Multiple Logistic Regression and Neural Networks as predictive tools for High Pressure Combustion Regimes

Jorge García-Rivas Carmona[a,b]

[a]*Instituto Superior Técnico, Universidade de Lisboa, Av. Rovisco Pais, 1049-001, Lisbon, Portugal*
[b]*Escuela Técnica Superior de Ingeniería Aeronáutica y del Espacio, Universidad Politécnica de Madrid, Pl. Cardenal Cisneros, 28040, Madrid, Spain*

Abstract

This paper presents a novel approach to study H_2/O_2/He high-pressure combustion in ESTHER shock-tube facility. ESTHER shock-tube provides filling pressures up to 100 bar, extreme initial conditions that have very few studies in the literature. Multiple Logistic Regression and Neural Networks are used to model different combustion regimes, among which detonation holds major interest for being non desirable in the shock-tube performance, as well as compromising for the integrity of the facility. The results of the statistical analysis of the sample of experiments available in ESTHER database shows that the dilution with inert gas, helium, is the leading parameter in detonation formation. The H_2/O_2 must be carefully diluted in order to avoid detonation risks but also to allow an efficient combustion. Another key factor in detonation phenomenon is the optical system configuration. The shock-tube uses an infrarred laser as external source of energy, which is also statistically decisive since one configuration provides a way sharper energy gradient. The results also provide regression models and a network structure to predict some of the variables involved in the combustion, such as pressure ratio or consumption velocity. These robust prediction tools can benefit ESTHER project allowing more efficient experiments, which will translate into time and expenses savings, along as improving themselves as statitiscal models thanks to each new experiment, which will feed the considered database in a feedback process.

Keywords: High-pressure combustion, Statistical analysis, ESTHER

1. Introduction

The use of hydrogen as fuel in ground power generation might be one of the solutions to face pollutants emission, since water is the only product of hydrogen-oxygen combustion. However, safety hazards in duct systems and the explosive nature of hydrogen-oxygen mixtures still pose significant technical challenges. The dilution of H_2/O_2 mixtures by inert gases such as helium, is one of the solutions to avoid explosive events. Yet, the ignition and propagation of premixed H_2/O_2/He flames at very high-pressure conditions, similar to those found inside ground power generation and internal combustion engines, is still poorly understood. Theoretical and several modelling strategies are used to understand the phenomena, yet the process is very complex and difficult to capture numerically and theoretically. As it will be discussed throughout this manuscript, new original strategies can be successfully applied to emulate ignition regimes even when theory and modelling are still not capable of fully describing ignition at high pressure conditions or are too computationally expensive.

The work developed in this thesis was motivated by the need of accurately predict the combustion regimes inside the combustion driver section of the new European Shock-Tube for High Enthalpy Research (ESTHER) facility of the European Space Agency (ESA). This facility is under final stage commissioning at IPFN/IST, and opens new perspectives to study the combustion phenomena at extremely high-pressure conditions. A prototype of the ESTHER combustion chamber was constructed to qualify both the gas filling system and the driver mechanical design, and nowadays is being used to study in detail deflagration and detonation regimes of H_2/O_2/He mixtures at gas filling pressures up to 100 bar.

It is, therefore, of major interest to predict the shot outcome regarding safety hazards for all the range of the experimental input parameters of ESTHER. This will help understand the physics and design combustion experiments at high pressure conditions as well as save expenses and time. In this novel approach to high-pressure combustion, Multiple Logistic Regression (MLR) and Artificial Neural Networks (ANNs) are employed as predicting tool for the combustion regime occuring inside the combustion driver section of the ESTHER facility.

Multiple Logistic Regression has been widely used during the past years in statistical studies such as malignant/cancerous cells detection [1] or demographic analysis [2], among other topics. Some examples of MLR applied to combustion field can be found in Meng *et al.* [3] and Wang *et al.* [4], who used this technique to study and measure CO_2 emissions and furnace combustion state, respectively, with successful results. On the other hand, the hitech neural networks tool is gaining importance in the 21st century as part of our society, accompanying technology as it evolves. Tosun *et al.* [5] compared Logistic Regression and Neural Networks methods in internal combustion diesel engine performances, concluding that the latter offers more accurate predictions for all the parameters consid-

ered. A similar study was carried out by Pino-Mejas *et al.* [6], who used both methods to predict energy consumption, energy demand and CO_2 emissions in Santiago, Chile. As many other authors of recent studies, Tosun [5] and Pino-Mejas [6] agree in the potential that these statistical algorithms have in the future of propulsion and energy engineering.

2. Experimental Data

2.1. Combustion-driver of ESTHER facility

The whole ESTHER experiment involves many processes and studies, yet this work focuses on the driver combustion section. A schematic representation of this shock-tube can be found in Fig. 1. The combustion takes place in a conventional combustion chamber where the H_2, O_2 and He are pumped in. Once the mixture becomes homogeneous inside the combustion chamber, an infrared laser provides the energy needed to trigger the ignition. The pressure rise is meant to burst a diaphragm that connects the chamber with an intermediate compression tube, filled with helium. This compression tube is connected to the shock-tube test section through a second diaphragm that will also burst when reached by the shock wave. At the end of the shock tube section, a dump tank recovers the H_2O resulting from the combustion process and any gas steam flow. A complete description of the system and experimental setup can be found in Lino da Silva et al. [7].

2.2. Independent and Dependent Variables

ESTHER facility holds a relatively new experiment, and so far the experimental campaign includes around 180 shots. This sample offers a large number of data regarding the combustion, however one should be cautious on the selection of variables that will be part of the statistic analysis. Therefore, a first step is to evaluate the problem variables and to try to reduce the complexity of the system. The main difference between the two sets of variables in the combustion problem is that independent variables are set by the investigators, as inputs of the experiments, whereas dependent variables are the outcome of the experiment, meaning that they are completely unknown until the experiment has been performed. The statistical analysis aims at predicting these dependent variables, in order to increase experimental efficiency and to develop physical models that will hopefully mimic the results observed in high-pressure ignition experiments of ESTHER facility.

The experimental variables considered for this statistical analysis are gathered in the next sections.

2.2.1. Independent Variables

- Initial temperature (T_0): refers to the gas temperature inside the combustion chamber at the beginning of the shot. It is equal to the ambient temperature in the laboratory.
- Initial pressure (p_0): the gas pressure inside the combustion chamber. It takes into account the ambient pressure.
- Equivalence ratio (ϕ): by definition, the equivalence ratio of a mixture is given by

$$\phi = \frac{n_{fuel}/n_{oxidizer}}{(\frac{n_{fuel}}{n_{oxidizer}})_{st}} \tag{1}$$

  where the subindex *st* refers to the stoichiometric mixture. The hydrogen/oxygen mixture combustion reaction is given by

$$H_2 + \frac{1}{2}O_2 \rightarrow H_2O \tag{2}$$

  so the equivalence ratio of the gas mixture in ESTHER is given by the formula

$$\phi = \frac{n_{H_2}}{2n_{O_2}} \tag{3}$$

- Helium dilution factor (*dil*): this variable makes reference to the amount of diluent gas, He, within the total mixture. Helium is necessary to absorb heat from the hydrogen/oxygen combustion and to avoid high risks of detonations. It is defined as

$$dil = \frac{n_{He}}{n_{total}} = \frac{n_{He}}{n_{O_2} + n_{H_2} + n_{He}} \tag{4}$$

- Ignition source, (*lens*): ESTHER uses an infrared laser as external source of energy. This variable represents the two different configurations of this optical system, namely without lens, so the laser provides energy to a relatively wide volume of gas mixture, and with lens, so the laser is focalized onto a point of the gas mixture. This last configuration provides a sharper energy gradient.

2.2.2. Dependent variables

- Ignition regime: this is the most important variable in the statistical study, and represents how the combustion evolved after the ignition of the mixture. It involves two different phenomena: detonation and dual slope. Detonation occurs when the flame front velocity reaches sonic conditions. If this flame front velocity is subsonic, ignition is associated with a deflagration. The dual slope is a phenomenon that can be recognized in the combustion pressure-time curve, and is related to tulip flames phenomena, as reported by Dunn-Rankin and Sawyer [8]. These two phenomena provide four different ignition regimes: single slope deflagration, dual slope deflagration, single slope detonation and dual slope detonation.
- Pressure ratio (p_{ratio}): the pressure ratio stands for the ratio between the maximum and initial value of the gas pressure recorded during the combustion, as given by

$$p_{ratio} = \frac{p_{max}}{p_0} \tag{5}$$

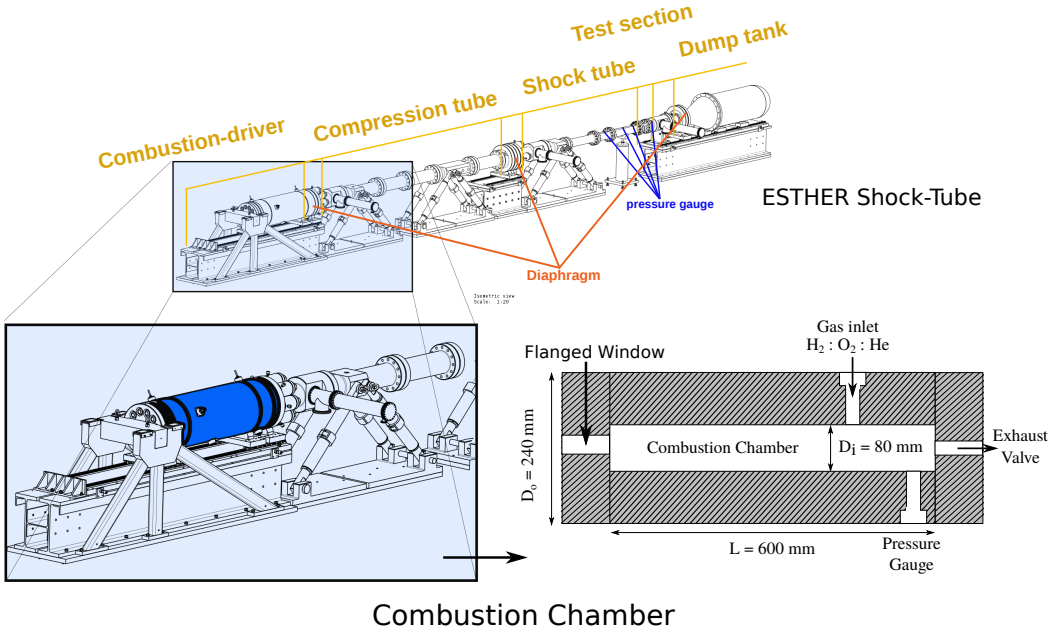

Figure 1: Schematic representation of the closed combustion chamber.

- Consumption velocity (v_{ch}): this variable represents the gas consumption velocity. It is defined in terms of the length of the combustion chamber (l_{ch}) and the elapsed time between initial gas pressure (when the laser is turned on) and maximum gas pressure (Δt_{ig}), as given by

$$v_{ch} = \frac{l_{ch}}{\Delta t_{ig}} \tag{6}$$

where $l_{ch} = 0,6\ m$.

3. Statistical Methods

3.1. Multiple Logistic Regression model

Multiple Logistic Regression (MLR) is a supervised statistical method that aims to model the probability of the dependent variables to occur, in case they are dichotomous or discrete, and to return an absolute value, if the dependent variable is a continuous magnitude, as function of the considered independent variables. Thus, the detonation and dual slope phenomena are predicted according to the following equation.

$$P(y) = g\left(\theta_0 + \sum_{j=1}^{n} \theta_j X_j\right) \tag{7}$$

where $P(y)$ is the probability of the phenomenon y to occur. X_i are the input variables of the model (independent variables), so $n = 5$ and θ_j are the weight coefficients assigned to each variable. θ_0 is called intercept term (bias). The function g is the sigmoid function, a special form of the logistic function, and makes the result of the regression belong in the interval of 0 and 1. This sigmoid function is given by

$$g(x) = \frac{e^x}{1 + e^x} \tag{8}$$

Otherwise, the dependent variables pressure ratio and consumption velocity can not be modelled as probabilities and hence multiple linear regression is used instead MLR. This different regression is given by the following expression.

$$y = \theta_0 + \sum_{j=1}^{n} \theta_j X_j \tag{9}$$

The variables in this equation have the same meaning to those seen in Eq. (7)This kind of regression does not use the sigmoid function, so the result belongs in the real numbers continuum.

This method iteratively updates the θ_j coefficients following the gradient descent method, which aims to reduce the cost function ($J(\theta)$). This function represents how difference are the current regression predictions and the real dependent variable to be predict, as given by

$$J(\theta) = \frac{1}{m} \sum_{i=1}^{m} \left[y^{(i)} \log\left(h_\theta\left(x^{(i)}\right)\right) + \left(\left(1 - y^{(i)}\right)\right) \log\left(1 - h_\theta\left(x^{(i)}\right)\right)\right] \tag{10}$$

where m is the number of experiments (combustion points) considered in the analysis, $m = 52$. $y^{(i)}$ refers to the value of the dependent variable of point i and $h_\theta\left(x^{(i)}\right)$ is the hypothesis, i.e. the prediction of $y^{(i)}$ using the current θ_j coefficients, as given by Eq.(7) or Eq. (9), depending on the dependent variable. The θ_j coefficients are assigned stepwise according to the

following rule.

$$\theta_j = \theta_j - \alpha \frac{\partial}{\partial \theta_j} J(\theta) \tag{11}$$

where α is a parameter that represents how big is the variation of the different θ_j coefficients at each iteration. It has been proven that, for normalized variables, the value $\alpha = 0,1$ offers a good compromise regarding computational speed and convergence efficiency [9].

3.2. Artificial Neural Networks model

An Artificial Neural Network is a mathematical model which is inspired by the structure and functional aspects of a biological nervous system. It is a very powerful state-of-the-art tool in statistical analysis, which is starting to be widely used for a lot of different applications, setting the base of the so-called machine learning, branch of technology dedicated to make the computers capable of learning and improving. Neural networks can be roughly described as many MLR concatenated and linked between them, being capable to building stronger links between the provided variables in order to predict the result with high level of accuracy.

The neural network structure adopted in this work is the Multiple Layer Perceptron. MLP networks have an input layer, followed by one or more hidden layers and an output. Just one hidden layer is considered in this analysis, as shown by Fig. 2.

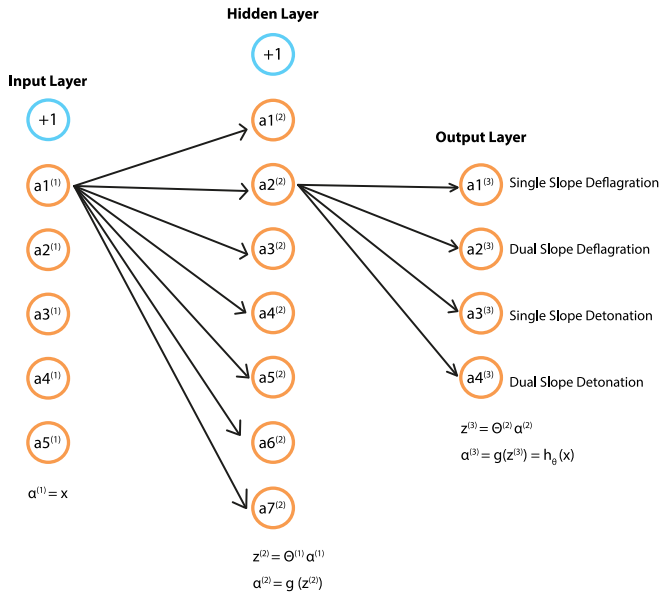

Figure 2: Neural network structure.

Figure 2 schematically represents the neural network developed in this project. Each layer has some artificial neurons (nodes), represented as $as_l^{(l)}$ (where s_l refers to the number of nodes in layer l), a weight matrix (Θ) that includes the layer's intercept term (given by the +1 nodes), and an output vector. This ouput vector is computed according the equations in the figure, where $z^{(l)}$ is the vector of nodes in layer l before undergoing the sigmoid transformation, so $g\left(z^{(l)}\right) = a^{(l)}$ The input layer receives the independent variables as inputs for the neural network, whereas the output layer returns the results of the predictions, the output of network. The hidden layer is the one in between, which has no particular physical meaning. The developed neural network faces a multiclass classification of the variable ignition regime, so the network's output is the probability of each possible ignition regime to occur. Therefore, the output layer vector is given by the following expression:

$$f(T_0, p_0, \phi, dil, lens)_k = g\left(\Theta_{0k}^{(2)} + \sum_{s=1}^{S} \Theta_{si}^{(2)} g\left(\Theta_{0s}^{(1)} + \sum_{j=1}^{n} \Theta_{jk}^{(1)} X_j\right)\right) \tag{12}$$

where i refers to each possible output (the different ignition regimes) and S is the number of nodes considered in the hidden layer. The subindices of the Θ matrices represent, respectively, the row and column within the matrix whereas the superindices refers to the layer which the Θ matrix is assigned to. The similarities between Eq. (7) and Eq. (12) are notable, which remarks the fact that a neural network can be modelled as many MLR. The learning method used by neural networks is known as backpropagation algorithm, and has been proven to provide robust results for analyzing complex problems. It is based on the backward propagation (from output layer to input layer) of the weight coefficients updating. The learning algorithm used in this work is backpropagation gradient descent with regularization factor (λ). This regularization factor adds a term to the cost function in order to prevent overfitting [9], as given by:

$$\begin{aligned} J(\Theta) = -\frac{1}{m}\left[\sum_{i=1}^{m}\sum_{k=1}^{K} y_k^{(i)} \log\left(\left(h_\Theta\left(x^{(i)}\right)\right)_k\right) + \left(1 - y_k^{(i)}\right) \log\left(\left(1 - h_\Theta\left(x^{(i)}\right)\right)_k\right)\right] \\ + \frac{\lambda}{2m}\sum_{l=1}^{L-1}\sum_{i=1}^{s_l}\sum_{j=1}^{s_{l+1}}\left(\Theta_{ji}^{(l)}\right)^2 \end{aligned} \tag{13}$$

where K represents the number of possible output units, i.e. the possible ignition regimes (four possible regimes are eligibles, so $K = 4$). Lastly, L represents the number of layers. As seen in Eq. (10), $y_k^{(i)}$ is the value od the dependent variable ignition regime of experiment i, regarding each possible regime (so $y_k^{(i)} = 1$ if the regime k was returned by experiment i, and $y_k^{(i)} = 0$ otherwise). $\left(h_\Theta\left(x^{(i)}\right)\right)$ is the prediction hypothesis, which refers to the prediction of $y_k^{(i)}$ using the Θ coefficients matrices, as given by Eq. (12). Although Eq. (13) looks more complex, it has the same function of Eq. (10). The updating of the Θ matrices follows the same rule of Eq. (11), yet the gradient of the cost function is obtained with the backpropagation technique. At each iteration, the residuals δ of each node are computed following the expressions.

$$\delta^{(L)} = a(L) - y^{(i)} \tag{14}$$

$$\delta^{(l-1)} = \Theta^{(l-1)} \delta^{(l)} g'\left(z^{(l-1)}\right) \tag{15}$$

where l is the layer number and L refers to the output layer. $a^{(L)}$ is the vector of nodes in layer L values, as given by Eq. (12). g'

is the gradient of the sigmoid function. Once computed, the cost function gradients are updated according the following rule.

$$\Delta_{ij}^{(l)} := \Delta_{ij}^{(l)} + a_j^{(l)} \delta_i^{(l+1)} \tag{16}$$

$$\frac{\partial}{\partial \Theta_{ij}^{(l)}} J(\Theta) := \frac{1}{m} \Delta_{ij}^{(l)} + \lambda \Theta_{ij}^{(l)} \tag{17}$$

where $\Delta_{ij}^{(l)}$ represents the accumulative error of each node in the network.

4. Results and Discussion

Before performing the statistical analysis, the experimental variables are normalized according to the following expression:

$$x' := \frac{x - \mu_i}{s_i} \tag{18}$$

where μ_i is the mean value of the variable i and s_i is the considered range of this variable i ($s_i = x_{i_{max}} - x_{i_{min}}$). Table 1 shows the μ_i and s_i values of each independent variable.

Variable	μ_i		s_i	
Initial temperature, T_0	19,9	C	14,0	C
Initial gas pressure,p_0	38,2	bar	88,0	bar
Equivalence ratio, ϕ	0,95		1,00	
Helium dilution, *dil*	0,72		0,20	
Use of lens, *lens*	0,44		1	

Table 1: Normalization parameters.

4.1. Ignition Regime

The results of the Multiple Logistic Regression analysis for the variable ignition regime, which differenciates detonation and dual slope phenomena, are shown in Table 2. The table provides the θ coefficients as well as the p-values of each independent variable. These p-values help determine the significance of the results in the Null hypothesis tests (Null hypothesis: the independent variable has no correlation with the dependent variable). A small p-value (typically $\leq$ 0,05) indicates strong evidence against the null hypothesis, which means that the independent variable is statistically correlated to the dependent variable. Conversely, a large p-value (> 0,05) indicates weak evidence against the null hypothesis, so the latter applies, which means that the independent variable does not correlate with the dependent variable, from a statistical standpoint.

In addition to the results on Table 2, Eq. (19) and Eq. (20) show the regression equations obtained after the MLR analysis.

$$P(Detonation) = g\left(-1,88 - 0,31T_0' - 0,16p_0' - 1,92\phi' - 4,99dil' + 0,79lens'\right) \tag{19}$$

$$P(DualSlope) = g\left(-3,10 - 6,90T_0' - 2,60p_0' - 0,18\phi' + 9,40dil' - 3,90lens'\right) \tag{20}$$

	Detonation		Dual slope	
Variable	θ_i	p-values	θ_i	p-values
Intercept term	-1,88	10^{-5}	-3,10	10^{-3}
Initial temperature, T_0'	-0,31	0,84	-6,90	0,02
Initial gas pressure, p_0'	-0,16	0,93	-2,60	0,37
Equivalence ratio, ϕ'	-1,92	0,30	-0,18	0,96
Helium dilution , dil'	-4,99	0,05	9,40	0,18
Use of lens, *lens*′	0,79	0,49	-3,90	0,02

Table 2: MLR coefficients for dependent variables Detonation and Dual Slope.

On the other hand, Fig. 3 shows the deflagration/detonation regime probability distribution returned by the analysis with Neural Networks. The range considered for initial pressure is relatively small compared with the rest of experimental variables. in addition, MLR results show that it is poorly correlated with detontaion (p-value = 0,84). Consequently, the influence of this variable is not discussed in this abstract.

4.1.1. Effect of dilution by Helium

As given by Table 2, helium dilution (p-value = 0,05) is the most important value to statistically predict whether detonation or deflagration occur. Indeed, values of helium dilution below the average (*dil* = 0,72) lead the regression likely to predict detonation, as can be inferred from the assigned coefficient (θ_i = -4,99). The coefficient assigned to *dil*, θ_i = -4,99, is the largest within the estimated coefficients, yet the variations of helium dilution are usually small with respect to the mean value (*dil* = 0,72). Therefore, an increase of 0,01 in the dilution factor means that the prediction of detonation decreases 3,9% (Δdil = 0,01 $\rightarrow$ ΔP(Detonation) = -3,9%). Considering Eq. (20), although the helium dilution p-value is above the 0,05 threshold, its assigned coefficient (θ_i = 9,40) is large enough to induce important changes in the dual slope predictions. In this case, high dilutions factors trigger the dual slope phenomena. The estimated coefficient (θ_i = 9,40) means an increase of 3,2% per 0,01 dilution factor (Δdil = 0,01 $\rightarrow$ ΔP(Dual Slope) = 3,2%).

Figures 3a) and 3b) show the influence of helium dilution in the probability of an deflagration/detonation occurrence. Figure 3b) shows a monotonic decrease of the single slope detonation probability from near 100% to 30% in the range of dilution factors within the interval of 0,57 and 0,69. For higher dilutions values (from 0,69 up to 0,77), the ignition regime most likely to occur when lens are not used is the single slope detonation (Fig. 3). Moreover, it can be appreciated that both regimes involving dual slope (dual slope deflagration and dual slope detonation) present an increase for dilution factors above 0,72. The change in the optical configuration enhances the results observed in the case without lens. Single slope detonation probability is enhanced for the range of *dil* considered. From *dil* = 0,68 up to *dil* = 0,77, single slope detonation probability monotonically decreases from 95% to 15%. These results completely agree with those seen in the MLR analysis, so it can be concluded that low values of helium dilution (*dil* < 0,70) does not provide the heat absorption necessary to avoid detonations of the

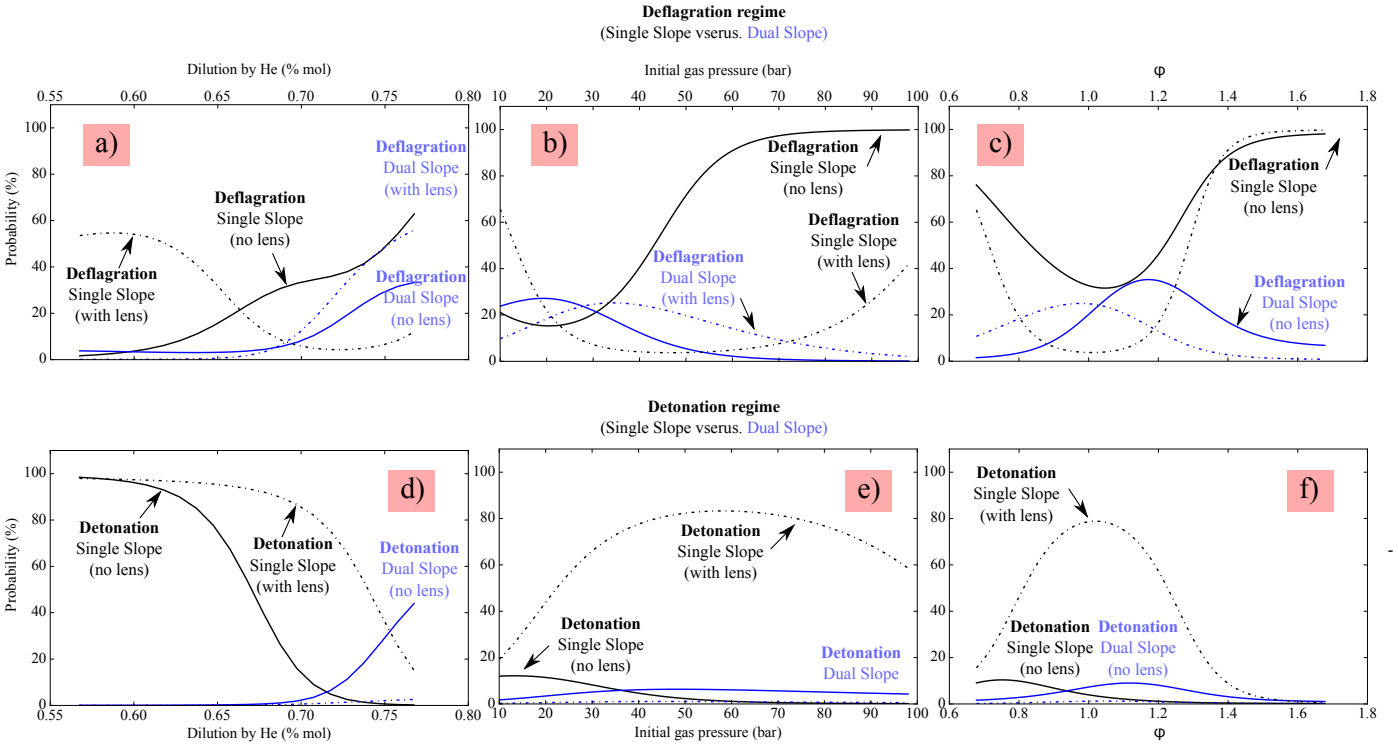

Figure 3: Deflagration regime probability considering both ignition sources and single/dual slope phenomenon in average conditions, regarding a) dilution by He, b) initial gas pressure and c) equivalence ratio. Detonation regime probability considering both ignition sources and single/dual slope phenomenon in average conditions, regarding d) dilution by He, e) initial gas pressure and f) equivalence ratio.

mixture. It is hence concluded that helium dilution is the independent variable statistically more correlated with detonation predictions.

4.1.2. Effect of initial gas pressure

Initial gas pressure (p-value = 0,93) are not strongly correlated to the event of a detonation, according to Table. 2. Unitary increases of initial gas pressure induce decreases in the probability of detonation to occur of Δp_0 = 1 bar → ΔP(Detonation) = -0,02%, while the rest of variables remain constant. Conversely, dual slope formation seems to be slightly dependent on the initial gas pressure (p-value = 0,37). The negative θ coefficient ($\theta_i = -2,60$) means that an unitary increase in these independent variables, while the rest of variables remain constant, provokes a decrease in the probability of dual slope of 0,4% (Δp_0 = 1 bar → ΔP(Dual Slope) = -0,4%).

In Figs. 3b) and 3d), the influence of the gas filling pressures on deflagration/detonation regime prediction is shown. Figure 3b) shows a monotonic increase of the single slope deflagration probability from 30% to 100% within the interval of 35 and 100 bar gas filling pressures. For p_0 below 35 bar, with a probability of 35%, the ignition regime most likely to occur is dual slope deflagration. However, the statistical analysis is not conclusive for these low values of the gas filling pressures without lens. Nevertheless, more consistent statistical results can be observed when the shot is performed with lens. The results on Fig. 3d) show that for a wide range of gas filling pressures (from 30 up to 100 bar), the ignition regime is more likely to be a single slope detonation, with a probability of around 80%. These results show that, from a statistical standpoint, for any gas pressure higher than 30 bar the ignition regime of an average shot with lens will be, with 80% probability, a single slope detonation, which means that, to change the ignition regime, is more relevant to change the optical system configuration (with or without lens) than varying the gas filling pressure. Nevertheless, this conclusion is only valid if the independent variables are in the vicinity of their respective average values.

4.1.3. Effect of mixture equivalence ratio

Table 2 shows that equivalence ratio is slightly more correlated than initial gas pressure, yet its p-value is still above the 0,05 threshold to validate the strong correlation. As seen for initial gas pressure, unitary increases of mixture equivalence ratio induce decreases in the probability of detonation ($\Delta\phi$ = 0,1 → ΔP(Detonation) = -1,2%), as given by the assigned θ coefficient (θ_i = -1,92). Conversely, dual slope formation seems to be statistically independent of the mixture equivalence ratio (p-value = 0,96). An increase of 0,1 in equivalence ratio means a decrease of 0,3% in dual slope probability ($\Delta\phi$ = 0,1 → ΔP(Dual Slope) = -0,3%) for average conditions, according to these regression models. However, their linear nature did not capture a possible non-linear evolution around the stoichiometric mixture.

The results shown in Figs. 3c) and 3f) are very expressive. In

Fig. 3c) the estimated probabilities show that, if the shot is performed without lens, for all the range of ϕ, the most probable ignition regime is the single slope deflagration. Yet, the single slope deflagration probability distribution with ϕ is not monotonic and reaches a minimum of 40% when the dual slope deflagration probability distribution reaches a peak, for values of ϕ in the range of 1 and 1,3. For equivalence ratios from 1,3 up to 1,7, the probability of single slope deflagration rises near 100%. The results on Fig. 3f) show that detonation has a probability of less than 10% for all the range of ϕ values. These results completely change when the shot is performed with lens. In fact, there is a high probability of single slope detonation for values of ϕ near the stoichiometric value (ϕ = 1); in the vicinity of ϕ = 1 the probability of detonation is about 80%. Moreover, the single slope deflagration regime is now shifted towards higher values of ϕ whereas near ϕ = 1 the probability of this regime is less than 5%. It is also interesting to note the quadratic evolution of the single slope detonation probability curve around the stoichiometric value. In fact, this result is explained by the physics of the problem. For mixtures with equivalence ratio ϕ around 1 the transfer of chemical energy into heat is more efficient and therefore the gas temperature increase is higher than that obtained for values of ϕ outside the vicinity of the value one. When the mixture is poor (excess of air/O_2, $\phi << 1$) or rich (excess of fuel, $\phi >> 1$) part of the chemical energy released during the combustion will be absorbed by the excess of air/O_2 or fuel/H_2 molecules and will be more difficult to trigger the mixture ignition. As observed in Figs. 3c) and3f), close to ϕ = 1 the ignition process is more efficient and faster, which may lead to a detonation regime. This also explains the differences observed for the case with lens and without lens. Without lens, the laser energy is deposited inside a volume larger than that when lens are used, which explains the higher detonation probability in this last case. The equivalence ratio enhances the energy density differences at the vicinity of ϕ = 1. This quadratic dependency could not have been captured by the linear regression used in the MLR method, although from a physical standpoint it should be modelled.

4.2. Pressure Ratio

Pressure ratio is a continuous variable and hence it is not a phenomenon that can happen or not. Multiple Linear Regression is used to predict the presumed pressure ratio returned by certain onset. This pressure ratio is therefore a non normalized discrete value. The regression values returned by the Multiple Linear Regression analysis is given by Table 3.

The regression equation for the dependent variable pressure ratio, which uses the coefficients θ_i seen in Table 3, is given by

$$p_{ratio} = 7,01 + 0,20T_0' - 1,21p_0' - 1,37\phi' - 5,11dil' + 0,41lens' \quad (21)$$

The p-values in Table 3 show that initial temperature (p-value = 0,74) and use of lens (p-value = 0,34) are poorly correlated with pressure ratio. Moreover, initial gas pressure (p-value = 0,14) and mixture equivalence ratio (p-value = 0,12) are slightly influent in the pressure ratio prediction. Conversely, helium dilution stands as the most statistically significant variable (p-value = 10^{-5}). The intercept term (θ_i = 7,01) shows that, for the average experimental onset, the pressure ratio prediction is 6,83 without lens (*lens* = 0 → p_{ratio} = 6,83) and 7,24 with lens (*lens* = 1 → p_{ratio} = 7,24). The use of lens slightly increases the pressure ratio prediction, but as observed in Table 3, the independent variable *lens* is poorly correlated with the pressure ratio (p-value = 0,34). Initial temperature (θ_i = 0,20), initial gas pressure (θ_i = -1,21) and equivalence ratio (θ_i = -1,37) are poorly correlated with pressure ratio. The mean increase produced in pressure ratio by an unitary increase of T_0 is 0,01 (ΔT_0 = 1C → Δp_{ratio} = 0,01), whereas an unitary rise in p_0 induces a decrease of -0,01 in pressure ratio (Δp_0 = 1 bar → Δp_{ratio} = 0,01) and an ϕ increase of 0,01 means a decrease of 0,12 ($\Delta\phi$ = 0,1 → Δp_{ratio} = -0,12). Unitary increases in helium dilution, which has been concluded as the most relevant independent variable, provokes a decrease in the pressure ratio prediction of 0,24 (Δdil = 0,01 → Δp_{ratio} = 0,24).

Variable	θ_i	p-values
Intercept term	7,01	0,00
Initial temperature, T_0	0,20	0,74
Initial gas pressure, p_0	-1,21	0,14
Equivalence ratio, ϕ	-1,37	0,12
Helium dilution, *dil*	-5,11	10^{-5}
Use of lens, *lens*	0,41	0,34

Table 3: MLR results for the dependent variable pressure ratio.

It can be concluded that dilution factor controls the strength of the ignition of the volatile H_2/O_2 mixture, and hence the pressure ratio obtained. As discussed in Castela et al. [10], the pressure ratio around eight is the theoretical limit between deflagration and detonation. Being so, regressions for Detonation and pressure ratio should present similarities. Indeed, both are very similar; the initial temperature has opposite influence yet very small, but the rest of independent variables present the same symbols in the respective coefficients, which means same influences.

4.3. Consumption velocity

Table 4 shows the coefficients returned by the regression analysis of the variable consumption velocity.

The consumption velocity regression equation is given by Eq. (22).

$$v_{ch} = 24,39 + 6,93T_0' - 44,37p_0' - 27,53\phi' - 105,92dil' - 2,92lens' \quad (22)$$

As stated in Table 4, the regression coefficients θ_i have velocity units, m/s, because consumption velocity is not normalized. The p-values shown in Table 4 reveal that the consumption velocity is the dependent variable more influenced by the experi-

Variable	θ_i		p-values
Intercept term	24,39	m/s	0,00
Initial temperature, T_0	6,93	m/s	0,37
Initial gas pressure, p_0	-44,37	m/s	10^{-5}
Equivalence ratio, ϕ	-27,53	m/s	0,01
Helium dilution, *dil*	-105,92	m/s	0,00
Use of lens, *lens*	-2,92	m/s	0,58

Table 4: MLR results for the dependent variable consumption velocity.

mental onset, from a statistical standpoint. Initial gas pressure (p-value = 10^{-5}), mixture equivalence ratio (p-value = 0,01) and helium dilution factor (p-value = 0,00) play a significant role as predictors of the mixture consumption velocity. As seen in the pressure ratio regression analysis (see Table 3), low gas pressures, low equivalence ratios and low dilutions produce higher velocity of mixture consumption. The average experimental onset returns a consumption velocity of 25,7 m/s without lens ($lens = 0 \rightarrow v_{ch}$ = 25,7 m/s) and 22,8 m/s with lens ($lens = 1 \rightarrow v_{ch}$ = 22,8 m/s). Initial temperature coefficient (θ_i = 6,93) induces a rise of 0,5 m/s if T_0 increases 1 C (ΔT_0 = 1 C $\rightarrow \Delta v_{ch}$ = 0,5 m/s). On the other hand, p_0, ϕ and *dil* are strongly correlated with the consumption velocity, as observed in their p-values (Table 4). An increase of 1 bar in p_0 means a mean decrease of 0,5 m/s in the consumption velocity (Δp_0 = 1 bar $\rightarrow$ Δv_{ch} = -0,5 m/s) and an unitary increase of ϕ causes a decrease of 2,5 m/s in the consumption velocity prediction ($\Delta\phi$ = 0,1 $\rightarrow$ Δv_{ch} = -2,5 m/s). Dilution factor also correlates with consumption velocity inversely. An increase of 0,01 in *dil*, while the rest of independent variables do not vary, means a decrease of 5,0 m/s in the consumption velocity (Δdil = 0,01 $\rightarrow \Delta v_{ch}$ = -5,0 m/s).

The regression returned by the MLR analysis has the same influence as the pressure ratio regression, which matches the intuition about higher pressure ratios involving higher consumption velocities. The only difference remains in the feature lens, since the focalization of the laser leads to lower consumption velocities according to this model, whereas produced higher pressure ratios. This could be a link with the tulip flames formation, and hence with dual slope phenomena.

5. Conclusions

High-pressure laser-induced combustion has been studied from a statistical standpoint using Multiple Logistic Regression and Neural Networks, in the scope of the ESTHER project development, which will provide support to Space exploration by validating spacecraft thermal protections. Regarding the variables considered for the detonation regime determination, the following conclusions were extracted from the MLR and Neural Network analysis:

- The initial gas pressure has been found not to be as relevant in the regime obtained after the shot as was firstly thought. It has moderate influence in the pressure ratio obtained, and it finds its strongest correlation with the consumption velocity. Both initial temperature and initial gas pressure, have opposite influences in pressure ratio and consumption velocity prediction, which makes us think that the gas mixture density could be a key factor in the propagation of the flame front and, hence, the formation of dual slope pressure-time curves.

- The equivalence ratio of the mixture is also important in the detonation process. Although the linear character of MLR shows a negative influence (so lean mixtures would produce higher detonation probabilities) the results provided by the neural network have shown that the influence of this independent variable is, indeed, non linear. The combustion efficiency reaches its maximum with the stoichiometric H2/O2 mixture and slightly richer mixtures. Higher combustion efficiency means that increases in pressure and temperature are achieved in the chamber after ignition.

- Both MLR and Neural Networks concluded that dilution is a key factor in the outcome of the combustion. The explosive behavior of the H2/O2 mixture makes helium completely necessary, since it has been proven that low helium dilution factors automatically lead to detonation. This dilution is, statistically speaking, the most important variable for detonation formation, as well as for the consumption velocity and the pressure peak achieved.

- Regarding the ignition source, the lens used to focalize the laser into a single point drastically changes the behavior of the combustion. The sharper temperature gradient induced often leads to instabillity and turbulence, which causes strong deflagration and detonations. The interaction between laser and the reactive gas mixture is still to be fully understood under extreme conditions. Yet, the statistical analysis has revealed that the focalized mode triggers detonation easily. Also, the ignition source is strongly related with the formation of tulip flames, taking into account that the initial temperature gradient determines the first stadiums in the development of the flame.

The multi-physics that drive ESTHER shock-tube combustion have been studied using statistical methods, which returned large data about the different interactions between the variables involved in the combustion and its possible outcome. This data provides the first step into the understanding of such a complex process and improvements has been achieved in this analysis. The implementation of the developed prediction tools in ESTHER shock tube will also allow more efficient shots, which will provide time and expenses savings as consequence.

6. Acknowledgment

IPFN activities received financial support from European Space Agency through contract 23086 “Kinetic Shock-Tube for Plan-

etary Exploration", and from Fundacão para a Ciência e Tecnologia through project UID/FIS/50010/2013.

References

[1] A. Tailor, D. Jurkovic, T. H. Bourne, W. P. Collins, S. Campbell, Sonographic prediction of malignancy in adnexal masses using multivariate logistic regression analysis, Ultrasound Obstet. Gynecol. 10 (1997) 41–47.

[2] A. A. A. Alsharif, B. Pradhan, Urban sprawl analysis of tripoli metropolitan city (Libya) using remote sensing data and multivariate logistic regression model, Journal of the Indian Society of Remote Sensing 42 (2014) 149–163.

[3] M. Meng, D. Niu, Modeling CO_2 emissions from fossil fuel combustion using the logistic equation, Energy 36 (2011) 3355–3359. 10.1016/j.energy.2011.03.032.

[4] Z. Wang, C. Song, T. Chen, Deep learning based monitoring of furnace combustion state and measurement of heat release rate, Energy 131 (2017) 106–112. 10.1016/j.energy.2017.05.012.

[5] E. Tosun, K. Aydin, M. Bilgili, Comparison of linear regression and artificial neural network model of a diesel engine fueled with biodiesel-alcohol mixtures, Alexandria Engineering Journal 55 (2016) 3081–3089. 10.1016/j.aej.2016.08.011.

[6] R. Pino-Mejas, A. Prez-Fargallo, C. Rubio-Bellido, J. A. Pulido-Arcas, Comparison of linear regression and artificial neural networks models to predict heating and cooling energy demand, energy consumption and CO_2 emissions, Energy 118 (2017) 24–36. 10.1016/j.energy.2016.12.022.

[7] M. Lino da Silva, B. Brotas de Carvalho, A. Smith, A. Marraffa, High-pressure h_2/he/o_2 combustion experiments for the design of the esther shock-tube driver, in: 46th AIAA Thermophysics Conference, 2016, pp. 2016–4156.

[8] D. Dunn-Rankin, R. F. Sawyer, Tulip flames: changes in shape of premixed flames propagating in closed tubes, Experiments in Fluids 24 (1998) 130–140.

[9] A. Ng, Machine learning, Coursera, N/A.

[10] M. L. Castela, M. Levy, M. L. da Silva, B. B. de Carvalho, A. Smith, L. Marraffa, Ignition of H_2/O_2/He mixtures in a high-pressure (up to 50 bar) bomb vessel, 8th European Combustion Meeting (2017).