

## 4.3 Conclusion

During this chapter, some calibration methods have shown better results than others. First, the outputs of the model were kept intact and the number of efficiencies have been strategically reduced. During this section, one calibration gave better results than the others: the one using only the primary and the secondary efficiencies to calibrate separately the primary and secondary mass flow rates respectively. In this context, most of the points were very accurately calibrated with a single set of efficiencies. However, a major problem arose from these results: the two mass flows were not sufficient as outputs, because there was no guarantee that a point found by the model belonged to the correct operating regime.

To remedy this, the critical pressure has been added in the outputs. With this supplementary output, the calibration with primary and secondary efficiencies still

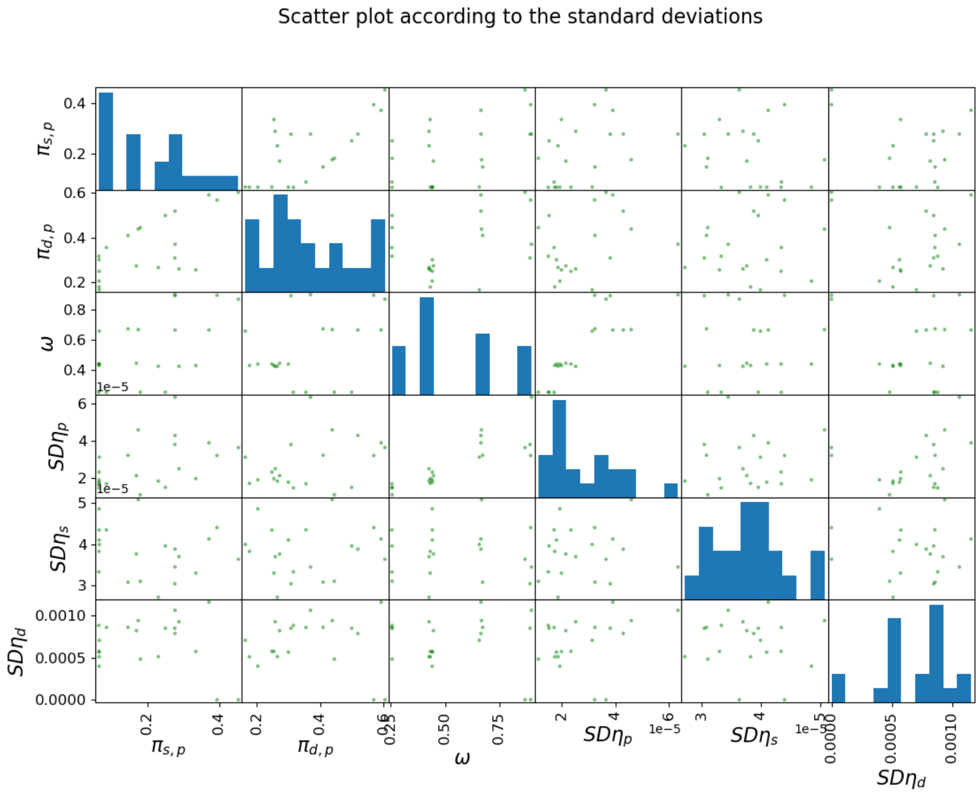


Figure 4.33: Correlations between the inputs of the model, the entrainment ratio and the means of the three efficiencies: On-design points

gave good results but, the critical pressure being not well calibrated, it resulted in almost to the same calibration as before. This case does therefore not bring much additional interest. On the contrary, calibration of the three outputs with three efficiencies together gives satisfying results in terms of accuracy but the standard deviation becomes again higher and the accuracy on the primary mass flow rate, which was easy to calibrate before, is decreased. Nevertheless, when looking closely at the operating conditions, it appears that the standard deviation is raised only for the off-design points. The above cases without standard deviation computed probably only on-design points. This would mean that off-design points would always have higher standard deviations because of the numerous possibilities to match a curve to critical points.

A way to go one step further could be to separate on- and off- design points and calibrate the off-design points with one or two more efficiencies. In this way, the on-design point would keep low standard deviation and that of the off-design points would decrease.

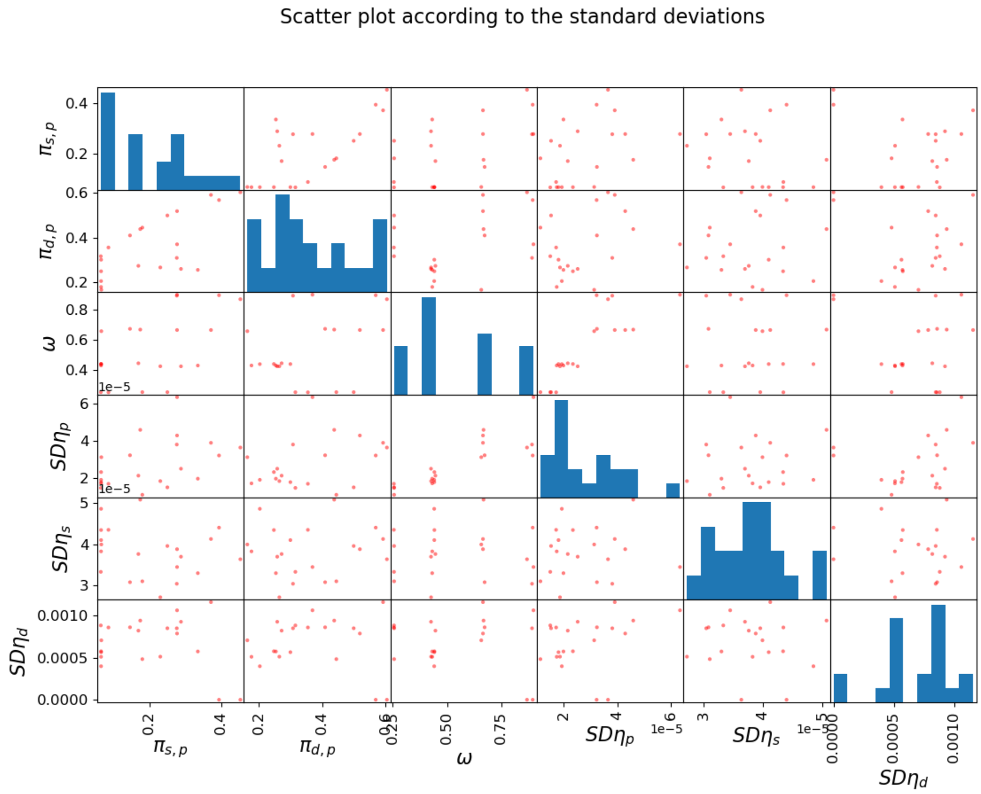


Figure 4.34: Correlations between the inputs of the model, the entrainment ratio and the standard deviations of the three efficiencies: Off-design points

Nevertheless, given that the calibration point per point seems to has achieved is limitations at the end of this chapter, a new way of proceed will be explored in the next chapter. In what following, the calibration will be made, no longer for a point, but with multiple points which compose a part of the curve. In that way the vicinity of the operating condition will be also taken into account. This process, in addition to gaining in robustness, could make the solutions more deterministic.

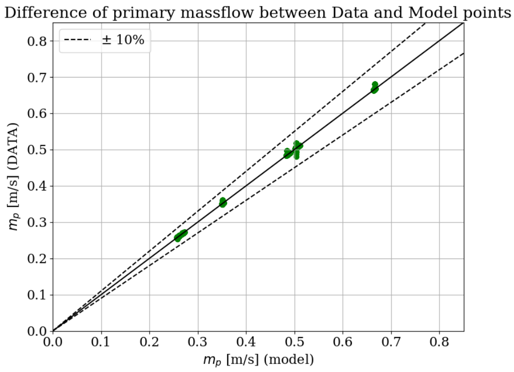


Figure 4.35: Error on the primary mass flow

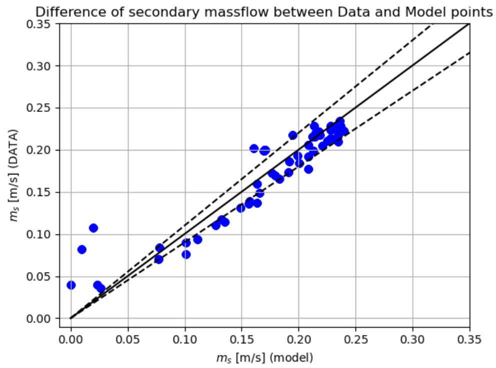


Figure 4.36: Error on the secondary mass flow

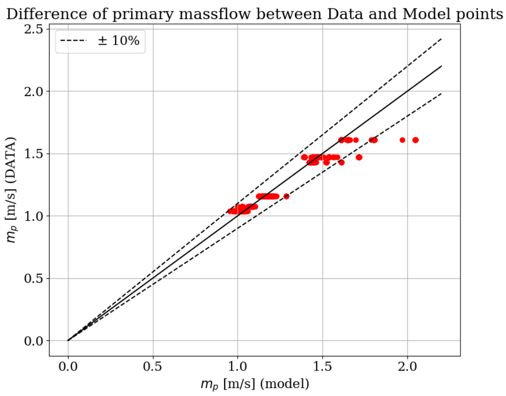


Figure 4.37: Error on the critical pressure

# Chapter 5

# Calibration using several operating points

The previous chapter has indicated that many points lead to indeterminate solutions where using three efficiencies and three outputs. However, from these results, another way to proceed can be found. Indeed, by calibrating with single points, the sensitivity as well as the dependency of each efficiency according to the mass flows and the other parameters are ignored. For example, it is interesting to see how the secondary mass flow rate is linked with the secondary efficiency by looking at which range of value for $\eta_s$ is linked with a mass flow rate.
Figures 5.1 and 5.2 compare the value of the secondary efficiency for each secondary mass flows. On the left $\eta_p$ and $\eta_s$ are only used for the calibration of the two mass flow rates and the critical pressure. On the right the result with calibration of the same outputs and the addition of $\eta_d$ into the parameters can be seen.

As can be seen, even if these two cases seem firstly to give similar results, the solutions according to the efficiencies do not appear to be the same. Indeed, in the right part of the figure, the value of the secondary efficiency can change significantly in the near vicinity of the solution for the secondary mass flow rate. This means that another solution can be found when changing a bit the mass flow rate. On the contrary, when using two efficiencies, the value of the secondary efficiency changes linearly according to the mass flow rate. It should be remembered that, for this case, there were no multiplicity of solutions.

In order to pass through this problem, a solution could be to enlarge the optimization zone to multiple points, or parts of the curves, instead of a single operating point. By this way, the zone near the point would be taken into account. The aim of this chapter is therefore to calibrate the model by taking into account a zone composed of multiple points in the cost function. To do so, the 6 curves composing the

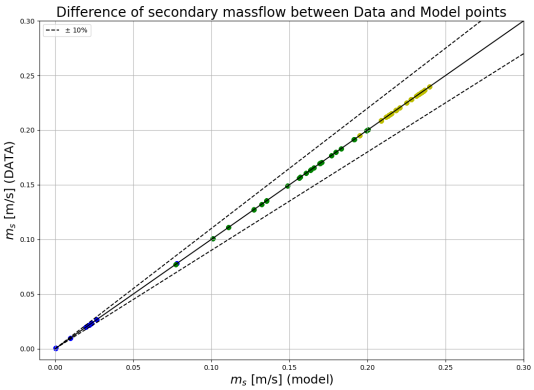


Figure 5.1: Error on the secondary mass flow with two efficiencies. The color mean: red: $\eta_s$>0.75, yellow : 0.75>$\eta_s$>0.5, green: 0.5>$\eta_s$>0.2 and blue: 0.2>$\eta_s$

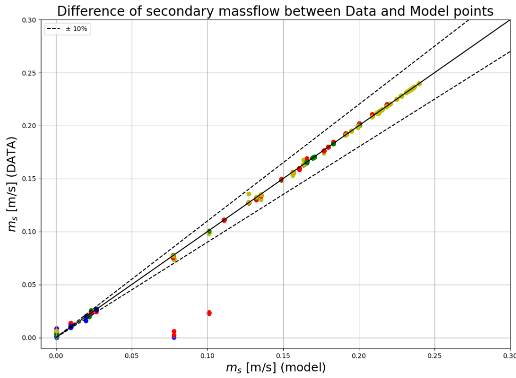


Figure 5.2: Error on the secondary mass flow with three efficiencies. The color mean: red: $\eta_s$>0.75, yellow : 0.75>$\eta_s$>0.5, green: 0.5>$\eta_s$>0.2 and blue: 0.2>$\eta_s$

data-set will be firstly re-sampled, by interpolation, in order to be able to take into account the regions around the points. The purpose of this is to become more deterministic. Indeed, by making a larger set for the cost function, the optimizer could make the difference between two good solutions when calibrating for a single point. That is also why a sorting of solutions will be made at the end of this chapter.

# 5.1 Multiple points calibration

## 5.1.1 Methodology

The data-set is composed of 63 points that belong to 6 different operating curves. In order to model the whole curves, they are firstly resampled by interpolation according to the back pressure. To do so, a linear interpolation with the function interp1d from the scipy.interpolate package [39] is made for each parameter of the data-set. The obtained curves are after resampled with one hundred points uniformly distributed with respect to the back pressure. In addition to having information for the parameters everywhere in the parameter space, this resampling allows to have the same amount of points for each curves. In this way, they have the same weight in the results. Indeed, the curve composed of 6 points was underrepresented in the foregoing results compared with the curve composed of 17 points. This is no longer the case in this chapter.

The calibration is carried out with the same procedure as in chapter 4 with the exception that the cost function is composed of multiple operating points. The calibration is done at first for $\eta_p$ and after for $\eta_s$ and $\eta_d$ like in chapter 4. This way of calibrating has been chosen, at first, with three efficiencies because it was shown that using two of them was insufficient in terms of accuracy. Moreover, given the good accuracy and insignificant standard deviation obtained for $\eta_p$, it is better not to sacrifice good prediction on something simple as the primary mass flow rate. In addition to this, breaking the calibration in two parts saves a lot of computation time, especially for this part of the study where the optimization is made on about ten times more points than in the previous chapter. Thanks to this, a few hours or even a day are saved.
The cost functions of the optimization problem are then :

$$\min_{\eta_p} \sum_{i=1}^{x} J_1(\eta_p) = ||m^p_{p,i} - m_{p,i}|| \tag{5.1}$$

$$\min_{\eta_s,\eta_d,\eta_p} \sum_{i=1}^{x} J_2(\eta_s;\eta_d;\eta_p) = ||m^p_{s,i} - m_{s,i}||/m_{s,i} + ||p^p_{c,i} - p_{c,i}||/10^5 \tag{5.2}$$

In these equations, x represents the number of points for which the optimization will be made and i is the indice of the point composing the hundred points of the curve.
The number x of points has to be chosen strategically to find a trade-off between robustness and accuracy of the solutions. From what has been tested, x=20 has seemed to be a good equilibrium. In this way, the curves are each cut in 5 and there are 30 sets of points to optimize, which represents a bit less than the half of the size of the original data-set. After that, calibration with x=50 has been done to observe if accuracy could be maintained with as much points. Moreover, by dividing the curve in two, the separation corresponds more or less to a separation between on- and off-design, which could be already sufficient.

After this, another way to proceed has been try. First, given that the data-set was containing 63 points, the choice to take x=10 has been made to divide each curve by ten and deal with approximately the same amount of calibration points as in chapter 4. By this way, a good accuracy is ensured. During the process with x=20 and 50, it could have happened that some bad results were found by the optimizer, because of local minima, while better results were found as well for a same set of points. Given that some results found by the optimizer are significantly better than others in terms of accuracy, those better results are only kept after the process for this step. In this way, the standard deviation of the problematic points could be decreased to insignificant values. More precisely, when the optimizer has found a solution below 5% of error according to the outputs, all the results having an error above this limit are deleted from the possible solutions. Given that x = 10, the optimizer is able to find an accurate solution for the vast majority of the set of points.

## 5.1.2 Results and comments

### 5.1.2.1 Calibration with x=20

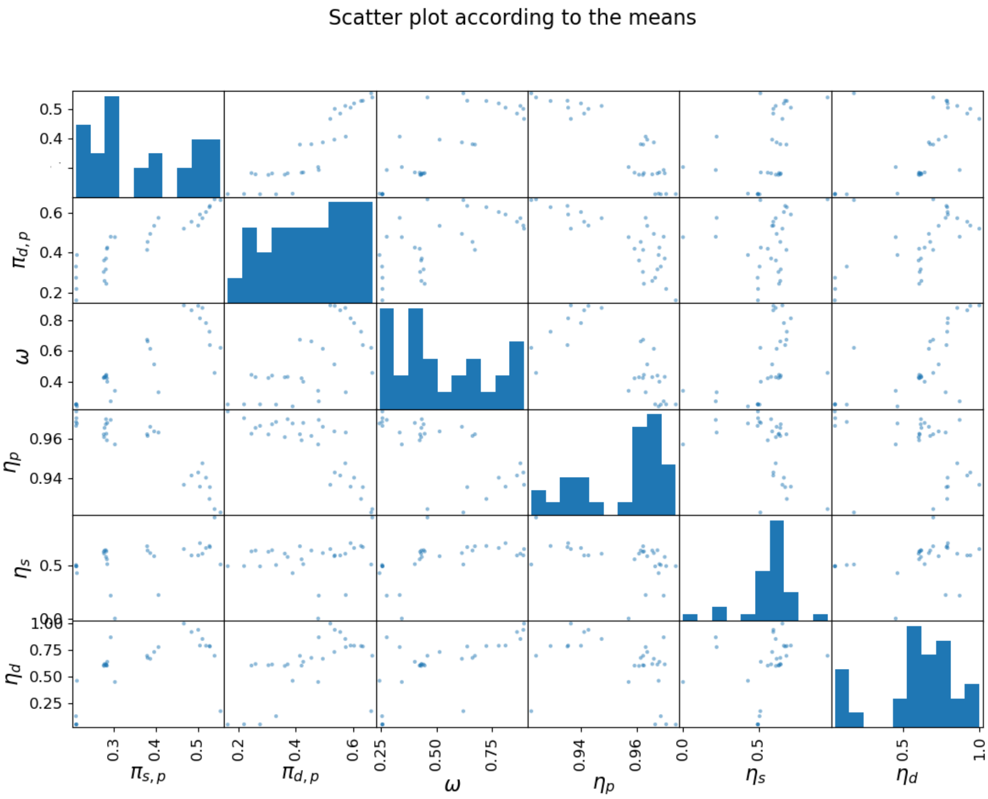


Figure 5.3: Scatter plot of the 30 mean values for the efficiencies and parameters obtained with x=20

The figures 5.3 and 5.4 represent the scatter plots obtained for the 30 calibration results with x=20. On the mean scatter plot (figures 5.3), it can be seen that the primary efficiency is relatively constant until the compression ratio $\pi_{s,p}$ or the entrainment ratio $\omega$ achieves high values. It should be noticed that this is also the case with the ratio $\pi_{d,p}$ which means that the value of the primary efficiency tends to decrease for off-design operating conditions. This last observation has also been

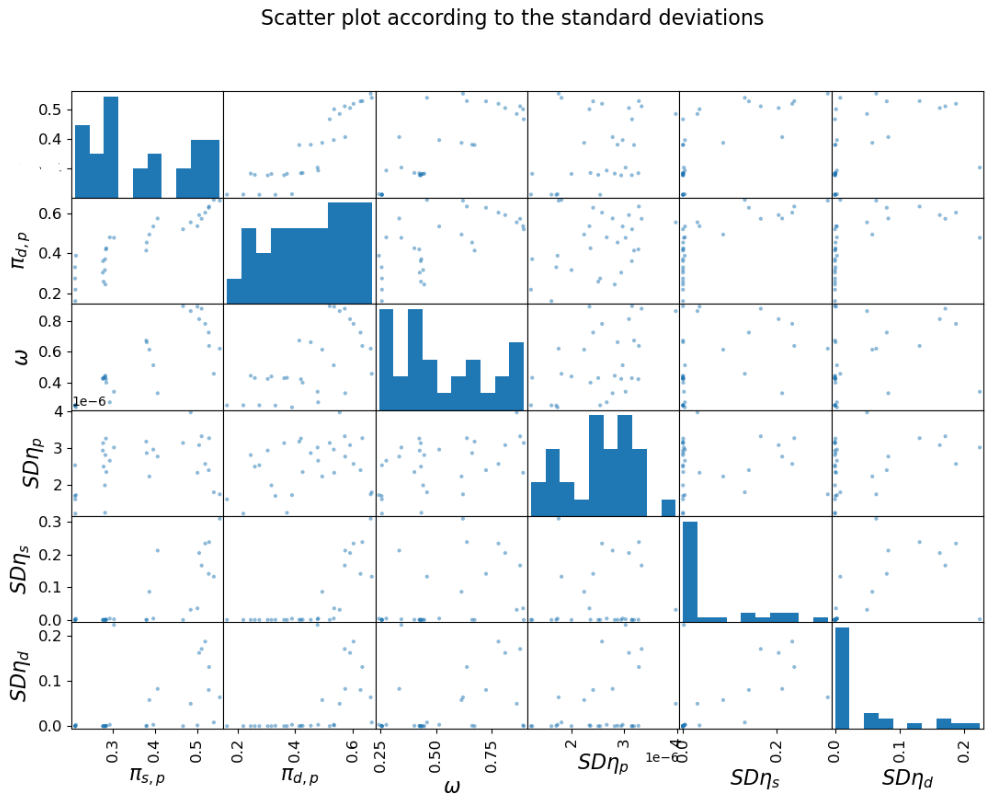


Figure 5.4: Scatter plot of the 30 means values for the efficiencies and parameters obtained with x=20

seen during the process. Another observation that could be made is the quite low values that the secondary efficiency takes in general.
The standard deviation remains insignificant for the $\eta_p$ as expected. However, even if most of their values no longer have variations, there is still a significant part of the solutions that have unacceptably high standard deviations. As seen from the ratio $\pi_{d,p}$, all of these points are off-design as it was also the case at the end of chapter 4. Nevertheless, proportionally, these results are less numerous.

Figures 5.5, 5.6 and 5.7 represent the accuracy of the results for primary, secondary and critical pressure respectively. The values are computed according to their mean values along the 20 points of the curves considered for a solution.
Globally, the model predicts fairly well the outputs. The primary mass flow rate is almost perfectly estimated, as usual, but the secondary mass flow rate has sometimes been miscalculated. The horizontal line formed by the bad computed points indicates that there are multiple combinations for the same set of points.

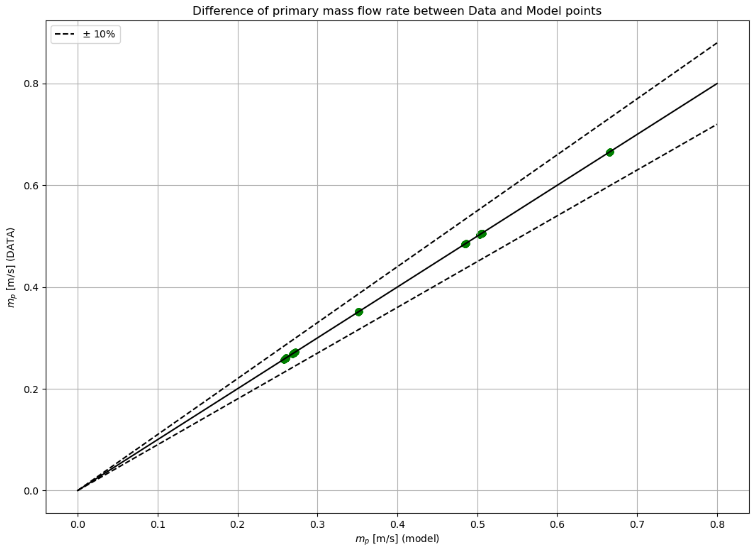


Figure 5.5: Error on the primary mass flow

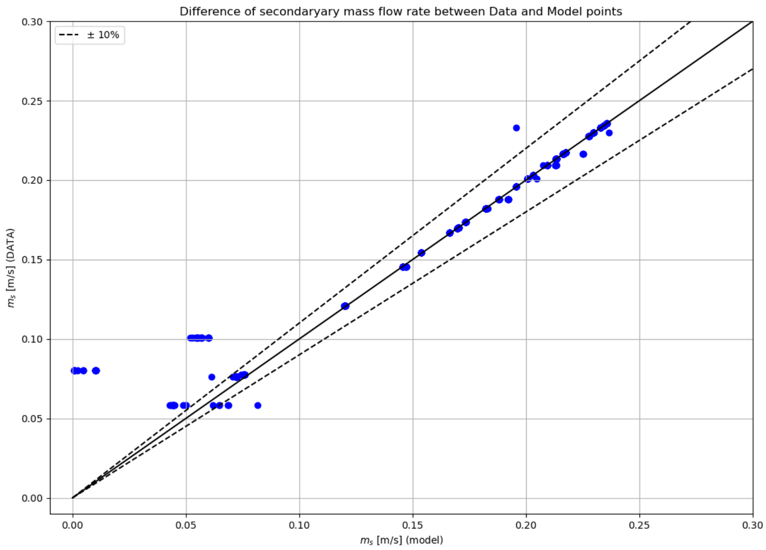


Figure 5.6: Error on the secondary mass flow

The higher values of the secondary mass flow rates are nevertheless calibrated with a good accuracy. It can be observed that some points have a bad accuracy while

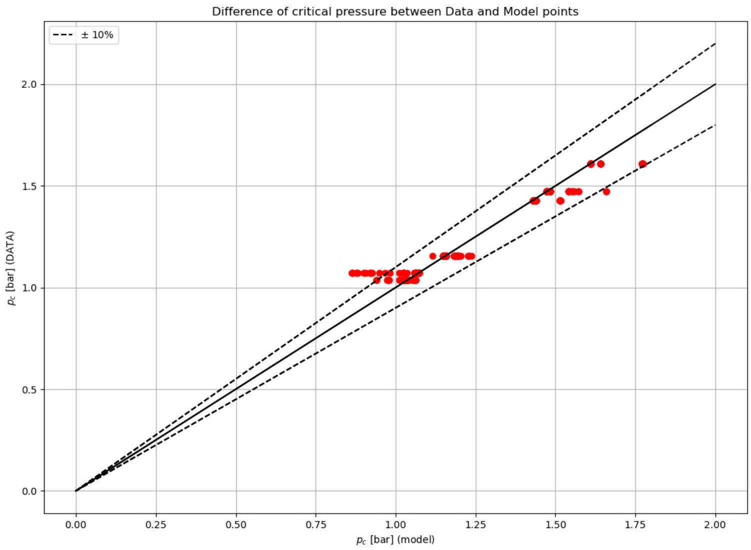


Figure 5.7: Error on the critical pressure

other points which are probably in another combination according to efficiencies have a good one. These points could cause a rising for the standard deviation for some results. By deleting them, there should be less standard deviation in the results.
The critical pressure, as for it, is computed with better accuracy as previously. Given that one critical pressure corresponds to one curve, This result suggests that the sampling had a beneficial effect for the calibration of the critical pressure. This may be due to the fact that, by taking a less restrictive region than an operating point, points for which the calibration is not accurate are less visible because the mean value seems correct.

#### 5.1.2.2 Calibration with x=50

The calibration 50 points per 50 points takes half a curve into account in the cost function. Given that 20 points gave quite good accuracy, 50 has been tried to see if the standard deviation could be again reduced by taking more points into account. By this separation, it can be assumed that the first part accounts for the on-design part of the curve and the second accounts for the off-design operation. In figures 5.8 and 5.9 correlations plots with x=50 for the means and standard deviation for the efficiencies respectively are shown.

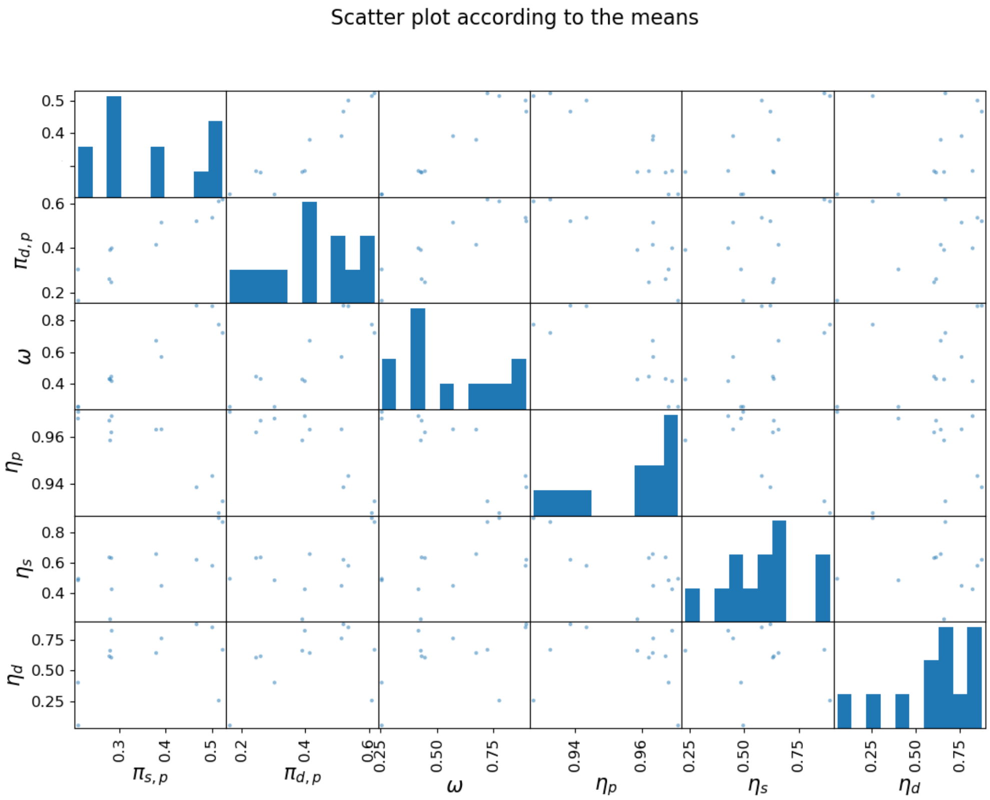


Figure 5.8: Scatter plot of the 12 mean values for the efficiencies and parameters obtained with x=50

On the means scatter plot, not many new links can be observed. Still, it can be observed that the $\eta_s$ has a slight linear dependency with the pressure and entrainment ratios. In the middle values of those, the secondary efficiency seems to be able to take two different values.
The standard deviation has significantly reduced for $\eta_d$ compared with the case with x=20. It remains nevertheless too high for 4 points. This is also and even worthy the case for $\eta_s$. It can be seen on the part of the figure which links the standard deviation for $\eta_s$ and $\eta_d$ that these four problematic sets of points are the same for both.

Figures 5.10, 5.11 and 5.12 draw the accuracy for the calibration with x=50. Even if the critical pressure as well as the primary mass flow rates are well estimated, taking x=50 is too much as seen from the inaccuracy of the calibration of the secondary mass flow rate. However, the points with the same ordinates but different abscissa suggest again that some better solutions than others are found for several points.

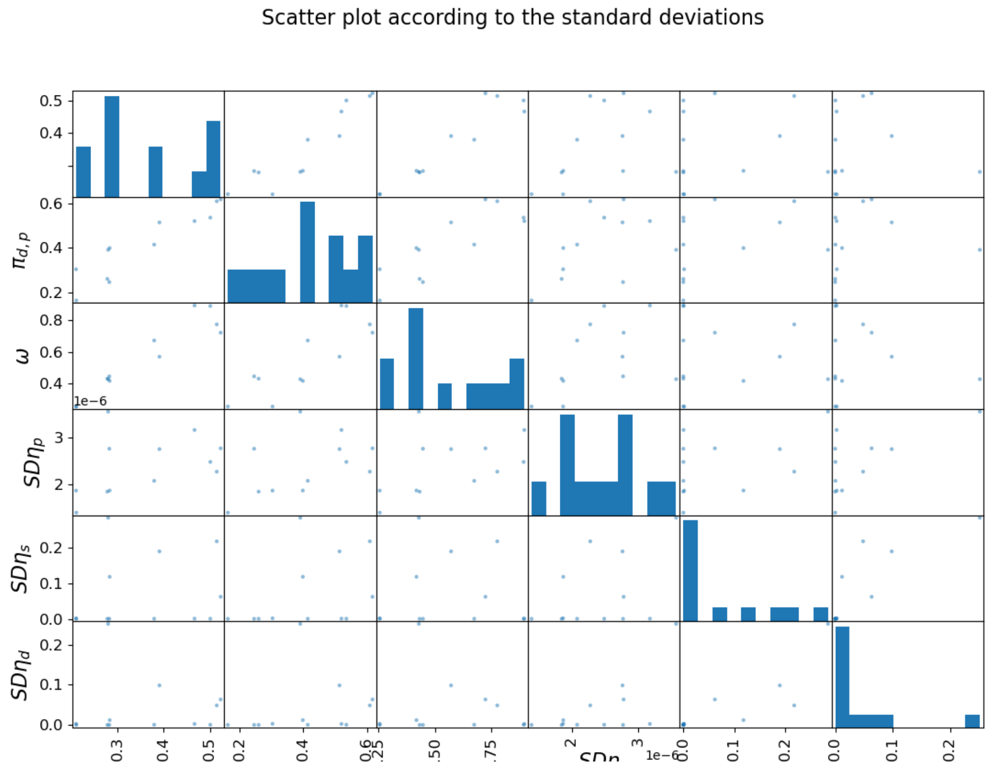


Figure 5.9: Scatter plot of the 12 means values for the efficiencies and parameters obtained with x=50

#### 5.1.2.3 Calibration with x=10 and with deletion of less accurate results

Before analysing the scatter plot, several cases observed during the process should be mentioned. First, the operation conditions for which bad solutions were deleted the most were for off-design applications and especially when the entrainment ratio was low. Indeed, it has been seen in the previous results that those regions contained solutions with bad and less bad accuracy. The deleted solutions are often corresponding to points at the border of the efficiencies space i.e. with an efficiency equal to 1.0 or 0.0. This is probably due the fact that a minima is achieved in the border of the authorized domain from the point of view of the optimizer. Finally, the bad accuracy results were representing a small part of the solutions of a set of points (about 10% to 15%). That was however enough to make significantly change their standard deviation in some cases.

Figures 5.13 and 5.13 illustrate the correlations for the means and standard deviation for efficiencies respectively. To better understand these results, the accuracy plots of figures 5.15, 5.16 and 5.17 should be firstly observed. Indeed, it can be seen that

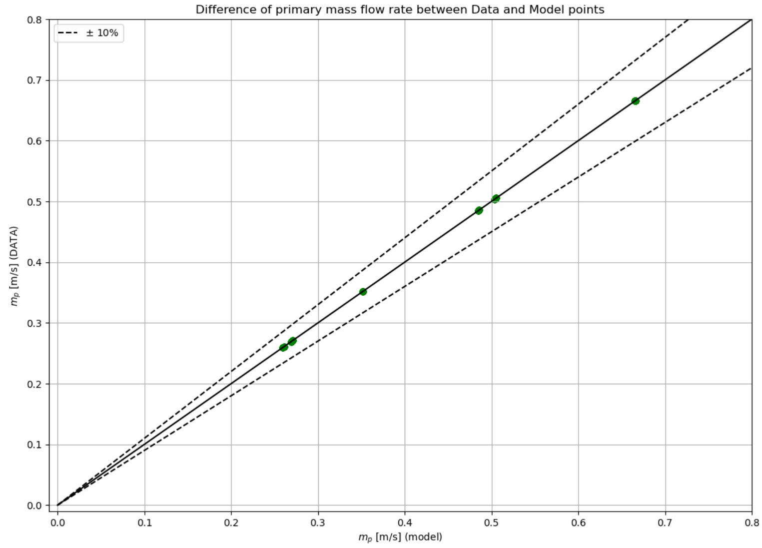


Figure 5.10: Error on the primary mass flow

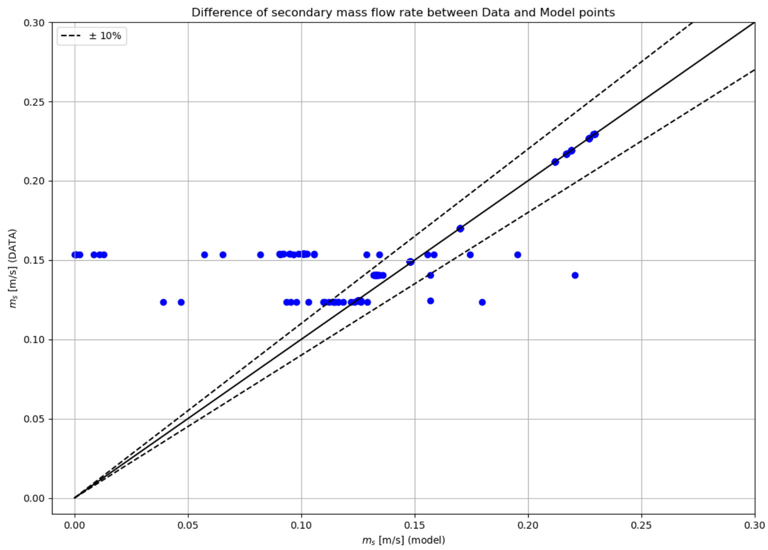


Figure 5.11: Error on the secondary mass flow

there are several sets of points for which no sufficiently good solutions are found. Most of these solutions correspond to low secondary mass flow rates. For those, the

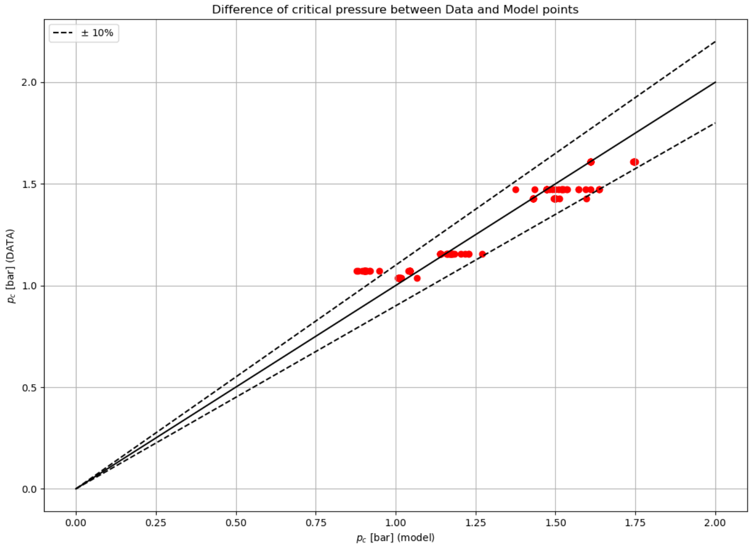


Figure 5.12: Error on the critical pressure

model predicts a wide variation (often more than 100%) for the secondary mass flow rate. The standard deviation for the efficiencies is therefore high as well. This nevertheless raises the fact that the three efficiencies are not enough to calibrate correctly the set of points with low secondary mass flow rate.
It can be seen on the standard deviation scatter plot (figure 5.16) that this way of proceeding has been quite successful for most of the points. The standard deviation remains however high for 7 sets of points out of a total of 60. As stated just before, these sets are, among others, those for which the secondary mass flow rate is badly estimated. However, there is still one on-design point which is part of the seven.
On the means scatter plot (figure 5.13), points seem more ordered than in chapter 4. Indeed, efficiencies are showing some dependencies such as curves appearing on the scatter plot when expressing efficiencies according to pressure ratio $\pi_{d,p}$ and entrainment ratio.

### 5.1.3 Conclusion

In this chapter, a calibration by taking into account sets of neighbouring points instead of points has been made. To do so, three cases have been considered.
First, a case for which the calibration has been made by taking into account 20 per 20 points in the cost function has been carried out. This calibration has shown sufficient accuracy but still too high standard variation for a significant part of the

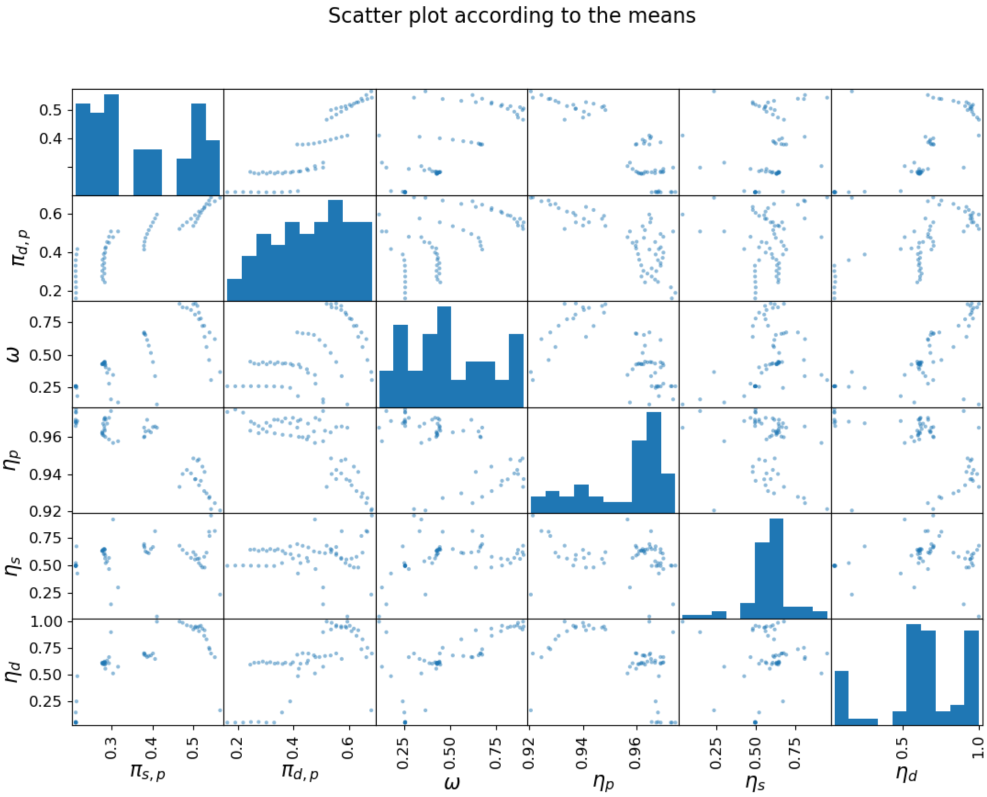


Figure 5.13: Scatter plot of the 12 means values for the efficiencies and parameters obtained with x=50

solutions. In the hope of decreasing the standard deviation, the next case has been to take no longer 20 but 50 points in the cost function. In this case, the curves were divided only by 2. This process has however given inaccurate results with still too high standard deviation for some sets of points. Nevertheless, from this experiment, several conclusions could be gathered. Indeed, the presence of such high standard deviations, even with 50 points in the cost function, suggests that multiple solutions exist no matter the size of the region that is taken into account in the cost function.

More importantly, the combination for some set with bad accuracy and high standard deviation has raised the question of knowing if it would be these bad solutions that cause the high standard deviation. Indeed, given that bad solutions were sometimes found while accurate solutions were found too, a logical choice was to delete the bad solutions from the results to avoid them from polluting the set by producing high standard deviations.

The third case was therefore made with 10 points in the cost function and with

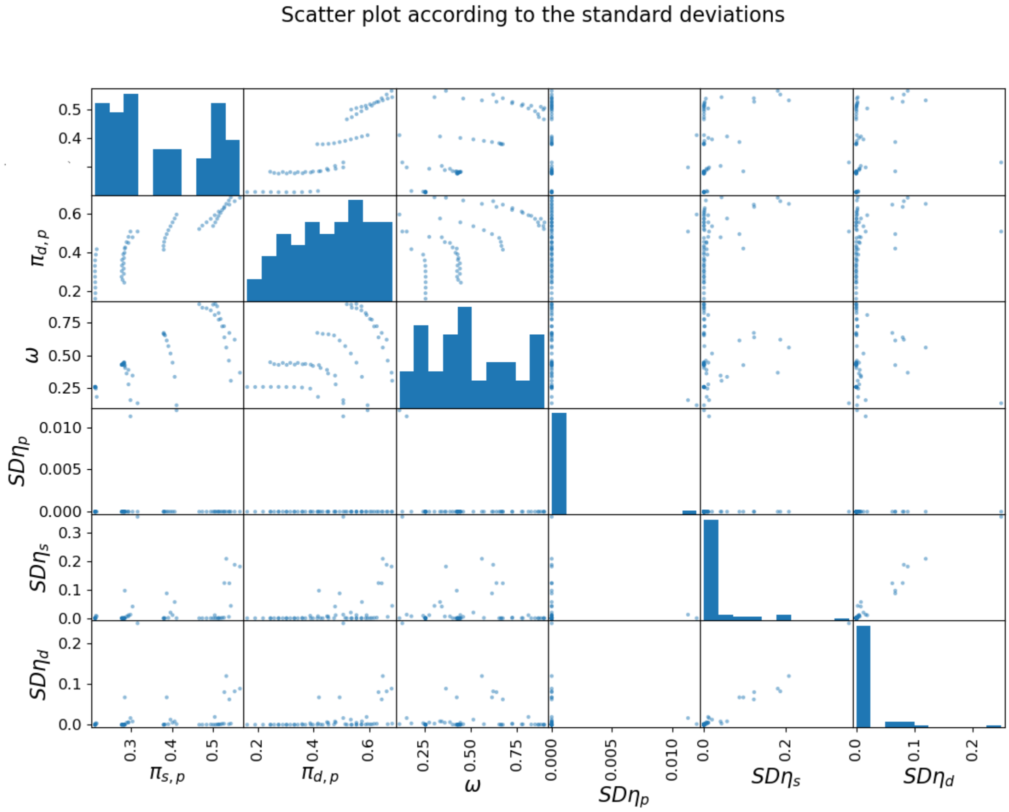


Figure 5.14: Scatter plot of the 12 mean values for the efficiencies and parameters obtained with x=50

deletion of bad results. This last experiment has given quite high accuracy and low standard deviation except for several problematic points for which the optimizer is not able to find a sufficiently accurate solution. This case dis also make appear clearer trends in the means scatter plot.

To go further, as seen with the results of the last case, several paths can be considered. First, by lowering the limit of what is considered as a good accuracy, for example to 1%, other less accurate solutions could be deleted. However, it raises the question of knowing at which accuracy a solution could be considered as good. Moreover, this would not help to solve the low secondary mass flow rate cases. Another path could be to consider more points in the cost function, like 20 or 50, and add an efficiency in the calibration for the previously less accurate points. All this with the same deletion as before. In this way, the additional efficiency should brings accuracy while keeping the standard deviation as low as possible with the deletion and the large number of points in the cost function.

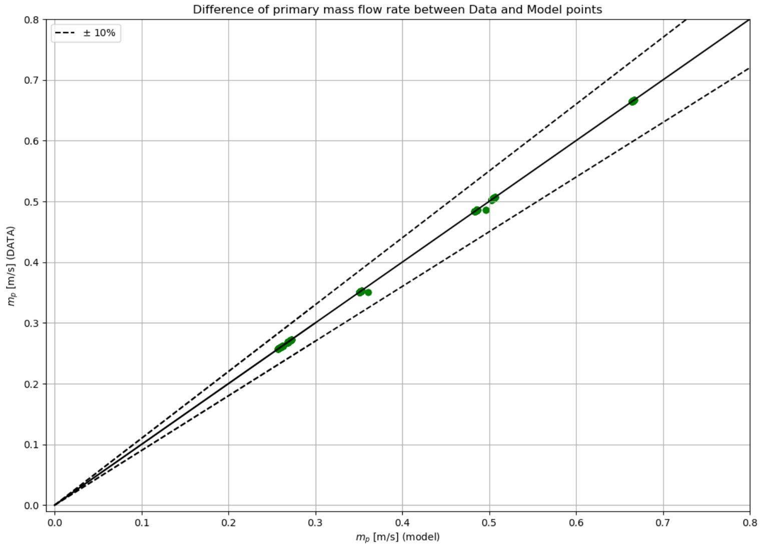


Figure 5.15: Error on the primary mass flow

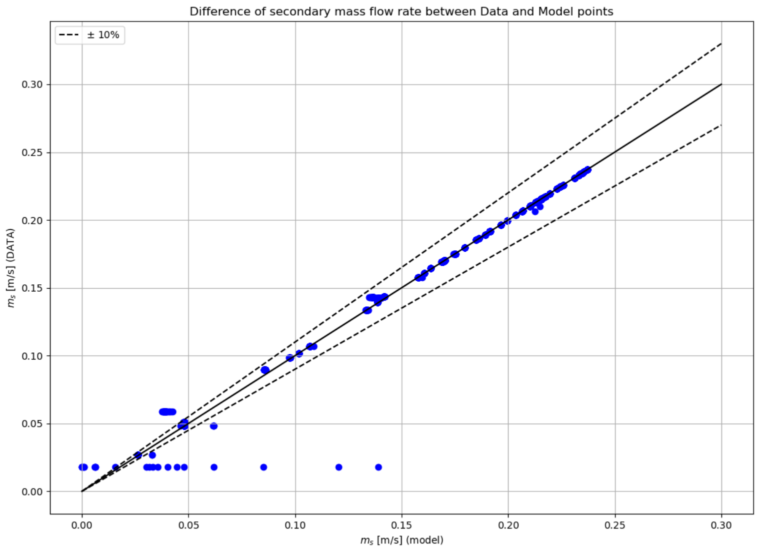


Figure 5.16: Error on the secondary mass flow

To conclude, the results in this part show that it is better to consider a point and its neighbouring points rather than a single point. The standard deviation has

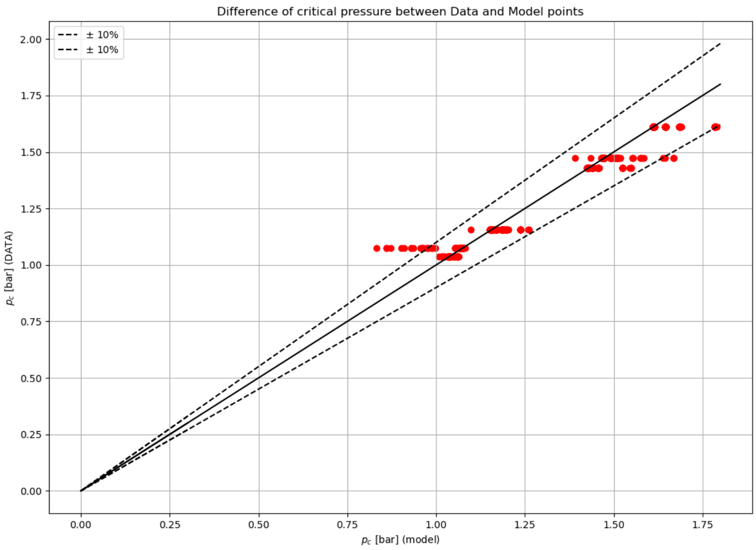


Figure 5.17: Error on the critical pressure

globally decreased while keeping a good accuracy except in the case with x=50. The optimizer is however experiencing difficulties for the low secondary mass flow rates. For these points, they may be either not considered or taken separately with more efficiencies to be well calibrated. As for the means, even if the trends in the scatter plots may actually not be sufficient to launch a machine learning model, the improvements by drawing the links between parameters are encouraging to go further in this direction.

# Chapter 6

# Conclusion and perspectives

In this work, various strategies have been adopted by playing with the efficiencies and the parameters of a thermodynamic model [1] in order to make this one usable for a further machine learning model.

In section 1.4 the main questions were asked that this thesis was intended to respond. In the first part of this conclusion, the achievements of this work will be presented by answering these questions. This part will be followed by the research perspectives that could be made in order to continue this work.

## 6.1 Achievements

The first question was whether multiple combinations of the five efficiencies could exist with other fixed outputs to predict the mass flow rates. The short answer to this question is : yes they can! Chapter 3 answered the first question in multiple ways. It has been proven that the entrainment ratio could be predicted with multiple sets of efficiencies for an entire operating curve and it was shown as well that the entrainment ratio alone to calibrate was not sufficient. Moreover when calibrating for both mass flows and for a single operating point, there were multiple possible ways to calibrate as well. In addition to that, strong links between $\eta_s$ and $\eta_{py}$ and between $\eta d$ and $\eta_m$ have been discovered. It was also shown that the primary efficiency was keeping a single solution in both cases.

The conclusions gathered in chapter 3 led to another question : Are there too many efficiencies defined in the model ? Is each of them relevant ? Followed by a second question : Are the mass flow rates alone, accompanied by the efficiencies, sufficient to characterise all the states on the ejector ? The answer is found in chapter 4. The answer to these questions are not that obvious. The first part of this chapter has shown that a calibration could be done accurately and without multiplicity of solutions with only two efficiencies and with a data-set composed of on and off-design points. This means that, yes there are too many efficiencies defined in the model but maybe not for all the operating conditions. Indeed, the second part of the chapter has showed that both same mass flow rates could correspond to totally different operating conditions. By showing this, the second part of chapter 4 has answered the second question of this chapter. Indeed, the critical pressure is needed in the outputs to distinguish at least an on-design point from on off-design point for the optimizer. However, even with this third output, the calibration has been done accurately with the help of three efficiencies. This chapter has brought also the fact that the calibration could be done separately for the primary mass flow rate with the help of only $\eta_p$. With this way of proceeding, the calibration has given more accurate results and has been achieved with less computation time. Finally, even if each definition of the efficiency is relevant, keeping the five for the calibration is not. Indeed, by putting arbitrarily the values of $\eta_{py}$ and $\eta_m$ to 1.0, one assumes that there is no friction loss during the mixing and the expanding of the primary flow which is not the case in a real ejector. Yet, the model can still be calibrated because the other efficiencies can balance the wrong values.

After the insufficiently satisfying results of chapter 4 about the multiplicity of solutions came a new question: what impact could taking into account the space in the vicinity of a point have on the calibration ? This question has been answered in chapter 5 where this new way of proceeding highlighted limitations of the optimizer which found inaccurate solutions when better ones could also be found for a same set of points. In addition to that, it has been shown that taking into account the point and its neighbours could make the calibration more deterministic while keeping correct accuracy until a certain limit. Moreover, some correlation has been observed between parameters and efficiencies found by calibration in this chapter which encourage the further use of a machine learning model.

## 6.2 Perspectives

As stated in section 1.4, the first goal of this work was to launch a machine learning model capable of finding the efficiencies when other input parameters were given to it. Although this model was not made true, this whole thesis was dedicated to find a way to be able to launch it by a proper way. Therefore, this section presents some elements to go further in this work.

In view of the last results of chapter 4 and chapter 5, the way of considering at first only on-design points could be planned. Given that, the multiplicity of solutions is no longer a problem for those points, by calibration, there could be clearer correlations appearing with the other parameters of the thermodynamic model. With these criterions, all the conditions would be combined to be able to launch an efficient machine learning model for the on-design regime. In addition to that, if adding an efficiency to the bad accurate points of chapter 4 gives some encouraging results, the machine learning model could be further expanded for off-design application with three efficiencies in on-design and four in off-design.

An improvement could be made on the model of Metsue et al. [1] by implementing also a back-flow mode. In this way, the case for which the entrainment ratio approaches zero would no longer be problematic.
Finally, another way to proceed could be to consider a calibration with the help of CFD results. In this way, one could extract the information of local phenomena as well as velocities pressures etc. In this way, the efficiencies could be obtained with their own formulas. There would be therefore only one single possible set of efficiencies. By collecting a sufficient amount of CFD simulations, a machine learning model could finally be launched. However, this approach would be costly due to the CFD simulations required.

# Appendix A

# Ejector dimensions

## A.1 Ejector geometry for Mazzeli et. al experiment

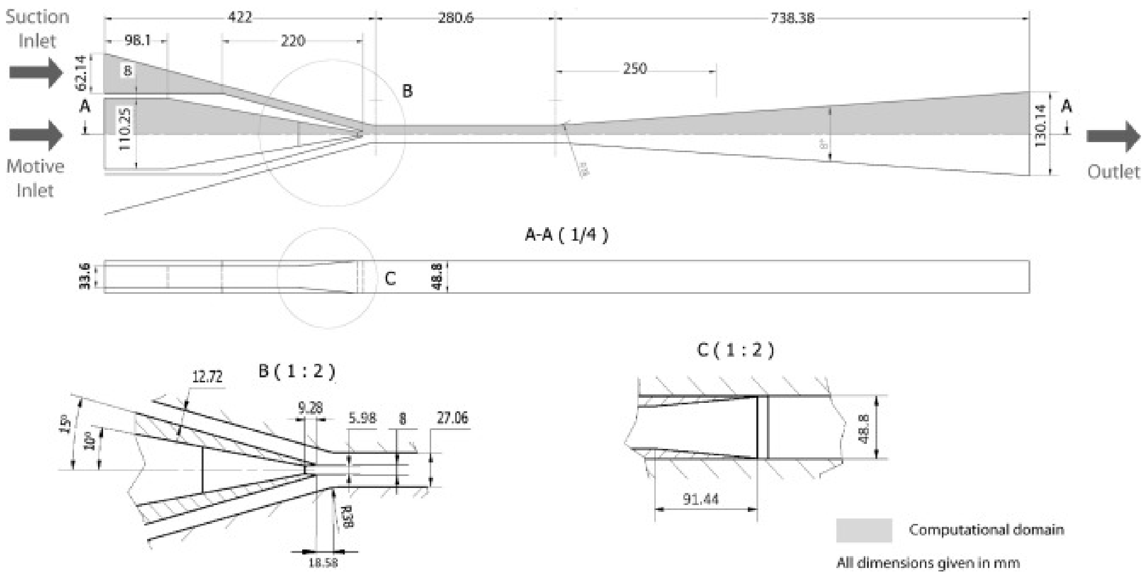


Figure A.1: Dimensioned drawing of ejector used for experiments [37]. All dimensions are in mm.

# Appendix B

# Flowcharts of the model of Metsue et al [1] for on-design and off-design operations

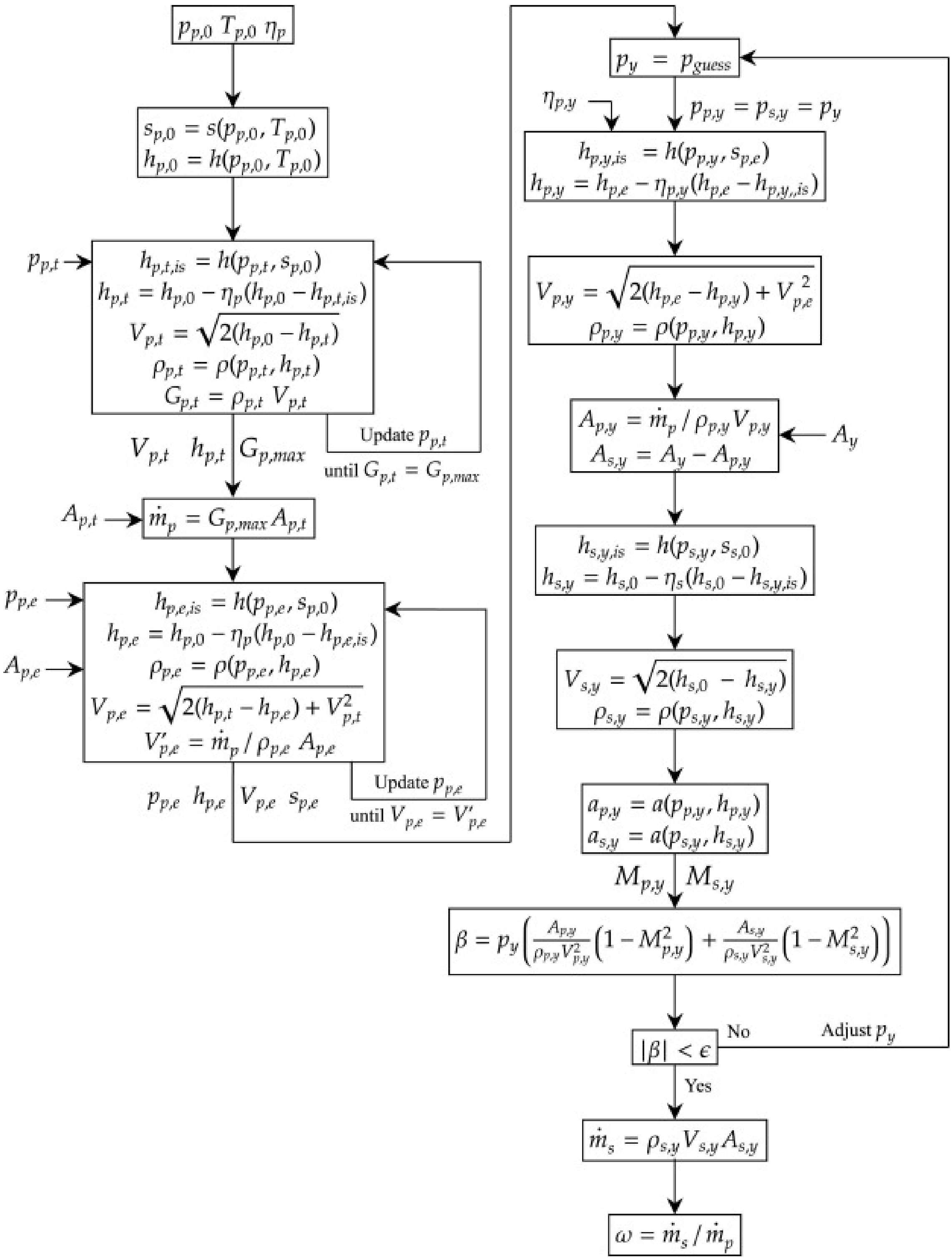


Figure B.1: Flowchart of the model of Metsue et al. [1] for on-design operations using the compound-choking criterion (fig 2.2).

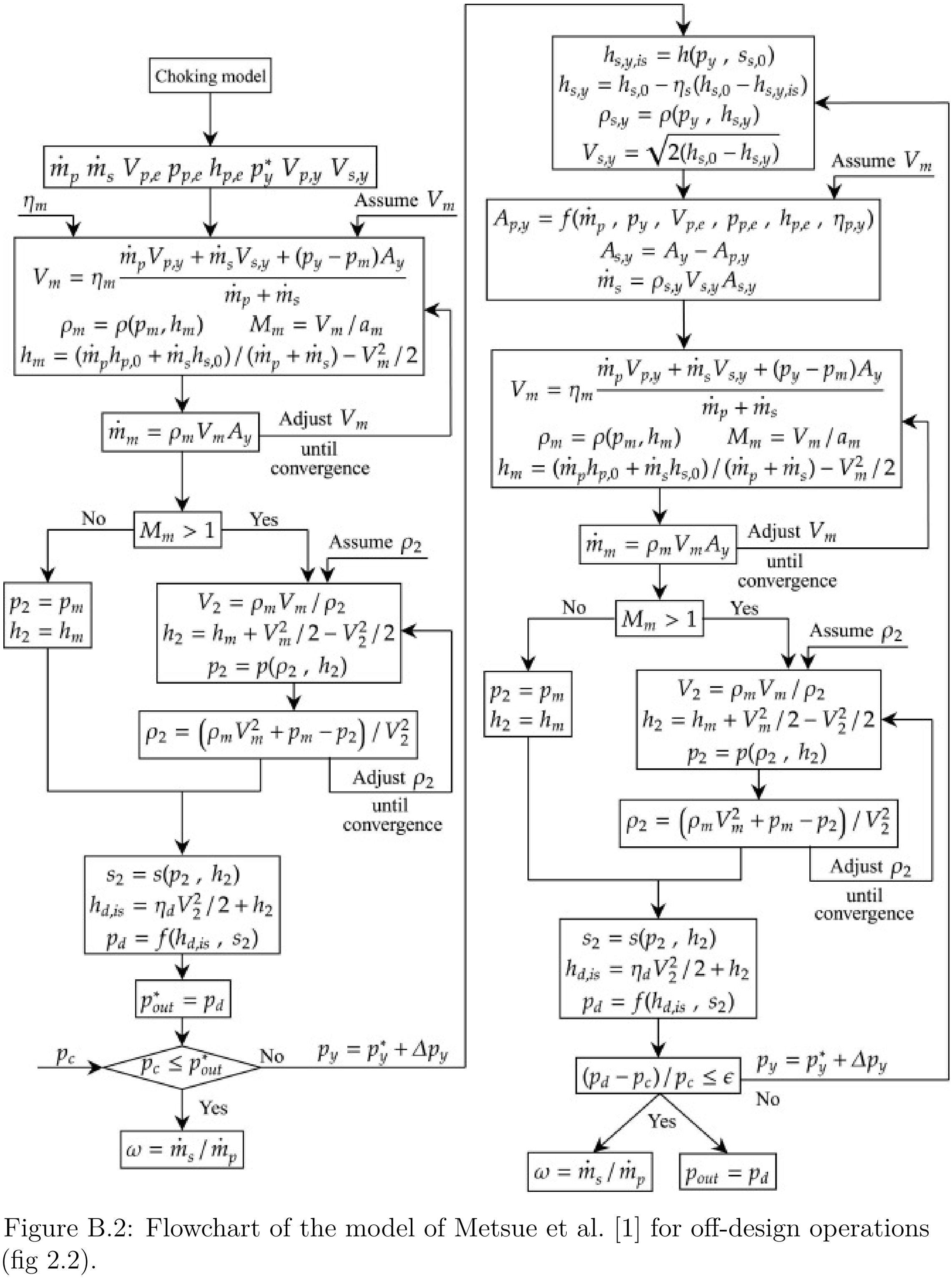


Figure B.2: Flowchart of the model of Metsue et al. [1] for off-design operations (fig 2.2).

# Bibliography


[1] A. Metsue, R. Debroeyer, S. Poncet, and Y. Bartosiewicz, “An improved thermodynamic model for supersonic real-gas ejectors using the compound-choking theory,” *Energy 238 121856*, 2021.

[2] F. D. BERKELEY, “Ejectors have a wide range of uses,” *GRAHAM MANUFACTURING COMPANY, INC., BATAVIA, N. Y*, 1958.

[3] D. E. I, “Recent developments in the design theories and applications of ejectors-a review,” *J. Inst. Energy*, vol. 68, pp. 65–79, 1995.

[4] Y. Choi and W. Soh, “Computational analysis of the flowfield ejector nozzle of a two-dimensional ejector nozzle,” *NASA: Washington, DC, USA*, 1990.

[5] J. Liu, L. Wang, L. Jia, and H. Xue., “Thermodynamic analysis of the steam ejector for desalination applications,” *Applied thermal engineering*, vol. 159, no. 113883, 2019.

[6] Z. Aidoun, K. Ameur, M. Falsafioon, and B. Messaoud, “Current advances in ejector modeling, experimentation and applications for refrigeration and heat pumps. part 1: Single-phase ejectors,” *Inventions*, vol. 4, no. 15, 2019.

[7] B. Tashtoush, M. Al-Nimr, and M. Khasawneh, “A comprehensive review of ejector design, performance, and applications,” *Applied Energy*, vol. 240, pp. 138–172, 02 2019.

[8] Y. Bartosiewicz, Z. Aidoun, P. Desevaux, and Y. Mercadier, “Numerical and experimental investigations on supersonic ejectors,” *International Journal of Heat and Fluid Flow*, vol. 26, pp. 56–70, 2005.

[9] S. Gavioli, J.-M. Seynhaeve, and Y. Bartosiewicz, “Dynamic simulation of an ejector-based cooling system for residential solar air-conditioning,” vol. 6, 11 2012.

[10] Y. W. R. He, S.and Li, "Recent developments in the design theories and applications of ejectors-a review," *Renewable and Sustainable Energy Reviews*, vol. 13, pp. 1760–1780, 2009.

[11] E. L. F. Keenan, J.; Neumann, "A simple air ejector," *J. Appl. Mech. Trans. ASME*, vol. 64, pp. 75–81, 1942.

[12] E. L. F. Keenan, J.; Neumann, "An investigation of ejector design by analysis and experiment," *J. Appl. Mech. Trans. ASME*, vol. 72, pp. 299–309, 1950.

[13] X. D. C. H. E. Chen, Z.; Jin, "Ejector performance analysis under overall operating conditions considering adjustable nozzle structure," *Int. I. Refrig.*, vol. 84, pp. 274–286, 2017.

[14] W. Chen, C. Shi, S. Zhang, H. Chen, D. Chong, and J. Yan., "Theoretical analysis of ejector refrigeration system performance under overall modes," *Applied Energy*, vol. 185, p. 2074 – 2084, 2017. Clean, Efficient and Affordable Energy for a Sustainable Future.

[15] J. Munday and D. Bagster, "A new ejector theory applied to steam jet refrigeration," *Industrial Engineering Chemistry Process Design and Development*, vol. 16, 10 1977.

[16] J. Fabri and R. Siestrunck, "Supersonic air ejectors," *Advances in Applied Mechanics*, vol. 5, 1958.

[17] A. H. Bernstein, W. H., and C. Hevenor, "Compound-compressible nozzle flow," *Journal of Applied Mechanics*, vol. 34, no. 3, p. 548, 1967.

[18] F. LiZhao, C. Tian, C. Wu, and X. Wang, "Performance predictions of dry and wet vapors ejectors over entire operational range," *Energies*, vol. 10, no. 7, p. 1012, 2017.

[19] D. Buyadgie, O. Buyadgie, O. Drakhnia, P. Brodetsky3, and V. Maisotsenko, "Solar low-pressure turbo-ejector maisotsenko cycle-based power system for electricity, heating, cooling and distillation," *International Journal of Low-Carbon Technologies*, vol. 10, no. 2, pp. 157–164, 2015.

[20] Y. S. R. S. Kim HD, Setoguchi T, "Navier-stokes computations of the supersonic ejector-diffuser system with a second throat," *J Therm Sci*, vol. 8, pp. 79–83, 1999.

[21] N. RS, "Computational fluid dynamics analysis of diffuser performance in gas-powered jet pumps," *Int J Heat Fluid Flow*, vol. 14, pp. 401–7, 1993.

[22] O. Lamberts, P. Chatelain, and Y. Bartosiewicz, “Numerical and experimental evidence of the fabri-choking in a supersonic ejector,” *International Journal of Heat and Fluid Flow*, vol. 69, pp. 194–209, 2018.

[23] D. Brezgin, K. Aronson, F. Mazzelli, and A. Milazzo, “The surface roughness effect on the performance of supersonic ejectors,” *Thermophys. Aeromech.*, vol. 24, p. 553–561, 2017.

[24] W. Chen, C. Shi, S. Zhang, H. Chen, D. Chong, and J. Yan, “Theoretical analysis of ejector refrigeration system performance under overall modes,” *Applied Energy*, vol. 185, 02 2016.

[25] S. He, Y. Li, and R. Wang, “Progress of mathematical modeling on ejectors,” *Renewable and Sustainable Energy Reviews*, vol. 13, pp. 1760–1780, 10 2009.

[26] A. Metsue, Y. Bartosiewicz, and S. Poncet, “Investigation on ejector design for co2 heat pump applications using dymola,” 09 2021.

[27] J. Chang, C. Wang, and V. Petrenko, “A 1-d analysis of ejector performance,” *International Journal of Refrigeration*, vol. 22, pp. 354–364, 08 1999.

[28] A. Hemidi, F. Henry, S. Leclaire, J.-M. Seynhaeve, and Y. Bartosiewicz, “Cfd analysis of a supersonic air ejector. part ii: Relation between global operation and local flow features,” *Applied Thermal Engineering*, vol. 29, pp. 2990–2998, 10 2009.

[29] E. W. Lemmon, I. H. Bell, M. L. Huber, and M. O. McLinden, “NIST Standard Reference Database 23: Reference Fluid Thermodynamic and Transport Properties-REFPROP, Version 10.0, National Institute of Standards and Technology,” 2018.

[30] S. Croquer, y. Fang, A. Metsue, Y. Bartosiewicz, and S. Poncet, “Compound-choking theory for supersonic ejectors working with real gas,” *Energy*, vol. 227, 03 2021.

[31] A. Hemidi, F. Henry, S. Leclaire, J.-M. Seynhaeve, and Y. Bartosiewicz, “Cfd analysis of a supersonic air ejector. part i: Experimental validation of single-phase and two-phase operation,” *Applied Thermal Engineering*, vol. 29, pp. 1523–1531, 06 2009.

[32] A. Metsue, “An improved thermodynamic model for supersonic ejectors,” Master’s thesis, UCLouvain, 2020.

[33] P. Virtanen, R. Gommers, T. E. Oliphant, M. Haberland, T. Reddy, D. Cournapeau, E. Burovski, P. Peterson, W. Weckesser, J. Bright, S. J. van der Walt, M. Brett, J. Wilson, K. J. Millman, N. Mayorov, A. R. J. Nelson, E. Jones, R. Kern, E. Larson, C. J. Carey, İ. Polat, Y. Feng, E. W. Moore, J. VanderPlas, D. Laxalde, J. Perktold, R. Cimrman, I. Henriksen, E. A. Quintero, C. R. Harris, A. M. Archibald, A. H. Ribeiro, F. Pedregosa, P. van Mulbregt, and SciPy 1.0 Contributors, "SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python," *Nature Methods*, vol. 17, pp. 261–272, 2020.

[34] J. Nocedal and S. J. Wright, *Numerical Optimization.* Springer, 2006.

[35] J. NOCEDAL, "Updating quasi-newton matrices with limited storage," *Mathematics of Computation*, vol. 35, pp. 773–782, 01 1980.

[36] C. Zhu, R. Byrd, P. Lu, and J. Nocedal, "Algorithm 778: L-bfgs-b: Fortran subroutines for large-scale bound-constrained optimization," *ACM Trans. Math. Softw.*, vol. 23, pp. 550–560, 12 1997.

[37] F. Mazzelli, A. Little, S. Garimella, and Y. Bartosiewicz, "Computational and experimental analysis of supersonic air ejector: Turbulence modeling and assessment of 3d effects," *International Journal of Heat and Fluid Flow*, vol. 56, 12 2015.

[38] A. B. Downey, *Modeling and Simulation in Python.* No Starch Press, 2022.

[39] P. Virtanen, R. Gommers, T. E. Oliphant, M. Haberland, T. Reddy, D. Cournapeau, E. Burovski, P. Peterson, W. Weckesser, J. Bright, S. J. van der Walt, M. Brett, J. Wilson, K. J. Millman, N. Mayorov, A. R. J. Nelson, E. Jones, R. Kern, E. Larson, C. J. Carey, İ. Polat, Y. Feng, E. W. Moore, J. VanderPlas, D. Laxalde, J. Perktold, R. Cimrman, I. Henriksen, E. A. Quintero, C. R. Harris, A. M. Archibald, A. H. Ribeiro, F. Pedregosa, P. van Mulbregt, and SciPy 1.0 Contributors, "SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python," *Nature Methods*, vol. 17, pp. 261–272, 2020.

**UNIVERSITÉ CATHOLIQUE DE LOUVAIN**
**École polytechnique de Louvain**
Rue Archimède, 1 bte L6.11.01, 1348 Louvain-la-Neuve, Belgique | **www.uclouvain.be/epl**