

POLITECNICO DI TORINO

Corso di Laurea Magistrale
in Automotive Engineering

Tesi di Laurea Magistrale

**Use of Machine Learning techniques and Neural Network
algorithms for the Laminar Burning Speed estimation**



Relatori

Daniela Misul

Mirko Baratta

Candidato

Salvatore Procopio

Anno Accademico 2017/2018

Index

Abstract	1
1. Introduction	2
1.1 Premixed turbulent combustion	3
1.2 Laminar burning speed existing models.....	4
1.3 The hydrogen contribution to CNG engines	12
2. Fundamentals of Machine Learning and Artificial Neural Networks.....	15
2.1 The different applications of Machine Learning	15
2.2 Artificial neural networks (ANN)	16
3. Laminar burning speed modelling	18
3.1 Modelling through Machine Learning.....	18
3.1.1 Flow-chart approach	18
3.2 Modelling through Neural Networks	28
4. Laminar burning speed results.....	32
4.1 Laminar burning speed vs. Fuel-air Equivalence ratio with Rational Quadratic GPR method	32
4.2 Laminar burning speed vs. pressure with Rational Quadratic GPR method.....	39
4.3 Laminar burning speed vs. Temperature with Rational Quadratic GPR method	40
4.4 Laminar burning speed vs. EGR with Rational Quadratic GPR method	41
4.5 Laminar burning speed vs. Fuel-air Equivalence ratio with ANN.....	42
4.6 Laminar burning speed vs. pressure with ANN	45
4.7 Laminar burning speed vs. Temperature with ANN.....	45
4.8 Laminar burning speed vs. EGR with ANN	49
4.9 Laminar burning speed vs. crank angle with Rational Quadratic GPR method	50
5. Conclusion and future development	56
Nomenclature	58
Subscripts.....	59
Bibliography	60

Figures index

Figure 1.1.1, <i>Representation of SI engine combustion event</i>	3
Figure 1.1.2, <i>SL vs equivalence ratio plot for an SI ICE fuelled by CNG</i>	4
Figure 1.2.1, <i>Magnification of laminar flame front</i>	6
Figure 1.2.2, <i>Representation of SI engine combustion event according to the Multi-zone model; V_b is showing the burned gas volume, instead V_u is the unburned gas volume. The point 1 is where the combustion starts thanks to the triggering event</i>	8
Figure 1.3.1, <i>Laminar burning speed vs equivalence ratio at 400 K, 0.2 MPa for three different cases</i>	14
Figure 2.1.1 <i>Iterative process to achieve satisfactory performance</i>	15
Figure 2.2.1, <i>Representation of a simple neural network</i>	17
Figure 3.1.1, <i>Partitioning of data according to the Holdout validation approach</i>	20
Figure 3.1.2, <i>Table provided by the Matlab App in which all the possible trained regression models are shown</i>	22
Figure 3.1.3, <i>Comparison between the available and predicted S_L for the first Trained model</i>	24
Figure 3.1.4, <i>S_L vs. equivalence ratio ϕ for the first trained model</i>	25
Figure 3.1.5, <i>Regression plot for the first Trained mode l</i>	25
Figure 3.1.6, <i>Comparison between the available and predicted S_L for the second Trained Model</i> ..	26
Figure 3.1.7, <i>S_L vs. equivalence ratio ϕ for the second Trained model</i>	26
Figure 3.1.8, <i>Regression plot for the second Trained model</i>	27
Figure 3.2.1, <i>For the first attempt design: NN performance and computation timing with Testing at 50%, Validation at 25% and Testing at 25%</i>	29
Figure 3.2.2, <i>For the first attempt design: a) NN performance and computation timing with Testing at 40%, Validation at 35% and Testing at 25%; b) NN performance and computation timing with Testing at 30%, Validation at 35% and Testing at 35%</i>	30
Figure 3.2.3, <i>For the second analysis: NN performance and computation timing with Testing at 30%, Validation at 35% and Testing at 35%</i>	31
Figure 4.1.1, <i>Laminar burning speed vs. Equivalence ratio: a) pure methane EGR 0%, b) pure methane 10% EGR, c) pure methane 20% EGR, d) comparison among previous cases</i>	34

Figure 4.1.2, <i>Laminar burning speed vs. equivalence ratio at 800 K: a) pure methane mixture with 0% EGR, b) comparison laminar burning speed of model vs. look-up table laminar burning speed at 0%EGR, c) pure methane mixture with 20% EGR, d) comparison laminar burning speed of model vs. look-up table laminar burning speed at 20%EGR</i>	36
Figure 4.1.3, <i>Laminar burning speed vs. equivalence ratio at 800 K: a) methane-hydrogen mixture with 0% EGR, b) comparison laminar burning speed of model vs. look-up table laminar burning speed at 0%EGR, c) methane-hydrogen mixture with 20% EGR, d) comparison laminar burning speed of model vs. look-up table laminar burning speed at 20%EGR</i>	38
Figure 4.2.1, <i>Laminar burning speed vs. pressure at 800 K: a) pure methane mixture with 0%,10%, 20% EGR; b) methane-hydrogen mixture with 0%,10%,20%EGR</i>	39
Figure 4.3.1, <i>Laminar burning speed vs. Temperature at $\phi=1$: a) pure methane mixture at 20 bar, 40 bar, 80 bar; b) methane-hydrogen mixture at 20 bar, 40 bar, 80 bar</i>	40
Figure 4.4.1, <i>Laminar burning speed vs. EGR at $\phi=1,800K$: a) pure methane mixture at 9.09 bar, 40 bar, 80 bar; b) methane-hydrogen mixture at 9.09 bar, 40 bar, 80 bar</i>	41
Figure 4.5.1, <i>Laminar burning speed vs. equivalence ratio at 800K, 0% EGR: a) for a pure methane mixture; b) for a mixture with 85% CH₄ and 15% H₂; c) for a mixture with 75% CH₄ and 25% H₂.</i>	43
Figure 4.5.2, <i>Laminar burning speed vs. equivalence ratio at 800K, 20% EGR: a) for a pure methane mixture; b) for a mixture with 85% CH₄ and 15% H₂; c) for a mixture with 75% CH₄ and 25% H₂.</i>	45
Figure 4.7.1, <i>Laminar burning speed vs. pressure at 800 K: a) pure methane mixture with 0%,10%, 20% EGR; b) ; b) for a mixture with 85% CH₄ and 15% H₂ with 0%,10%,20%EGR ; c) for a mixture with 75% CH₄ and 25% H₂ with 0%,10%,20%EGR</i>	48
Figure 4.8.1, <i>Laminar burning speed vs. EGR at $\phi=1, 800K$: a) pure methane mixture at 9.09 bar, 40 bar , 80 bar; b) mixture with 85% CH₄ and 15% H₂ at 9.09 bar, 40 bar, 80 bar; c)) mixture with 75% CH₄ and 25% H₂ at 9.09 bar, 40 bar, 80 bar</i>	50
Figure 4.9.1., <i>Laminar burning speed vs. Crank at 2000 rpm x 600 kPa for pure methane mixture comparing SL predicted through machine learning model and SL obtained from Huang et al. correlation</i>	51
Figure 4.9.2., <i>Laminar burning speed vs. Crank at 3000 rpm x 600 kPa for pure methane mixture comparing SL predicted through machine learning model and SL obtained from Huang et al. correlation</i>	52
Figure 4.9.3., <i>Laminar burning speed vs. Crank at 2000 rpm x 600 kPa for HCNG15 comparing SL predicted through machine learning model and SL obtained from Huang et al. correlation</i>	52

Figure 4.9.4., <i>Laminar burning speed vs. Crank at 3000 rpm x 600 kPa for HCNG15 comparing SL predicted through machine learning model and SL obtained from Huang et al. correlation</i>	53
Figure 4.9.5., <i>Laminar burning speed vs. Crank at 2000 rpm x 600 kPa for HCNG25 comparing SL predicted through machine learning model and SL obtained from Huang et al. correlation</i>	53
Figure 4.9.6., <i>Laminar burning speed vs. Crank at 2000 rpm x 600 kPa for HCNG25 comparing SL predicted through machine learning model and SL obtained from Huang et al. correlation</i>	54
Figure 4.9.7., <i>Laminar burning speed increment w.r.t. hydrogen rising concentration</i>	54

Abstract

This thesis project exploits the learning and computation capacity of Machine Learning algorithms, with the purpose of forecasting the Laminar Burning Speed (S_L) for gaseous fuelled Spark-Ignition Engines. In the specific, both compressed natural gas (CNG) and compressed natural gas with hydrogen (HCNG) will be evaluated. The analysis has the objective to provide a robust model able to predict, in an acceptable computation time, the S_L . The robustness is fundamental for the CFD computations, also as the saving time that we can gain during the simulating with respect to the requested one for the different flame speed correlation laws. For the creation of predictive model, a dataset is adopted considering the S_L dependence on Temperature (T, [K]), pressure (p, [bar]), EGR rate [%], equivalent air-fuel ratio (ϕ , [-]), available percentage of methane (% CH_4) and hydrogen (% H_2). However for an effective result, Rational Quadratic Gaussian Process Regression model is used due to its better performance than the other considered linear regression, regression tree, and support vector machine (SVM) models. In addition, the neural networks as well are considered for the forecasting of S_L because these are used for solving the regression problems with a high level of efficiency and accuracy. In the specific for this computation step, the Levenberg-Marquardt training algorithm is considered.

1. Introduction

The requirements of existing emission standards, which defines the maximum acceptable limit for the new vehicles exhaust emissions, influence the developing of new technologies. For example, to stay within the imposed restrictions, the elaboration of alternative fuels to petroleum fuels in internal combustion engines has been necessary. The natural gas, mainly made up of methane, is part of this family of fuels because of the low carbon percentage that means a lower level of emissions in terms of CO , CO_2 and HC . For what concerns the NO_x , they will also decrease for the least reached temperature in the cylinder. However, in the case of turbocharging engines the temperature, inside the cylinder, will surely increase with a consequence rising of NO_x emissions. Hence, an EGR system will be exploited to preventing NO_x formation because the residual gasses can act as diluents in the unburned gas mixture, and the peak temperature reached during the combustion process decreases with increasing of the residual concentration. By the way, the recirculated exhaust gasses amount must be kept under control since it produces a less reactive mixture with possible misfiring events. Indeed, at the end of the combustion process, the peak pressure and flame front normally slows down due to the dropping in peak pressure and temperature of mixture caused by expansion phase; but if the temperature goes down too fast due to a high residual concentration, the flame front can stop leaving a part of mixture unburned. Moreover, the natural gas has an anti-knock property that allows the exploiting of a spark ignition engine (SI engine) fuelled by methane at higher compression ratio. Therefore, a rising of thermal efficiency is expected. It should be noted that the methane burns slowly influencing in negative way the efficiency, the available power and the engine fuel consumption due to a certain cycle to cycle variations. Mixing hydrogen with the natural gas can help to address these issues thanks to its reactivity which ensures to improve the thermal efficiency, increase the burning velocity, extend the flammability limits and reduce the pollutant emissions because of a higher Hydrogen/Carbon atomic ratio. Thus, for the design of spark ignition engines an accurate computation for the laminar burning speeds is mandatory. It is a key parameter to describe the fuel properties, the occurring combustion phenomena and strictly related to the equivalence ratio as well as to temperature and pressure of the unburned gas mixture.

1.1 Premixed turbulent combustion

For the spark ignition engines, a low reactivity fuel¹, as gasoline or CNG, is compressed after mixing with air. Despite, in the end of compression, it may reach high temperature and pressure in order of 700 K and 25-30 bar, there is the necessity to trig the oxidation reaction by mean a spark properly due to the already mentioned low fuel reactivity. It can increase at several thousands of temperature Kelvin degree (6000-7000 K) the region where the electrical arc is taking place. Immediately after the first portion of charge is ignited, the heat is transferred to the surrounded layers of mixture meanly through convection increasing their temperature. In this way, the flame travels through the combustion chamber.

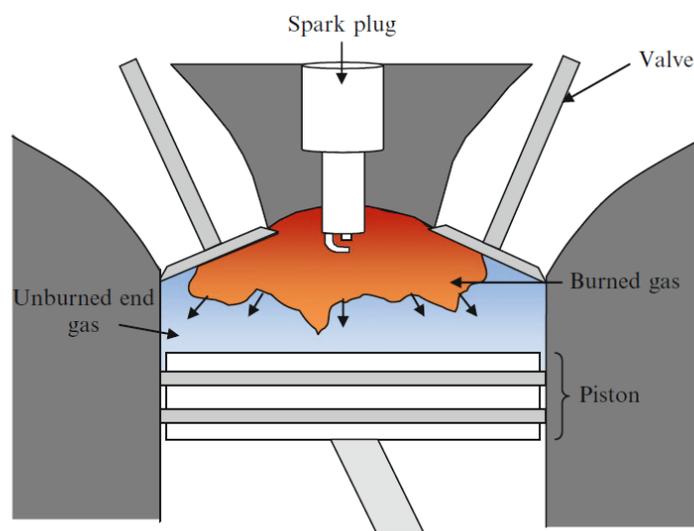


Figure 1.1.1, Representation of SI engine combustion event

Therefore, at the beginning of the combustion, the spark plug is encircled by unburned gasses; right after, the combustion chamber also contains burned gasses and at the interface between them and the unburned ones, the chemical reactions go on determining the growing of the burning region. These two zones are separated by a region called flame front. Actually, the flow process in the engine cylinder is turbulent and the charge motion within the cylinder is one of the major factors that controls the combustion process in spark ignition engines. The turbulence corrugates the flame front, increasing the heat exchange area and the flame propagation speed.

¹ characterized by a short hydrocarbon chain, rigid structure and high capability to evaporate

Indeed, as the engine spin speed goes up, also the turbulence intensity increases thereby positively affecting the burning speed and reducing the time available for combustion at higher engine regime.

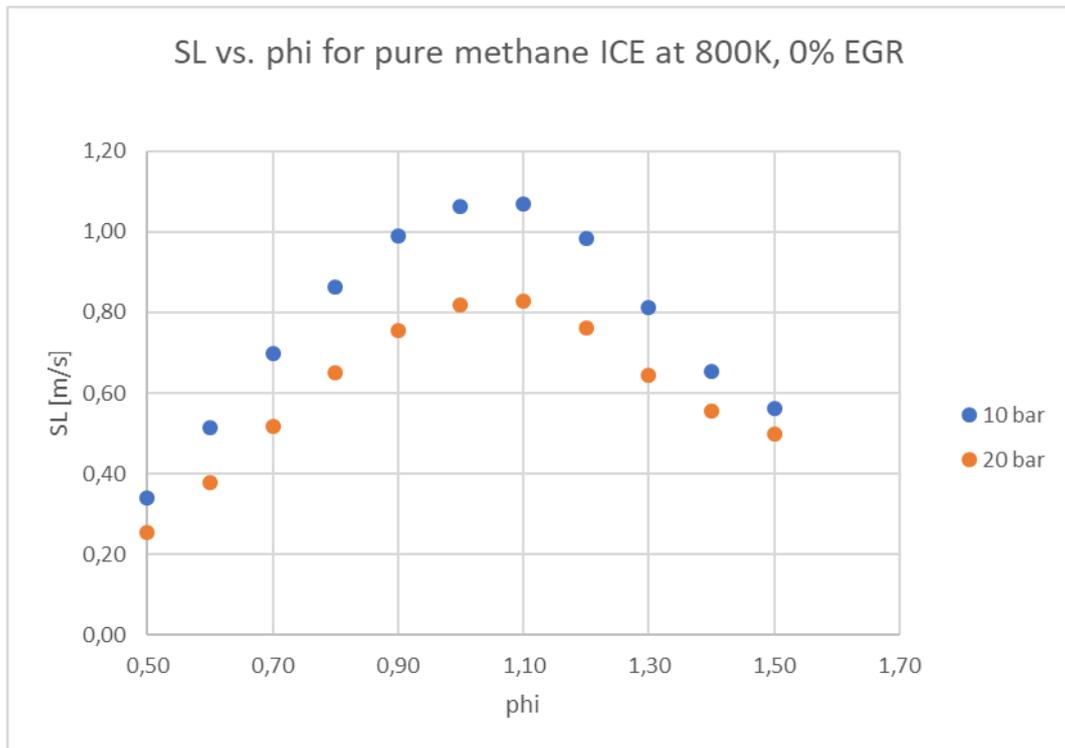


Figure 1.1.2, SL vs equivalence ratio plot for an SI ICE fuelled by CNG

The Figure 1.2 is showing the reason for which Naturally Aspirated Spark Ignition engines must work as much as possible close to the stoichiometric condition; the velocity of combustion has a maximum around an equivalence ratio about 1.1, and so they are throttled at part load to meet the decreasing of injected fuel.

1.2 Laminar burning speed existing models

Over the years, different chemical kinetic models and laminar flame speed correlations are developed for its prediction and to approximately describe the mixture behaviour in the engine combustion chamber. For making some evaluations and understanding how the combustion event

develops, it is possible to assume that the system is working under laminar flow conditions. In this way, there is the opportunity to neglect the effect of corrosion due to turbulence, which complicates the analysis. It is possible to imagine the combustion chamber subdivided in the simplest case in two macrozones, one where there are the unburned gases and the other in which the burned gases are host. The two zones are separated by a flame front assumed to be infinitesimal. By magnifying the flame front, it is possible to distinguish three layers:

- *Preheat zone*, where the fuel available, in the unburned region, reacts with oxygen creating the O, OH and H radicals without any important energy release. This is the initiating step promoted by the plug triggering activity. From the Figure 1.3, it is possible to see how the temperature gradient “T” grows with a concave upward trend thanks to the heat transferred from inner layer;
- *Inner layer*, where the chain-braking reaction takes place; the products of the reaction (radicals) contribute to the reactants (fuel and radical) of another reaction. This continuous transformation of products in reactants leads to a thermal explosion with an important release of energy
- *Oxidation layer* packed by hot products at the equilibrium temperature of the burned gases. In this layer, the temperature gradient has a concave downward trend due to the delivery of heat

Thus, all these reactions are exothermal and the maximum for the Heat Release Rate is reached in the inner layer where the thermal explosion determines the combustion event. Different methodologies can be used for the evaluation of flame speed, like that in which it propagates normally and relatively to the unburned gases. In this case, an unsteady spherical flame front spreading in the radial direction is considered, for which, in general, can be described by the following formula:

$$\frac{d\bar{r}_f}{dt} = \bar{V}_u + \bar{S}_{L,u} \quad (1)$$

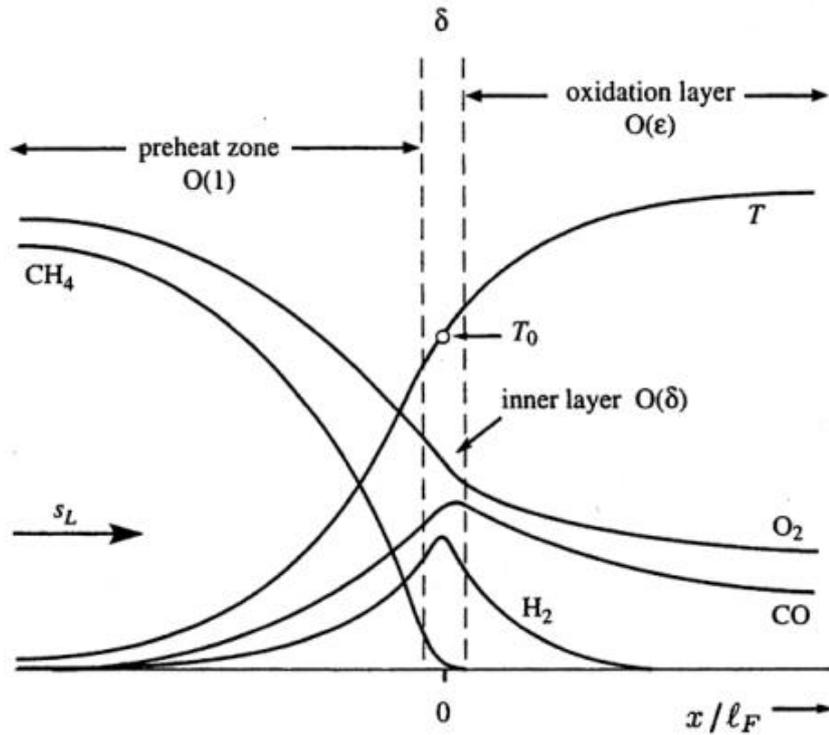


Figure 1.2.1, Magnification of laminar flame front

By considering a one-dimensional flow in radial dimension, the vectoral formula (1) can be simplified into the scalar one:

$$\frac{dr_f}{dt} = V_u + S_{L,u} \quad (2)$$

where the first term is the flame front velocity, the second one is the unburned gas velocity and the third one is the laminar burning speed measured at the unburned gas. Introducing the conservation of mass between the burned and unburned zone, it is possible to compute V_u :

$$V_u = \frac{\rho_u - \rho_b}{\rho_u} \frac{dr_f}{dt} \quad (3)$$

$\frac{\rho_u - \rho_b}{\rho_u}$ is showing the expansion effect² that is present when the unburned mass, within the flame front, becomes burned. As consequence, the laminar burning speed measured at the unburned gas is obtained as³:

$$S_{L,u} = \left(1 - \frac{\rho_u - \rho_b}{\rho_u}\right) \frac{dr_f}{dt} \quad (4)$$

For the simulation of SI engine cycles, different models are evaluated and classified as zero-dimensional model, quasi-dimensional model, multi-dimensional model and multi-zone model. The possibility and capability to predict a simplified description of three-dimensional phenomena occurring in the cylinder is the main difference between the quasi-dimensional and the zero-dimensional model, which is only based on thermodynamic equilibrium. The zero-dimensional model thermodynamic equations, representing the model structure, are:

- the mass conservation equation dependent on time
- the energy conservation equation dependent on time

Multi-zone models include some geometrical parameters in addition to the previous basic thermodynamic approach, like the interface radius (flame) which divides burned from unburned gases. In this case, the multi-zone can be also called as two-zone model. Therefore, combustion is triggered, several multiple burned gas zones are generated at each specific crank angles and the flame propagation starts. For example, the Figure 1.2.2 shows a multizone model for a SI engine, at a specific crank angle θ , with one unburned zone and six burned zones, where the point 1 is where the spark is taking place. Every burned and unburned gas zone has an uniform temperature and composition, but with an independent evolution during the expansion stroke. Also, the pressure is shared through the cylinder in uniform way. For the first thermodynamic law, an ideal gas inside the combustion chamber is considered. During the compression stroke, the mixture is homogeneous and made by induced air, fuel and residual gases⁴ coming from the previous engine cycle:

² $\rho_u - \rho_b > 0$

³ $V_b = 0$ for assumption, because the burned mass inside the flame front is not moving

⁴ They are computed by means of a correlation in which the residual gases are in function engine speed and average pressure during the exhaust and intake pressure strokes

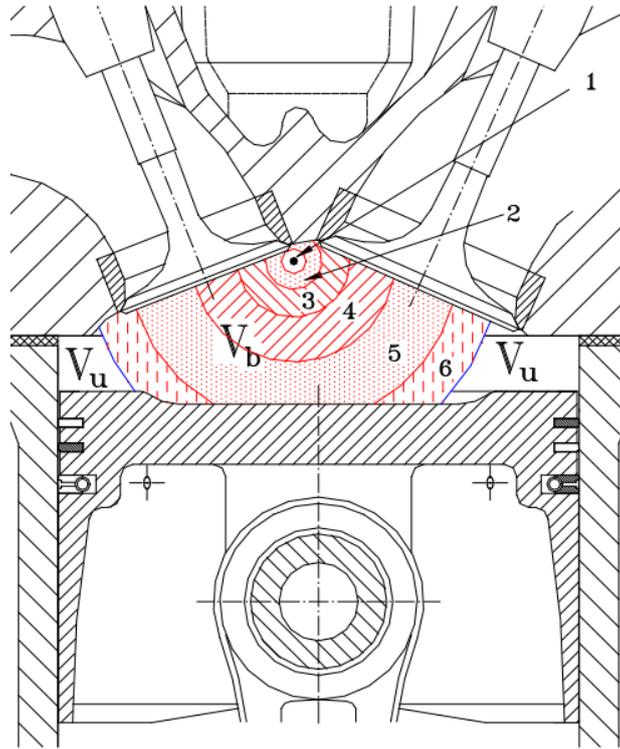


Figure 1.2.2 Representation of SI engine combustion event according to the Multi-zone model; V_b is showing the burned gas volume, instead V_u is the unburned gas volume. The point 1 is where the combustion starts thanks to the triggering event

$$m = m_a + m_f + m_r \quad (5)$$

For the conservation of mass, the flow through the crevices and all possible other leakages are neglect:

$$dm = d(m_a + m_f + m_r) = dm_u + dm_{b,n} = 0 \quad (6)$$

It is worth to highlight the presence of $dm_{b,n}$ in the formula: after the spark timing, the first burned single zone is generated and it grows in volume and mass since part of unburned gas is burning. At a specific crank angle, a new burned gas zone is created and the old one does not receive mass from the unburned gasses, unlike its volume continue to change due to its density variation. As consequence:

- dm_b from the first to the penultimate zone is equal to zero
- dV_b for each zone is always different from zero

For what concerns the energy conservation equation, it is necessary to write different equations considering one for unburned gas zone, n-equations for each i-th burned gas zones and one for the last one burned zone:

$$\text{For unburned gas zone:} \quad -q_u A_u \frac{d\theta}{\omega} + V_u dp = (1 - x_b) m dh_u \quad (7)$$

$$\text{For } i\text{-th burned gas zone:} \quad -q_{b,i} A_{b,i} \frac{d\theta}{\omega} + V_{b,i} dp = x_{b,i} m dh_{b,i} \quad (8)$$

$$\text{For the last burned gas zone:} \quad -q_{b,n} A_{b,n} \frac{d\theta}{\omega} + V_{b,n} dp = x_{b,n} m dh_{b,n} + m dx_{b,n} (h_{b,n} - h_u) \quad (9)$$

where in general q is the heat flux transferred to the wall, A is the surface across which the heat flux is exchanged, $\frac{d\theta}{\omega}$ is the time, x is the mass fraction⁵ and h the enthalpy. In all the equations, it is possible to distinguish three terms:

- the first one is the heat transferred
- $V dp$ is the work transferred⁶
- the change of enthalpy related to the specific zone.

In the equation (9), there is an additional fourth term that represents the presence of some unburned gasses entrained in the burned gas zone. These will change their chemical composition. However, if a prediction activity must be carried out, a sub-model for x_b is necessary. Moreover, the model owns a formulation of Newton's convection law that takes account of the fact that there is an unsteadiness of the exchanged heat flow from the bulk gas to the wall [12]. The heat transfer coefficient \tilde{h} is evaluated by means of the Woschni's correlation, for which it is in function of engine bore, the mean piston speed, the gas temperature, the pressure in the combustion chamber, the displacement, the temperature at the intake valve closing and the motored pressure. The multi-dimensional model is governed by the Navier-Stokes equations in addition to

⁵ $x_{b,i} = \frac{m_{b,i}}{m}$; $x_b = \sum_{i=1}^n x_{b,i}$; $m_u = (1 - x_b)m$

⁶ It is not $p dV$ because on the right of equation there is the enthalpy term

the conservation equation of mass and energy; these are also dependent on the spatial coordinate. However, in practical simulations the analytical correlations of the laminar flame speed as function of equivalence ratio, pressure and temperature are preferred. They are simply implemented in CFD codes, considering that the S_L is defined for any type of fuel, once the unburned mixture composition, pressure and temperature are known. Various forms of empirical and semi-empirical functional relationships have been proposed for the laminar burning speed [1,2]. The most used is the power law formula:

$$S_L(\phi, T_u, p_u) = S_{L0} \left(\frac{T_u}{T_0} \right)^\alpha \left(\frac{p_u}{p_0} \right)^\beta \quad (10)$$

where S_{L0} is the velocity for a specific equivalence ratio ϕ , measured at room condition (T_0, p_0) and α and β are respectively temperature and pressure exponents that could be constant or mixture strength-dependent. Gülder pitched an expression of S_{L0} :

$$S_{L0}(\phi) = Z W \phi^\eta e^{-\xi(\phi-\delta)^2} \quad (11)$$

where W , η and ξ are constants for a specific fuel and $Z=1$ for a single constituent fuel. In the case of natural gas, the mixture is a mix of hydrocarbon molecules and their volume fraction vary with geographical source and treatment applied during production or transportation [3]. Despite varying Z in the Gülder correlation, it has been demonstrated that it is not enough to evaluate the effects of variation of natural gas composition on its laminar flame speed [3]. A few years ago, Dirrenberger et al. [4] starting from a modified version of Equation (6), obtained a relationship valid for a natural gas composed by methane, ethane and propane. S_L estimated with this correlation has a good accuracy for lean and rich mixtures, but less acceptable close to stoichiometric condition. In the present thesis, the modified Huang et al. correlation is used, and it stipulates a dependence of laminar burning speed on fuel-air equivalence ratio, pressure, temperature and amount of residual gas in the mixture:

$$S_L(\phi, T_u, p, x_{res}) = S_{L0} \left(\frac{T_u}{T_0} \right)^\alpha \left(\frac{p_u}{p_0} \right)^\beta (1 - 1.5x_{res}) \quad (12)$$

The chemical kinetic analysis is useful to evaluate the chemical reaction rates; the mechanism of a chemical reaction can be thought as a sequence of events occurring as before reaching the final product. The chemical kinetic study contains different complex reactions that must be analysed through a specific mechanism. Indeed, the reaction mechanisms are designed as step-by-step descriptions of what is happening at molecular level during the reaction. Therefore, an overall reaction can be imagined to be composed of several elementary processes. For example for the oxidation reaction of methane molecule, it does not collide with two molecules of oxygen forming suddenly CO_2 and two molecules of H_2O_{vap} , but there will be a first step in which is involved a hydrogen of CH_4 which will be replaced with an oxygen molecule and so on. Then, several steps will follow each other up to the CO final oxidation step in CO_2 . Naturally, this procedure requires more time if the molecule is always more rigid and longer. Different mathematical models are exploited to describe and forecast a specific chemical reaction. A particular attention is given to the GRI 3.0 mechanism developed to obtain a detailed chemical reaction mechanism capable to describe all the important step of the natural gas flame propagation, ignition, NO formation and re-burn chemistry [7]. This mechanism can be adapted also for air-methane-hydrogen combustions. The dataset available, on which the laminar burning speed predictions will be done, is designed as a look-up-table generated by the result of the commercial simulation tool DARS used for 0D and 1D chemistry analysis, laminar flame speed computation and so on. Naturally, in this case DARS helped for the calculation of laminar burning speed for air-methane-hydrogen mixture. For this computation, a freely propagating flame model is used, and it assumes that the propagating flame front is flat and infinitely large. This gives the possibility to consider a constant system total energy and avoid evaluating the boundary conditions. As regards kinetic mechanism, the GRI 3.0 mechanism is used to validate the methane-hydrogen combustion and compute the corresponding mixture laminar burning speed over a wide range of pressure, temperature, equivalence ratio and residual fraction. The mechanism is represented by three files which contain information about chemistry and thermodynamic parameters.

1.3 The hydrogen contribution to CNG engines

The hydrocarbon fuels are a molecular combination of hydrogen and carbon atoms and the simplest one is the methane (CH_4); as it was mentioned in the previous subchapter, the methane is also the main constituent of natural gas in addition to ethane (C_2H_6), butane (C_4H_{10}) and propane (C_3H_8). Everyone is part of the light hydrocarbon fuel molecule since they have a number of carbon atoms less than five, meaning a low molecular weight. On the contrary, the gasoline⁷ and the diesel⁸ are heavy hydrocarbons, which are liquid at ambient conditions unlike the hydrogen that is gaseous. Moreover, it has also the lowest atomic weight with respect to any other substance, and consequently also a low-density value at whatever temperature and pressure conditions. The hydrogen at ambient conditions is in gaseous form, so it is possible to store it in high pressure tanks or in liquid state at 20 K. The expansion ratio considers the difference in volume that is present when the gas or liquid is stored on a storage system compared to one occupied at the atmospheric pressure and temperature by the same gas or liquid. When the hydrogen is kept in gaseous state at high pressure the expansion ratio is 1: 240, instead when it is stored in liquid form it rises to 1: 848. This means that from the hydrogen size storage system point of view, it is better to choose the first solution. However, this solution can take out a problem related to the safety passenger conditions: because in case of accident or explosion, the tank, being highly pressurized, would act as bullets. Thus, another possible solution is to produce the hydrogen on board, for example from methanol, avoiding this last issue. Another important parameter to take into account is the flammability of hydrogen, defined as upper and lower flammability limit. These two bound values represent respectively the lowest and the highest gas concentration that will support the propagation of flame when the mixture is ignited: below the lower flammability limit (LFL) there is not enough fuel to withstand the combustion, unlike above the upper flammability limit (UFL) the mixture inside the cylinder is too rich and there is not enough oxygen to trig the combustion. The hydrogen flammability has a wide range of concentrations in air varying from 4 to 75 %, and it tends to increase with the temperature. Thus, the leakages of hydrogen can lead to potential burning event and are not so simple to avoid because it is the smallest molecule naturally occurring. By the way, the hydrogen has an

⁷ An important constituent of gasoline is the heptane (C_7H_{16})

⁸ Fuel used for CI engines are extremely "flexible" since they are long chain like cetane ($C_{16}H_{34}$)

autoignition temperature higher than methane and gasoline, so it is difficult to ignite the mixture air/hydrogen without an external source and has also a decent anti knock property when it is exploited in ICE (see below the summary Table 1.1).

Fuel	Autoignition Temperature [°C]	Octane Number ⁹
Gasoline	230-480°C	87
Methane	540°C	125
Hydrogen	585°C	>130

Table 1.3.0.1

The hydrogen has also a higher flame and combustion propagation velocity than methane; so, it can solve the low methane propagating flame problem that impacts negatively the ICE performance. This sentence can be confirmed by Figure 1.4.1, where the comparison of laminar burning speed for an air-methane and air-hydrogen-methane combustion is shown. In addition, there is also the possibility to reduce the *CO*, *CO₂* and *HC emissions* due to lower carbon/hydrogen ratio; on the other hand, producing a higher combustion temperature, an EGR system can be introduced to take under control the *NO_x* emissions. To sum up, the laminar burning speed is in function of:

$$S_L (\phi, T_w, p_w, \% EGR, \% CH_4, \% H_2) \quad (13)$$

It is important to introduce also the dependence on the amount of recirculated exhaust gases in the combustion chamber because they make the mixture less reactive.

⁹ Octane Number describes the antiknock tendency

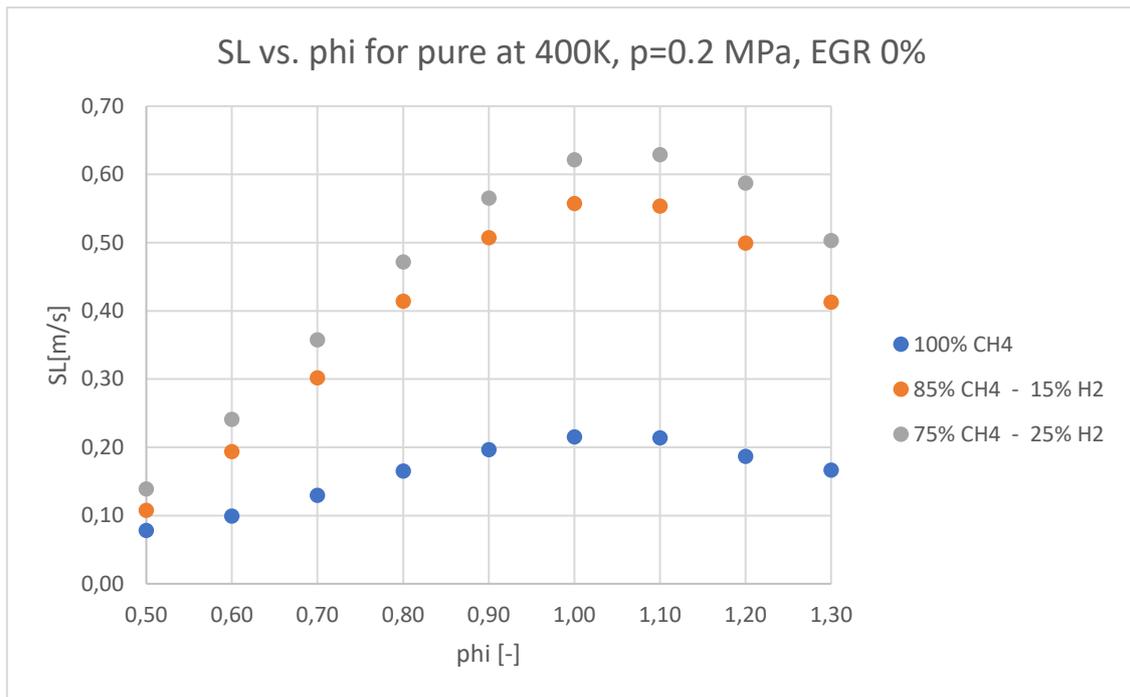


Figure 1.3.1 Laminar burning speed vs equivalence ratio at 400 K, 0.2 MPa for three different cases

Continuing with the theme of emissions, it is worth recalling that for the direct injection spark ignition engines (DISI) there is also a particulate number (PN) limit introduced from the 1st September 2017, in fact also a particulate filter must be adopted. In general, the main sources of particulate emissions in Gasoline Direct Injection (GDI) engines is the wall wetting, since the fuel spray hits the piston surface, the intake valve bottoms and the liners. This fuel impingement in a CNG DISI engine can be isolated due to the gas fuel directly injected into the cylinder. Moreover, from an experimental analysis the CNG DISI engine particle emissions showed an unannounced trend: PN concentration rises as fuel injection timing is advanced due to the fuel mixing quality; instead, the PN concentration dependence on spark timing remains unchanged with respect to GDI engine [5]. Hora et al. [6] performed some experiments on CNG and hydrogen-CNG (HCNG) fuelled engines; partial combustion of lubricating oil leads to formation of nano-particles in the engine combustion chamber. They showed that HCNG combustion emitted a lower amount of particulate compared to CNG because the flame speed of HCNG is higher than CNG one, resulting in better combustion.

2. Fundamentals of Machine Learning and Artificial Neural Networks

The aim of this second chapter is to provide an overview about the key concepts of machine learning algorithms that will be necessary to reach the final thesis goal.

2.1 The different applications of Machine Learning

Machine Learning exploits different statistical methods to give the machine (e.g. a computer) the possibility to learn directly from the data without any explicit knowledge. In general, the "training" is the first step in which the available data are examined for the following model creation. It is obtained through a specific algorithm, which analyses the supplied dataset. After the algorithm has found the best possible repetitive model, it is used to make estimates or to achieve any other pre-set objective [i].

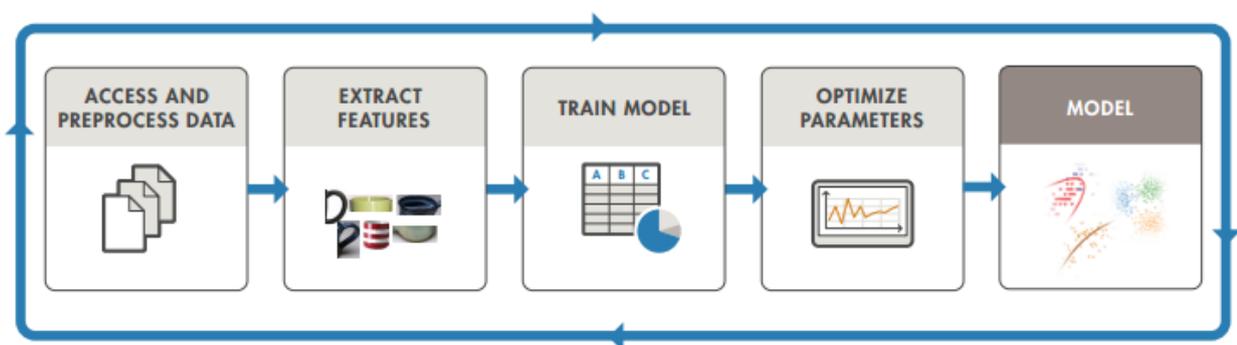


Figure 2.1.1 Iterative process to achieve satisfactory performance

Machine Learning finds many fields of application and can be used for: speech recognition, image processing, e-mail filtering, regression or classification problems and so on. The main activities of Machine Learning are subdivided in three categories:

- *Supervised Learning*, in which the machine receives data, in the form of input and output, with the final goal of extracting a general rule that associates them. In this case, the

response of the machine to a given problem is based on the "experience" accumulated in the training phase. There are two different types of supervised learning:

- *Classification*, the inputs are divided into two or more classes and the learning system must produce a model able to assign them a specific class (example: e-mail anti-spam filtering, where the inputs are the emails and the classes are "spam" and "not spam");
- *Regression*, starting from continuous input, machine uses a model capable of predicting a continuous output (example: predicting laminar burning speed S_L having as input T , p , EGR and ϕ);
- *Unsupervised Learning*, in which the computer aims to find a structure or a path in the input data provided. The machine will have to list all the information, organize it and learn its meaning and using to get the best results for different situations. Usually, for this type of learning, the *Clustering* method is considered [ii];
- *Reinforcement Learning*, where the machine is equipped with systems and tools able to improve its learning and understand what is happening around her thanks to the support provided by a series of elements such as sensors. This type of learning is used for autonomous cars that, in fact, make use of many sensors, cameras, GPS systems to create a detailed learning system that allows to take any decision [ii].

2.2 Artificial neural networks (ANN)

Artificial neural networks are machine learning approaches, seen as mathematical models composed of artificial neurons used to solve problems in various technological fields. The network is usually composed of a layer of input neurons (or nodes), which receive signals from the external environment or from one or more intermediate layers (hidden layers) organized in several levels. Each node can receive, process and send a signal to a following node [8]. The hidden layer is the intermediate layer where there are several hidden neurons that are neither in the input layer nor the output layer [9]. There are different artificial neural networks: Feed forward type, Recurrent neural network and Radial Basis Function neural network. In this project, a two-layer feed forward network with sigmoid hidden neurons and linear output neurons is exploited because an input-

output fitting problem with a big amount of data must be solved. The final layer is the output one where there is a single neuron capable of providing the above-mentioned output. The neural network can understand and process data of a certain degree of complexity and providing a function that would be difficult to extract through manual procedures.

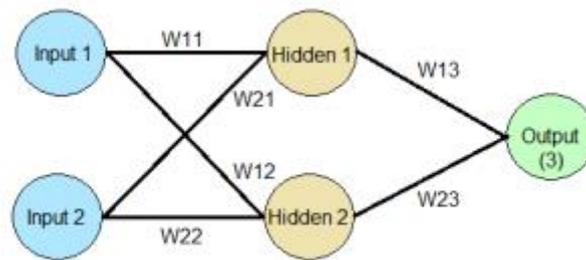


Figure 2.2.1, Representation of a simple neural network

Networks can also be used to solve Classification, Regression and Clustering problems. Moreover, these can work in parallel and so handle a lot of data. Despite their great efficiency and the ability to generate almost certainly acceptable results, due to the impossibility of examining step-by-step the path from input to output, neural networks are often criticized. In fact, we should accept the result as it is, without being able to explain why and how the result has been generated¹⁰.

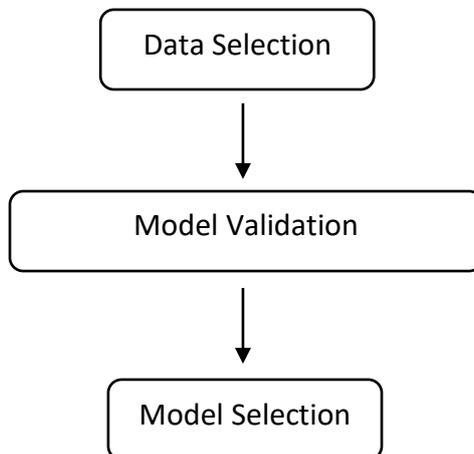
¹⁰ Indeed, the neural networks are considered as black boxes

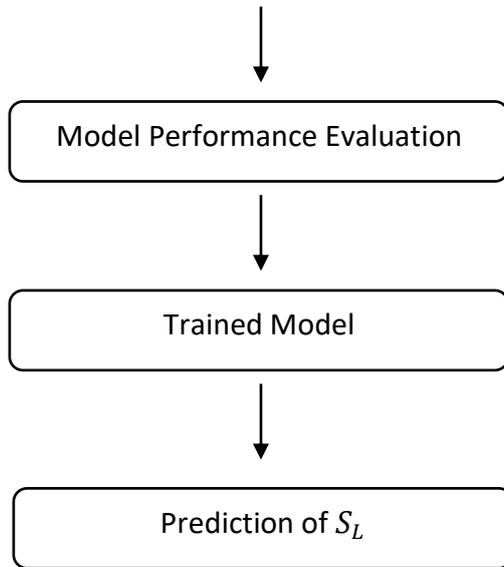
3. Laminar burning speed modelling

In this chapter, a model for methane-air and methane-hydrogen-air mixture based on regression machine learning algorithms will be created and implemented to predict the laminar burning speed S_L . For a first attempt design, it is considered a simpler model able to forecast the aforementioned parameter as function of temperature, pressure, equivalence ratio and amount of recirculated gases. This is useful to understand the prediction potentiality of different machine learning algorithms and how to select the best regression model according to the indicated performance parameters. The same first attempt design approach is exploited to get acquainted with artificial neural network model. After that, a more complicated model is developed considering a higher amount of input data and introducing other two independent variables, namely the percentage of methane and hydrogen in the mixture. Thus, at the end, the model is able to forecast a laminar burning speed as function of six independent variables as shown in the equation (7). The input data, exploited for the training of model, are extrapolated by means of the GRI 3.0 mechanism on DARS. The Regression Learner Matlab App and the Neural Network Matlab Fitting App are used for this project to reach the goal. Using these applications, it is possible to explore the data, select a specific model validation scheme, create it and evaluate the results.

3.1 Modelling through Machine Learning

3.1.1 Flow-chart approach





Data selection

For both models developed, the data come from a look up table, where the S_L values are taken out by using GRI 3.0 mechanism. These values of laminar burning speed are useful to have a comparison with the results coming from the regression models. By the way the data was organized in this way:

T_u [K]	p_u [bar]	EGR rate [%]	phi [-]	X_{CH_4} [%]	X_{H_2} [%]
300÷1600	1÷200	0÷100	0÷5	100%	0%
300÷1600	1÷200	0÷100	0÷5	85%	15%
300÷1600	1÷200	0÷100	0÷5	75%	25%

Table 3.0.1.1, Dataset used for the model training

For the first attempt design, the first data row of previous table is used to create the model. Instead for the second model, the approach is slightly different. First of all, an unique dataset¹¹ is built by combing all the three data row. However, not all the data will be exploited for the training of model; in fact, a random sample is created and it will be exploited for the training and

¹¹ The number of data is 153307

validation of algorithm. As consequence, the trained model will predict the laminar burning speed for the remaining data that are not involved in the previous sample. This is useful also to manage and control the problem of overfitting¹²; in fact, this issue is present when a model is able to work (or as in this case to make prediction) only with a specific dataset, and it may fail to adapt different data for the forecasting of future observations. Thus, overfitting occurs when a model remembers the training data instead of learning a trend that will be useful to generalize the analysis. To reduce the overfitting chance, there is the possibility to select different techniques like the cross-validation or the Holdout validation-based early stopped. The latter splits the data in two subsets, training and validation data, of which the corresponding classification errors¹³ are computed. When the validation error starts to rise while the training one is still decreasing, overfitting has occurred [10]. This is an approach quite exploited by neural network, but the cross-validation is more powerful and complex than the early stopped approach. The cross-validation splits data in k-partitions (called k-folds) instead of one single partition as seen before, but it is not suitable for abundant dataset.

Model validation

For the model validation, the “Holdout validation” is exploited because for the first attempt and the final model the data available abound. For this approach the dataset is subdivided in three portions:

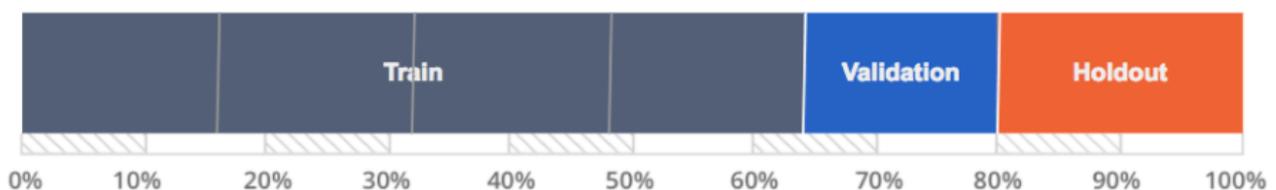


Figure 3.1.1, Partitioning of data according to the Holdout validation approach

¹² In machine learning, it can be called also overtraining

¹³ The workedclassification error is computed comparing the predictions and the targets (true values)

- *Training set* is a subset of the provided dataset from which the machine learning algorithm learns the relationship among inputs and the target;
- *Validation set* is another input data subsection to which the trained algorithm is applied for understanding how much it is reliable in defining the relationship inputs-predicted target. However, the amount of data dedicated to the validation is fixed at priori and if there are a lot of data it is possible to keep the 40% of them for the validation, so that the model can be selected basing on 40% of the available data;
- *Holdout set* gives a final estimate of model after the two previous steps. It is like a final check for our model.

Model Selection and Performance evaluation

The Regression Learner Matlab App gives the possibility to train a selection of models to find the best solution in terms of accuracy and operating time. The choice can be among:

- *Linear regression model*
- *Regression tree model*
- *Gaussian process regression (GPR)*
- *Supported vector machine (SVM)*

To assess the best performance of model, it is necessary to consider four statistic parameters:

- Root Mean Square Error (RMSE) measures the average magnitude of error, which is the difference between the estimated target value of model and the actual test value. It is always positive and the smaller the RMSE will be and the smaller will be the difference of predicted model value and the provided one by look-up-table;

$$RMSE = \sqrt{\frac{1}{n} \sum_{j=1}^n (y_j - \hat{y}_j)^2} \quad (14)$$

▼ History		
1.1	☆ Linear Regression Last change: Linear	RMSE: 1.1881 4/4 features
1.2	☆ Linear Regression Last change: Interactions Linear	RMSE: 1.0699 4/4 features
1.3	☆ Linear Regression Last change: Robust Linear	RMSE: 1.4632 4/4 features
1.4	☆ Stepwise Linear Regression Last change: Stepwise Linear	RMSE: 1.0699 4/4 features
1.5	☆ Tree Last change: Fine Tree	RMSE: 0.44222 4/4 features
1.6	☆ Tree Last change: Medium Tree	RMSE: 0.57872 4/4 features
1.7	☆ Tree Last change: Coarse Tree	RMSE: 0.70225 4/4 features
1.8	☆ SVM Last change: Linear SVM	RMSE: 1.353 4/4 features
1.9	☆ SVM Last change: Quadratic SVM	RMSE: 1.0922 4/4 features
1.10	☆ SVM Last change: Cubic SVM	Training <input type="checkbox"/> <input checked="" type="checkbox"/> <input type="checkbox"/> <input checked="" type="checkbox"/> 4/4 features
1.11	☆ SVM Last change: Fine Gaussian SVM	RMSE: 0.79844 4/4 features
1.12	☆ SVM Last change: Coarse Gaussian SVM	Training <input type="checkbox"/> <input checked="" type="checkbox"/> <input type="checkbox"/> <input checked="" type="checkbox"/> 4/4 features

Figure 3.1.2, Table provided by the Matlab App in which all the possible trained regression models are shown

- R-squared (R^2) is the principle goodness index of regression curve and it is also known as coefficient of determination. R^2 is always less than one and higher than zero. It must compare the trained model with the model where the response is constant and equals the mean of the training response;
- Mean squared error (MSE) that is the square of the RMSE, and so, as for the RMSE, the smaller it will be and the better it is;
- Mean absolute error (MAE), that is always positive and provide an information similar to RMSE but this last one, taking the square root of the average errors, gives a relatively high weight to large errors. Thus, RMSE is more useful when large errors are particularly undesirable.

$$MAE = \frac{1}{n} \sum_{j=1}^n |y_j - \hat{y}_j| \quad (15)$$

For holdout validation, the RMSE is calculated on the held-out observations. By the way, the best the best model for your goal is not the one is showing the best score in terms of MAE, RMSE and so on. Indeed, a model with slightly lower score could be better than another model, it all depends on final goal.

Trained Model

For both the two examined cases, the best regression model is the Rational Quadratic Gaussian Process Regression (GPR). It is a Bayesian approach¹⁴, namely a mathematical procedure that applies probabilities to statistical problems, and produces a posterior distribution of an unknown regression function "f". Bayesian statistic is strictly related to the concept of posterior probability $p(\vartheta|D)$, described as:

$$p(\vartheta|D) = \frac{p(\vartheta)p(D|\vartheta)}{p(D)} \quad (16)$$

where D is the outcome of the events, $p(\vartheta)$ is the prior (i.e. the function of probability distribution of a priori belief), $p(D|\vartheta)$ is the likelihood of observing the results given the prior distribution $p(\vartheta)$ and $p(D)$ is the evidence [iii]. The prediction problem is managed as a non-parametric regression problem¹⁵, such that it is possible to face up with the goal of machine learning that is the one to learn from a set of measurements a function "f" of the relationship existing between the predictor variables¹⁶ $x (\emptyset, T_u, p_u, \% EGR, \% CH_4, \% H_2)$ and the target¹⁷ y, defined as:

$$y = f(x) + \varepsilon \quad (17)$$

¹⁴ It is a mathematical procedure that applies probabilities to statistical problems.

¹⁵ It is a specific type of regression analysis, where the predictors are created according to information derived from the dataset.

¹⁶ Input data or covariates

¹⁷ In this case study, it is the laminar burning speed

where ε is the error term. A Gaussian process (GP) defines a probability distribution over functions, described as:

$$f(x) \sim GP(m(x), k(x_i, x_j)) \quad (18)$$

where $m(x)$ and $k(x_i, x_j)$ are the mean and the covariance function of the regression function "f" [11]. The covariance function $k(x_i, x_j)$, called also Kernel function, determines how two outputs y_i and y_j are expected to be at the inputs x_i and x_j . These x_i and x_j are column vectors and so the Kernel function determines how much the target at point x_i is affected by targets at other point x_j with $i \neq j, i = 1, 2, \dots, n$. For different kernel functions, the kernel parameters are strictly related to the signal standard deviation σ_f and the characteristic length scale σ_l . The characteristic length scales define how far away the values of x_i are from the responses to become unrelated [iv]. In this work, the Rational Quadratic Kernel is evaluated and it is defined in this way:

$$k(x_i, x_j, \vartheta) = \sigma_f^2 \left(1 + \frac{r^2}{2\alpha\sigma_l^2} \right) \quad (19)$$

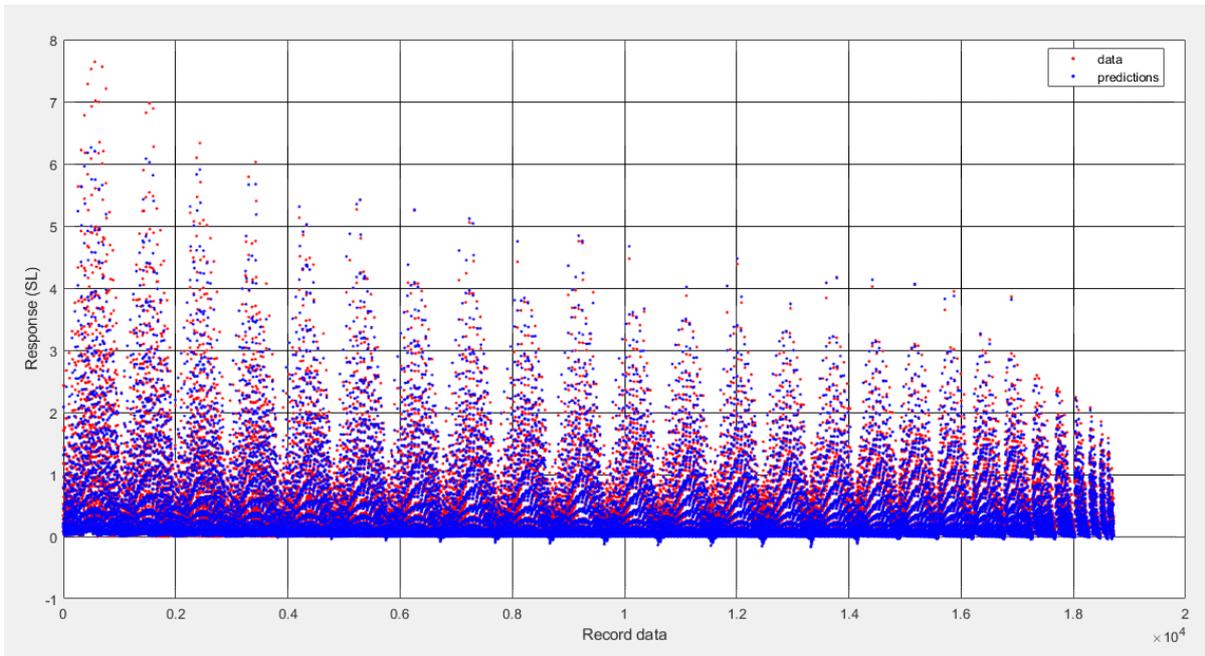


Figure 3.1.3, Comparison between the available and predicted S_L for the first Trained model

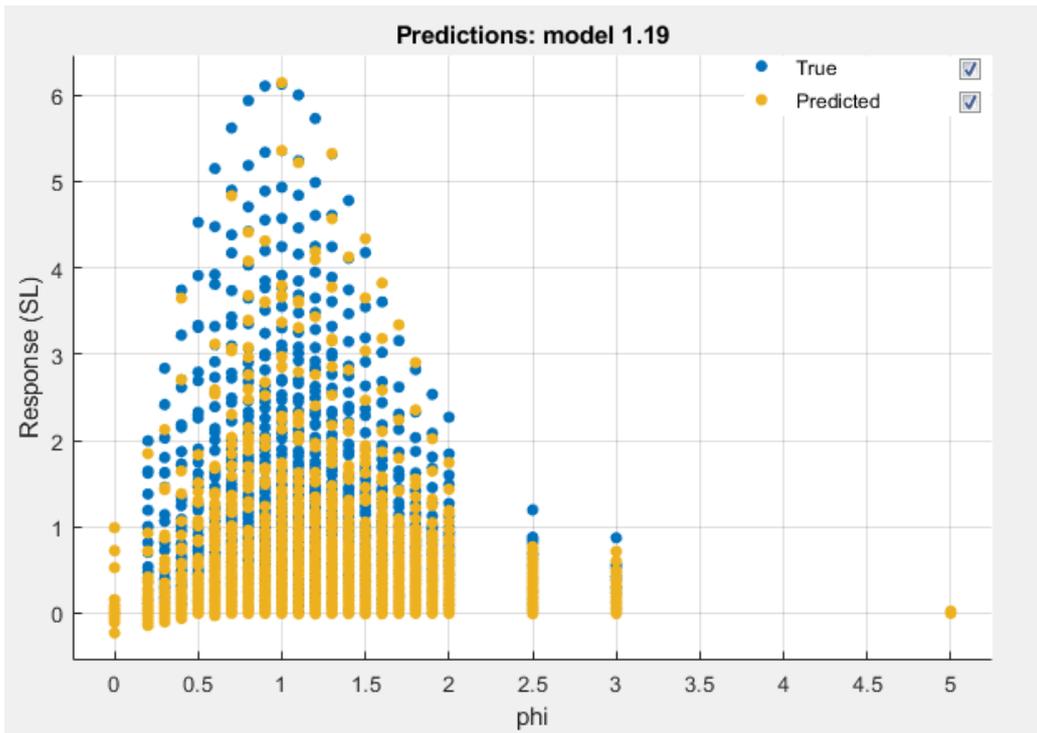


Figure 3.1.4, S_L vs. equivalence ratio ϕ for the first trained model

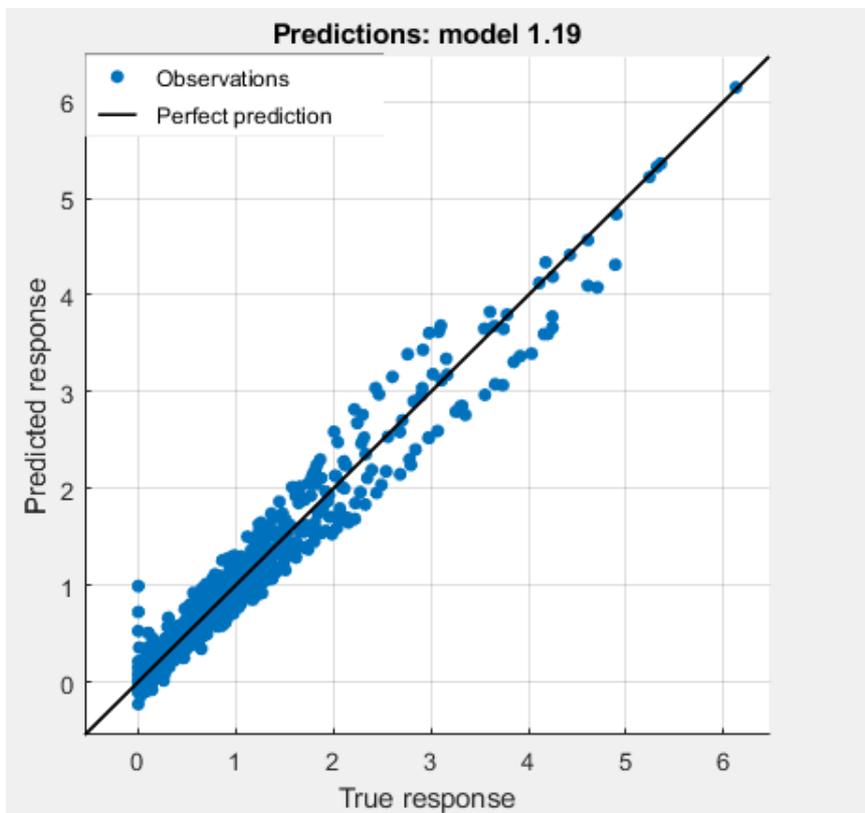


Figure 3.1.5, Regression plot for the first Trained model

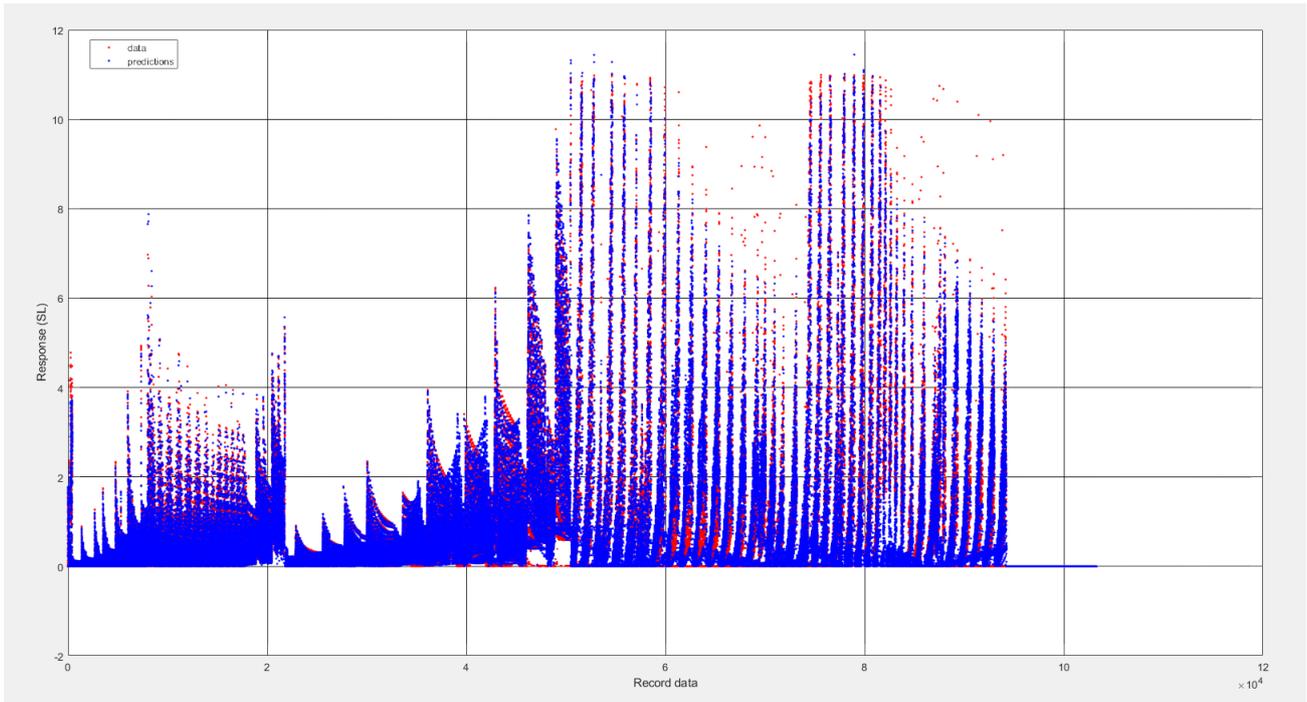


Figure 3.1.6, Comparison between the available and predicted S_L for the second Trained Model

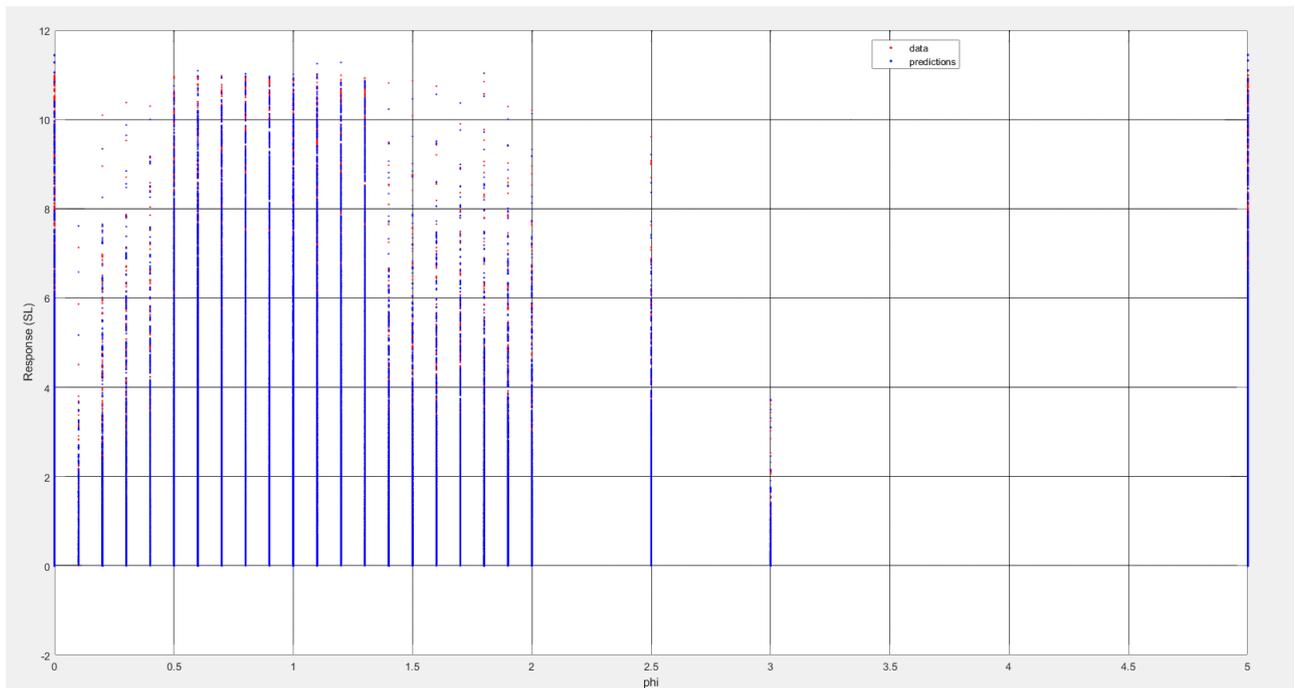


Figure 3.1.7, S_L vs. equivalence ratio ϕ for the second Trained model

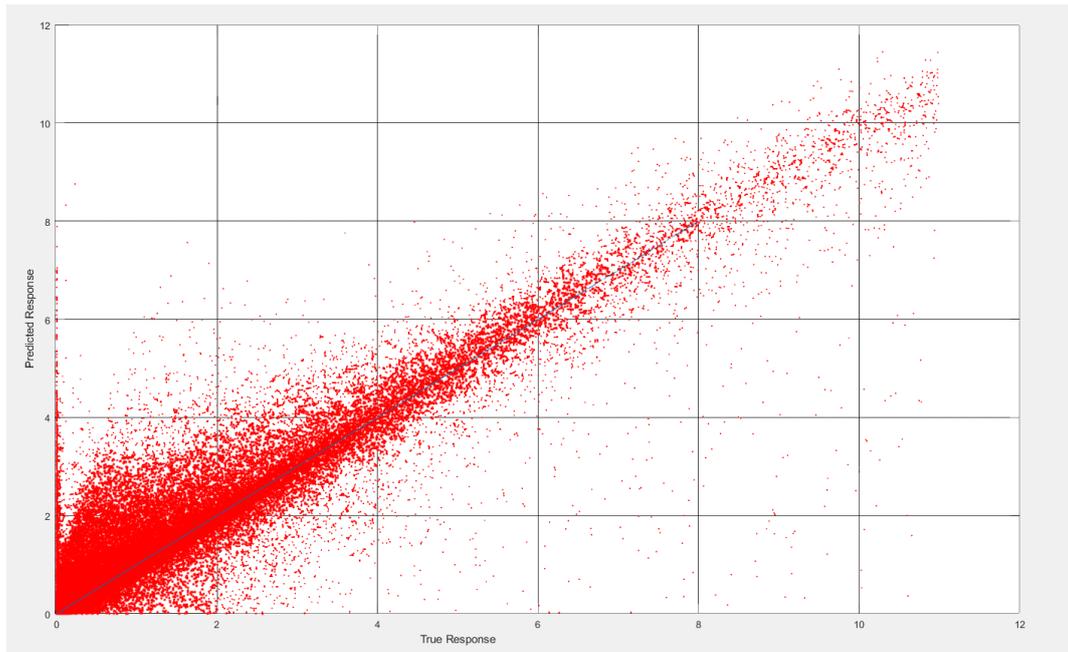


Figure 3.1.8, Regression plot for the second Trained model

Predicted Laminar burning speed

The Regression Learner Matlab app gives the possibility to extract a code, through which there will be the possibility to obtain a regression model and its RMSE like the one seen during the exploiting of the application. Thus in input, there is the necessity to provide a table containing the same predictor and response columns as provided in the app. In output, all the information about the trained model will be available, and naturally there will be the possibility to make predictions with the returned model using a new dataset, that must contain at least the same predictor columns as used during the training. Indeed, as it has been said in the Data Selection section, for the second model not all the data will be exploited for the training activity and the laminar burning speed predictions will be executed on the remaining data that are not involved in training and validation step.

3.2 Modelling through Neural Networks

The curve fitting is a process used to construct a mathematical function able to have the best correspondence on a series of data; so, this curve fitting can mean carrying out both an interpolation among points or a "smoothing" activity where a flat function is built on the data available. This field is strictly related to the regression analysis and to the research of a mathematic curve which can be suitable with the observed data. These curves can give a hand for highlighting a data trend, hypothesizing a specific function pattern where the data are not available and so on. The Neural Network is a tool that can solve a fitting problem in a restricted time interval because its algorithm is faster than the conventional iterative one. For this project, the Neural Fitting Matlab app is used for solving an input-output fitting problem through a two-layer forward neural network with ten hidden neurons. The steps that must be followed are:

- the *data selection*
- the *creation and training of model*
- the *evaluation of its performance* through the parameters seen in the previous Machine Learning section.

However, the datasets exploited are the same that are shown in the Machine learning Data selection section. Thus, there will be two case studies:

1. a first attempt design, needed to understand the capability of the neural network algorithm
2. a second analysis, where two different datasets will be exploited for the training and the forecasting of Laminar Burning Speed S_L

For what concerns *the data selection*, there is always the necessity to subdivide and provide the data in input and output (target) form. To carry out this, a simple Matlab code can be created and executed to split the data in two portions: input and output matrices. After that, there is the necessity to select the amount of data that will be used for:

- the training phase of model
- validation phase, useful for evaluating the network generalization
- testing phase, for measuring the network performance after the training.

For this work, 30% of data are exploited for the training phase, the 35 % for the validation and the 35% for the testing. The splitting is chosen for obtaining the best trade-off between performance and computation time as Figure 3.2.1 and 3.2.2 are showing.

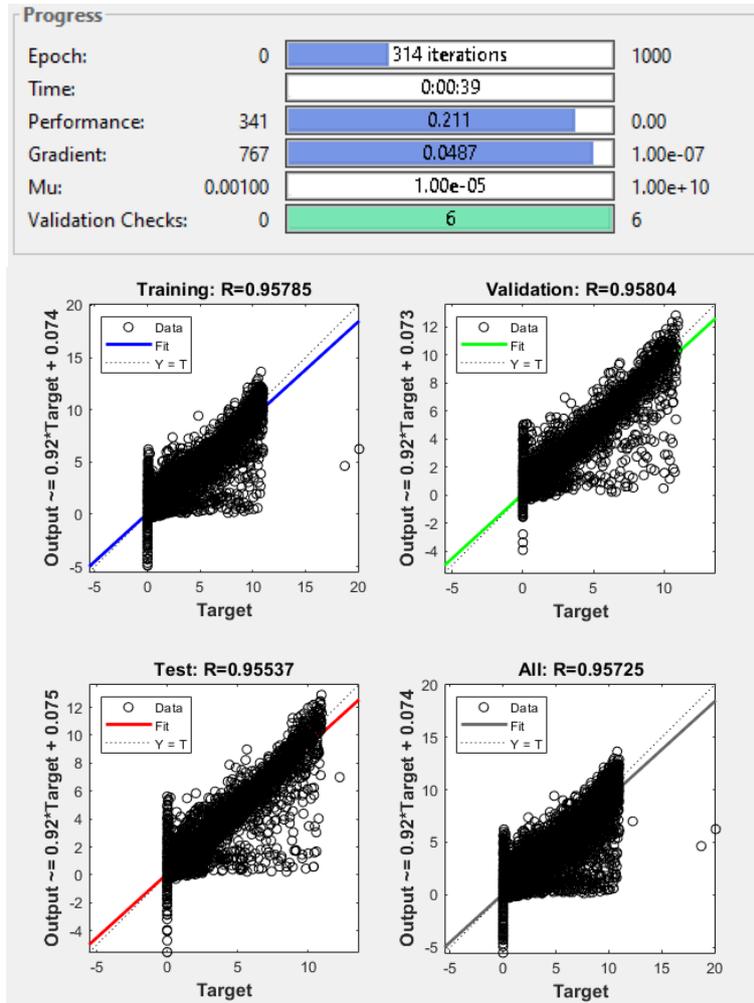


Figure 3.2.1, For the first attempt design: NN performance and computation timing with Testing at 50% , Validation at 25% and Testing at 25%

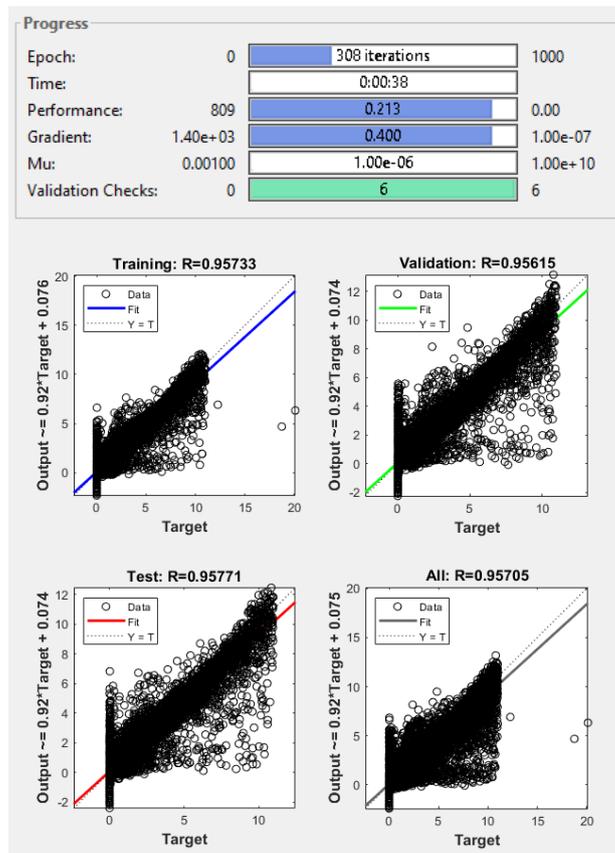
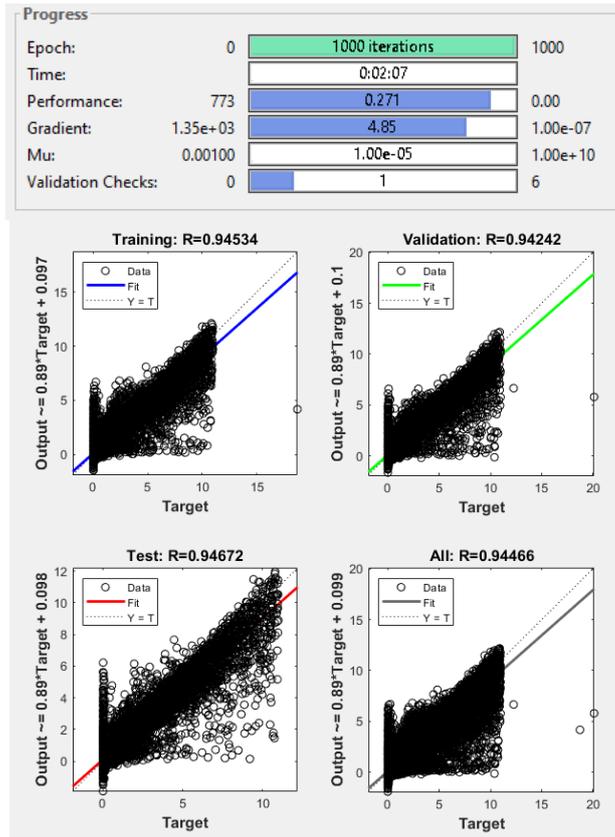


Figure 3.2.2, For the first attempt design: a) NN performance and computation timing with Testing at 40%, Validation at 35% and Testing at 25%; b) NN performance and computation timing with Testing at 30%, Validation at 35% and Testing at 35%.

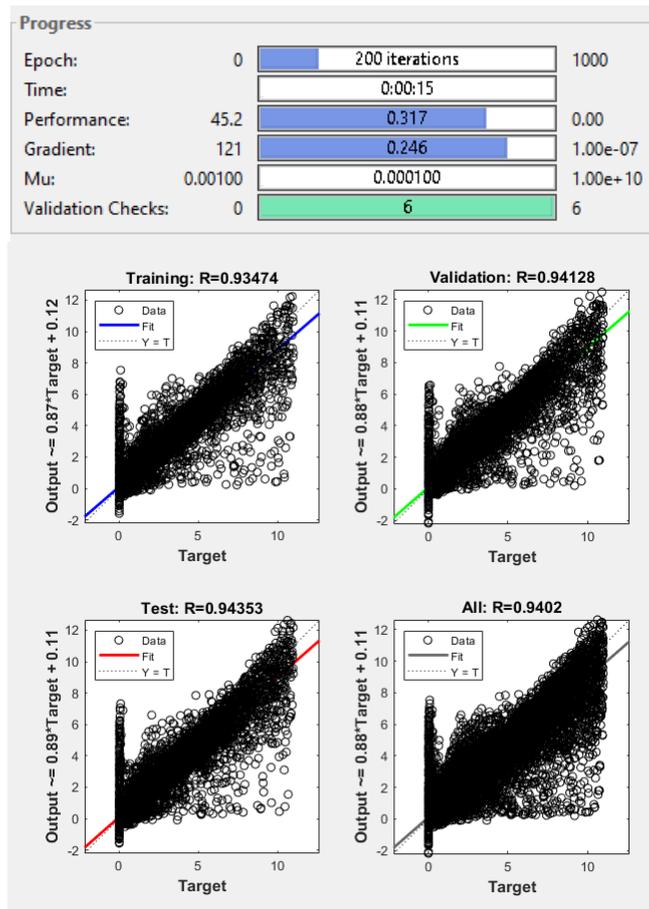


Figure 3.2.3, For the second analysis: NN performance and computation timing with Testing at 30%, Validation at 35% and Testing at 35%

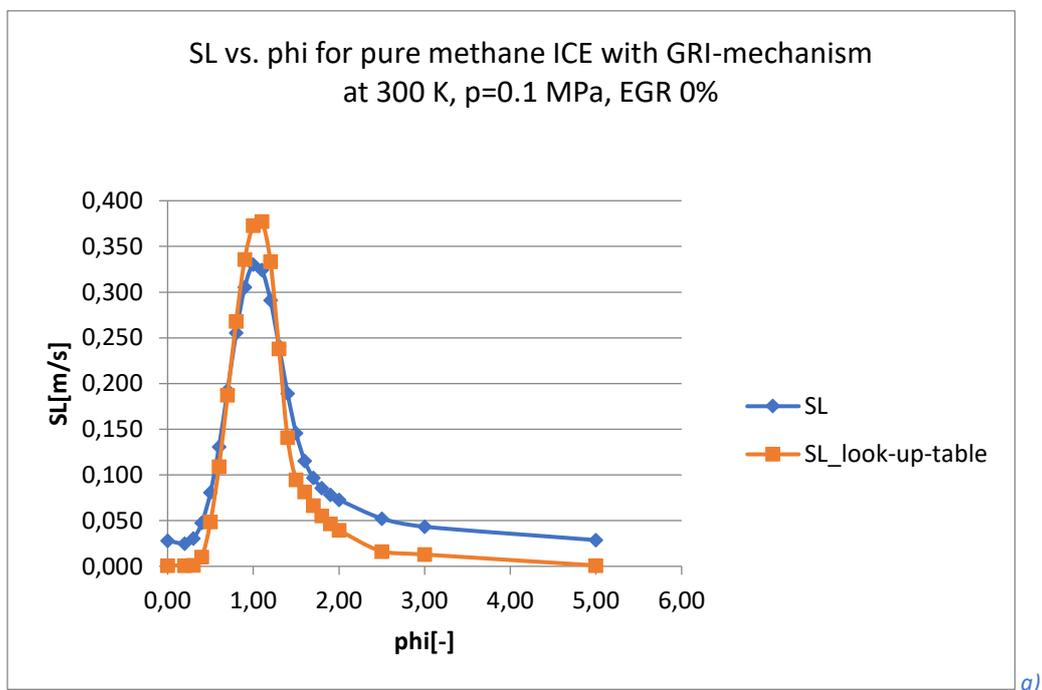
For the training, the Levenberg-Marquardt algorithm, i.e. damped least squares, is used to manage the curve fitting problem through a minimization approach. This regression technique searches the function such that it minimizes the sum of distances squares among the observed data and the one of function itself. The Figure 3.2.1 and 3.2.2 are showing the lines of least squares. The regression plots provide not only an information about the accuracy of model, by means of the value of R, the slope and offset of line, but also it is showing a comparison between each generated data with its known target. Naturally, when the amount of data is increasing, it is difficult to appreciate in the plot the spreading of data and the presence of errors among model output and the targets.

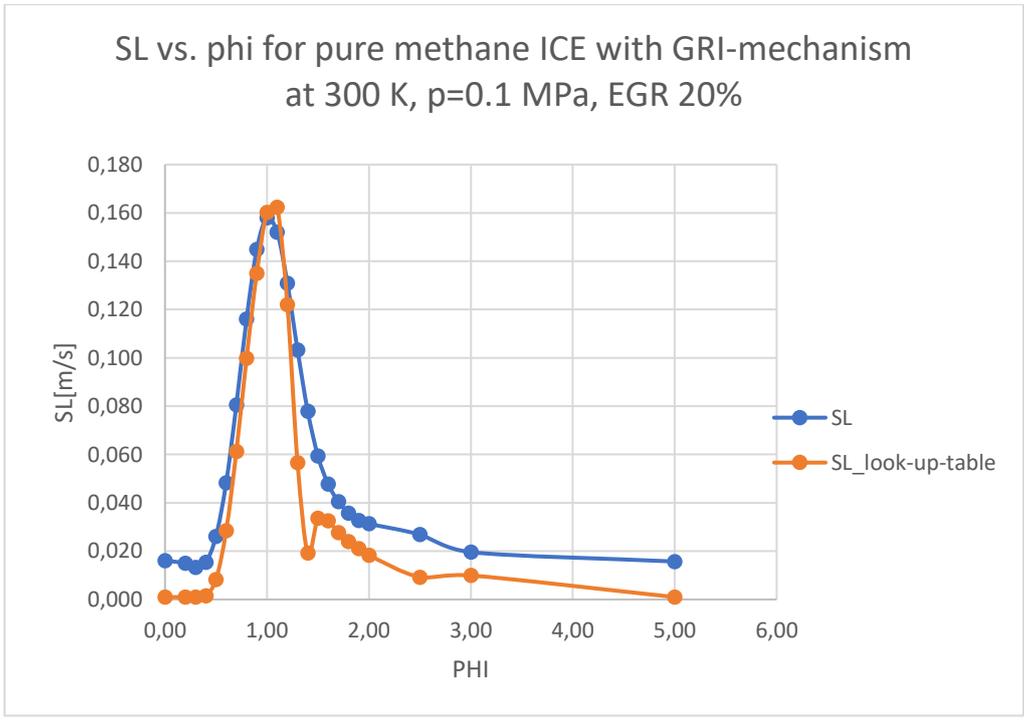
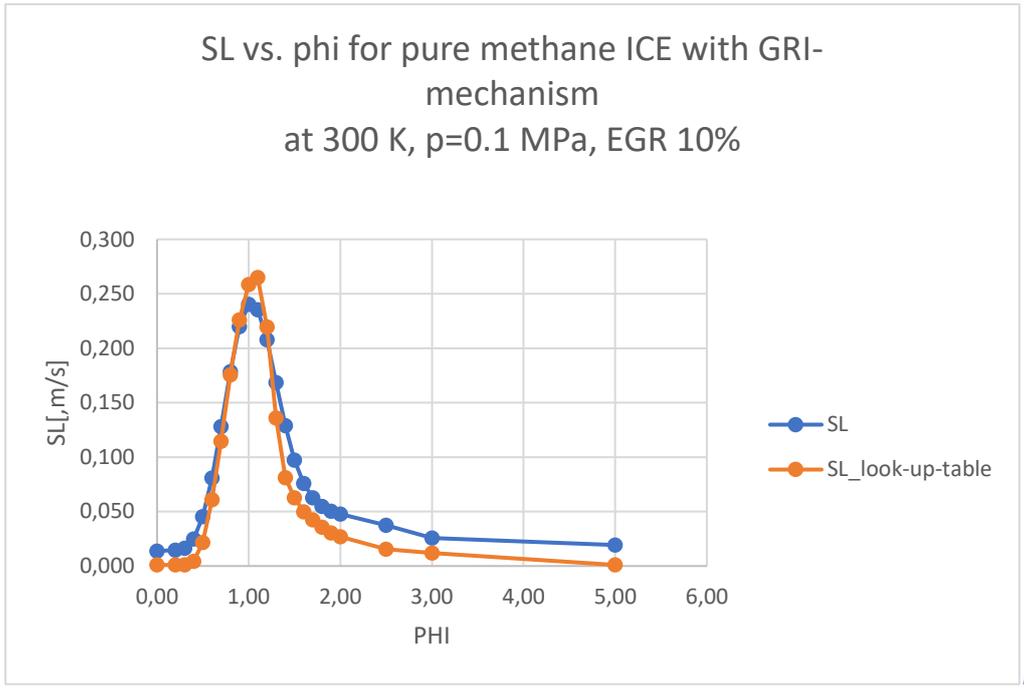
4. Laminar burning speed results

In this chapter, the physic aspect of models will be analyzed to understand if the previous algorithms are not only appreciable from a mathematical and statistical point of view.

4.1 Laminar burning speed vs. Fuel-air Equivalence ratio with Rational Quadratic GPR method

In this section, the analysis of the equivalence ratio effect on the laminar burning speed will be treated. The plots refer to the laminar burning speed predicted with Rational Quadratic GPR method considering the second case model analysis. In the first plots, it is possible to see the difference in trend between the S_L computed by models and the one provided by the look-up table considering the 3.0 GRI mechanism; the level of temperature and pressure is fixed correspondingly at 300 K and 1 bar.





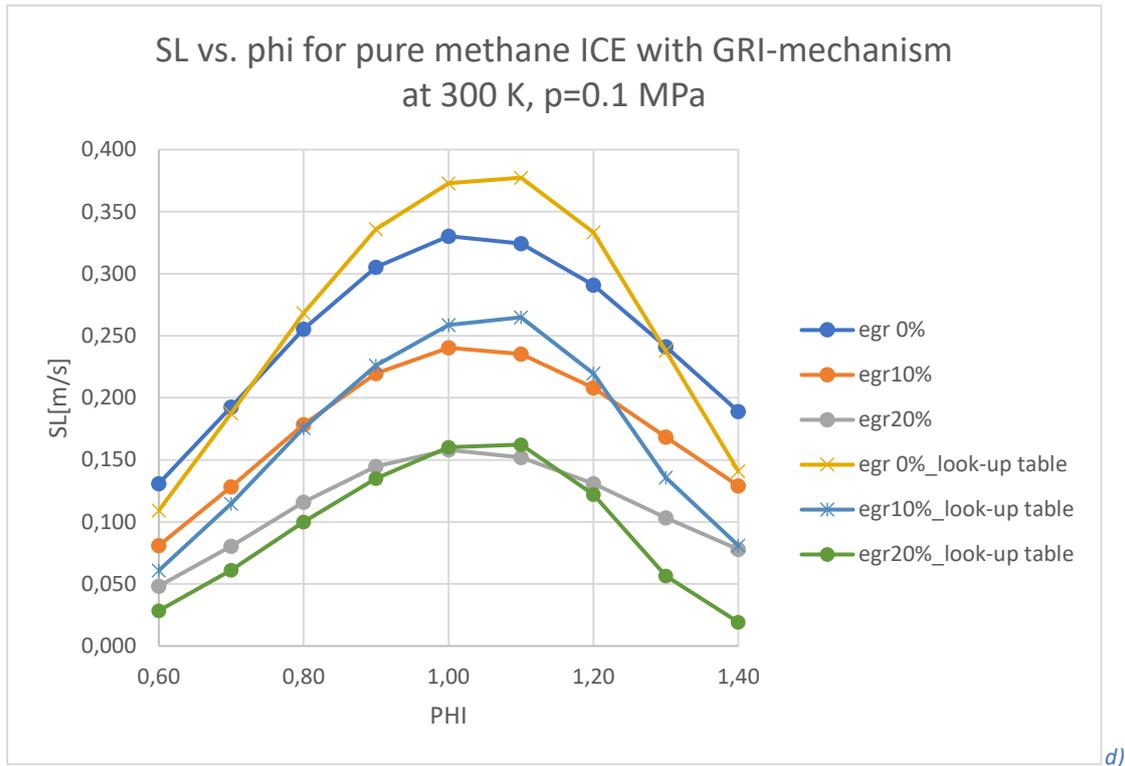
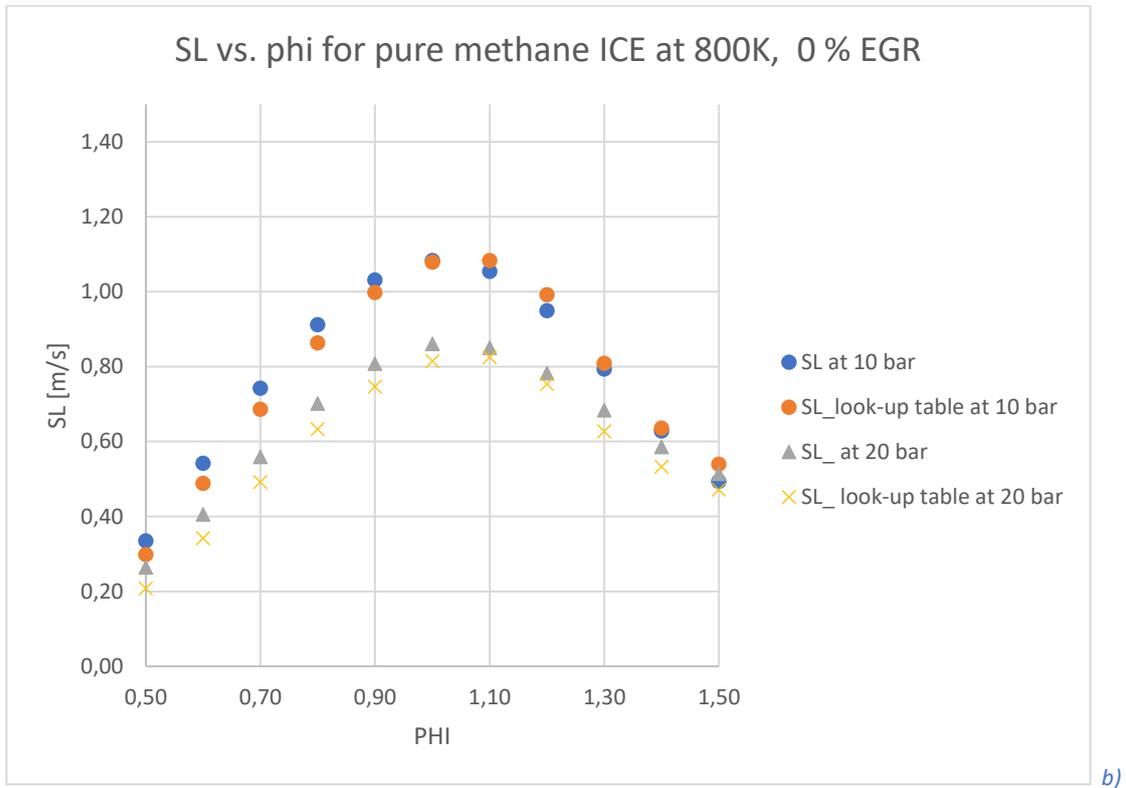
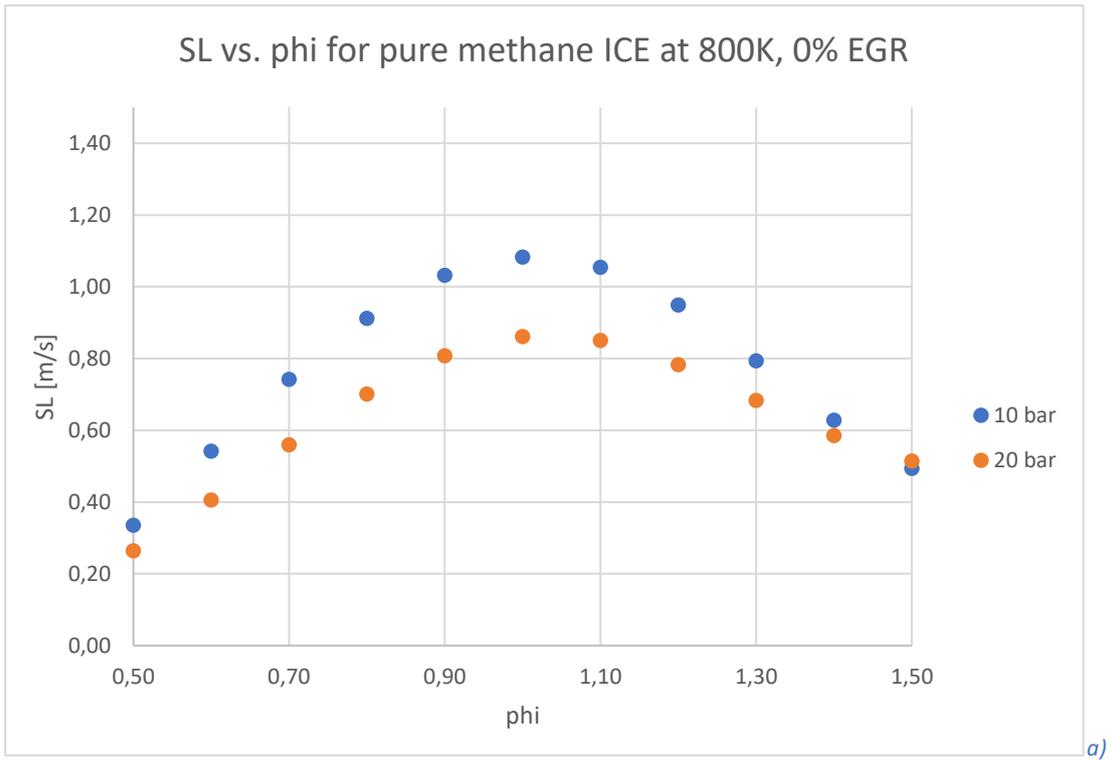


Figure 4.1.1, Laminar burning speed vs. Equivalence ratio: a) pure methane EGR 0%, b) pure methane 10% EGR, c) pure methane 20% EGR, d) comparison among previous cases

It is possible to see, especially in the Figure 4.1.1 c), how the discontinuity points, present in the S_L look-up table, are smoothed by model computation. In the following plots, the fuel-air equivalence ratio ranges between 0.5 and 1.5 and two different mixtures are considered to evaluate how the laminar flame speed trend changes:

- pure methane (100% CH_4)
- methane-hydrogen with 85% CH_4 , 15% H_2

Moreover, in the figures there will be a comparison among curves that are moving in function of pressure level. Depending on the Figure 4.1.2 and 4.1.3, it is possible to highlight how the curves move downward as the pressure rises with a stronger value change for the laminar burning speed at low pressure events.



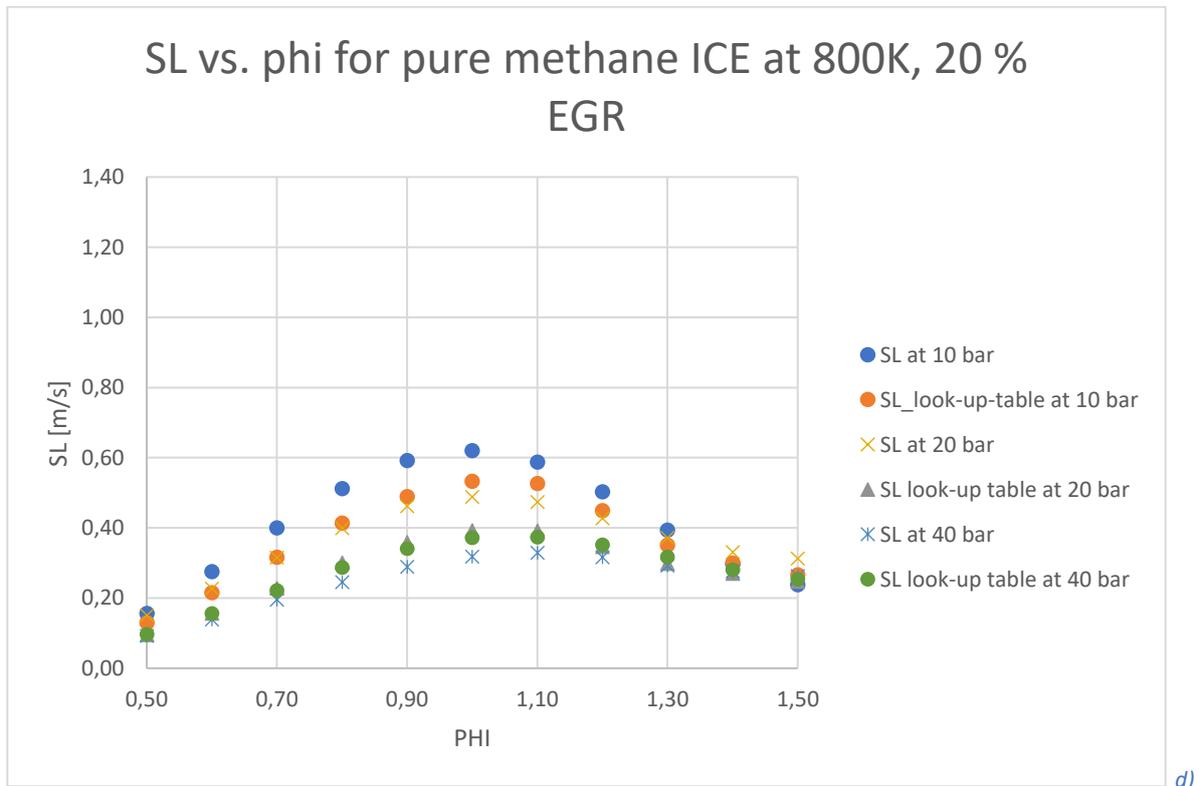
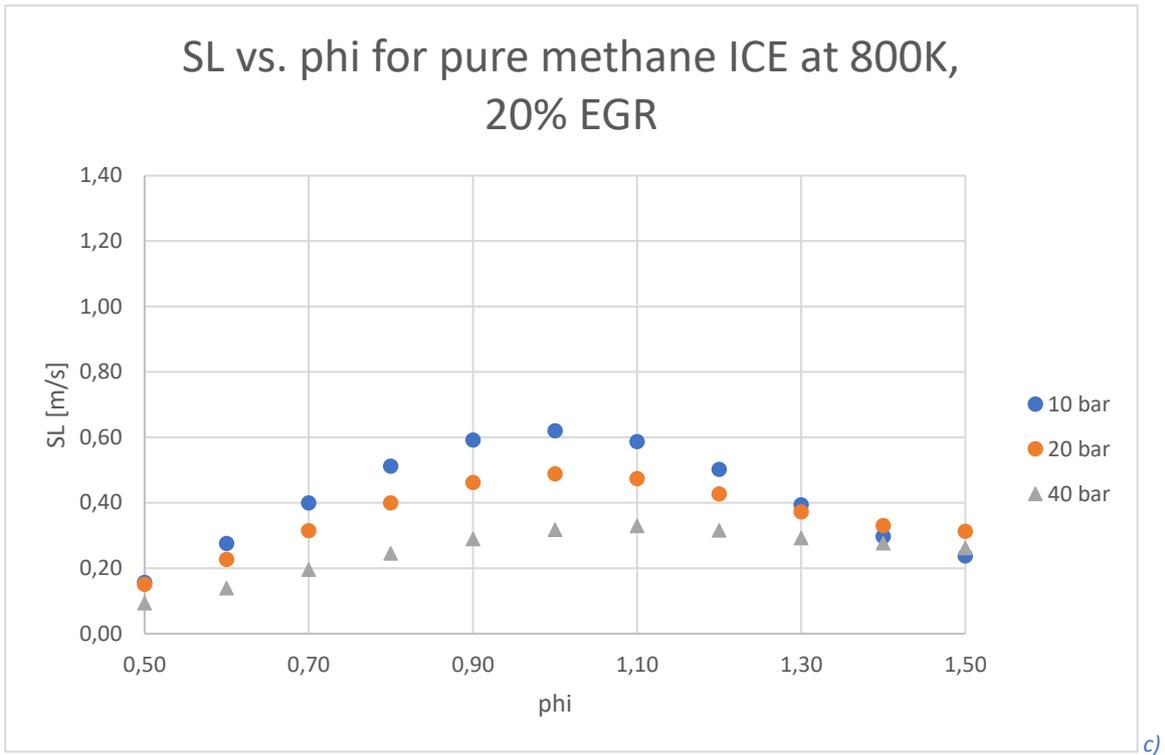
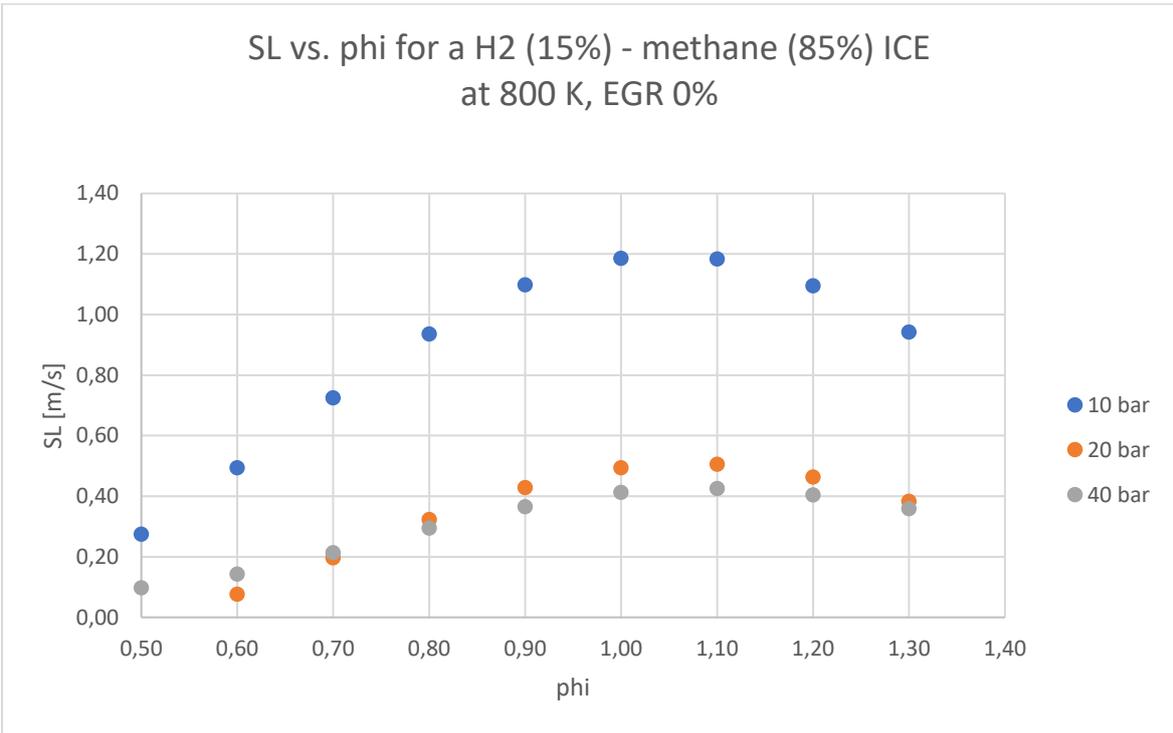
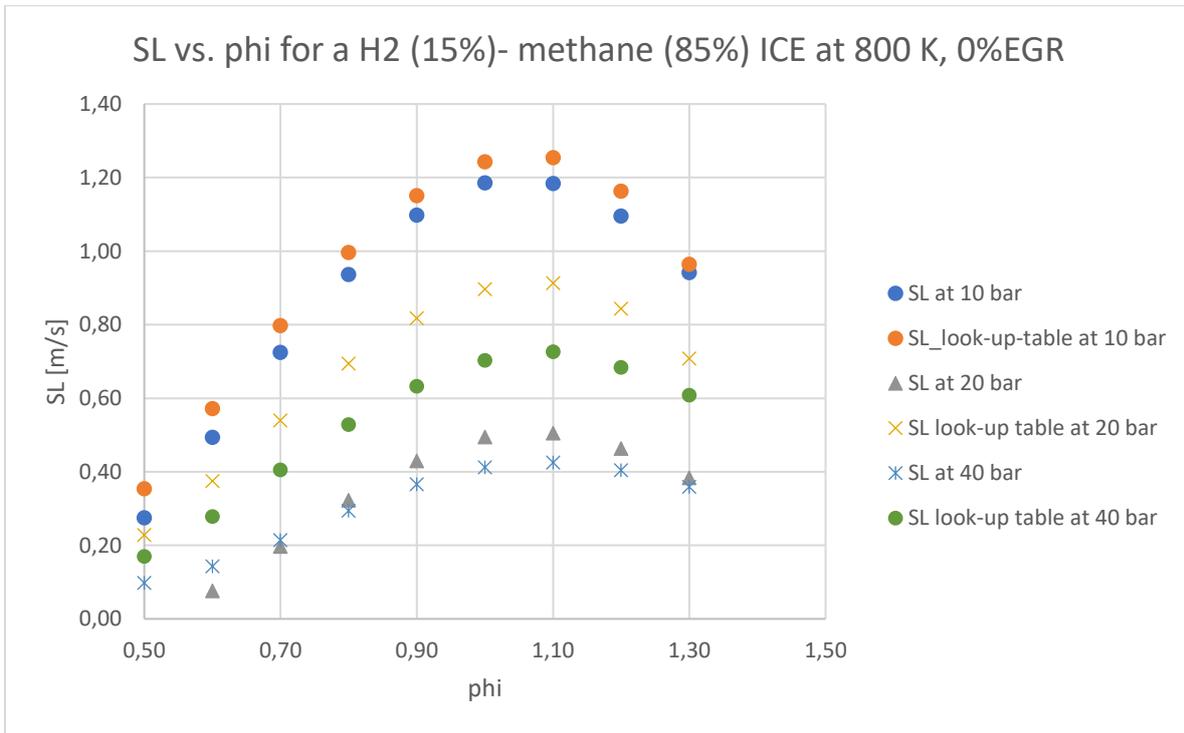


Figure 4.1.2, Laminar burning speed vs. equivalence ratio at 800 K: a) pure methane mixture with 0% EGR, b) comparison laminar burning speed of model vs. look-up table laminar burning speed at 0%EGR, c) pure methane mixture with 20% EGR, d) comparison laminar burning speed of model vs. look-up table laminar burning speed at 20%EGR



a)



b)

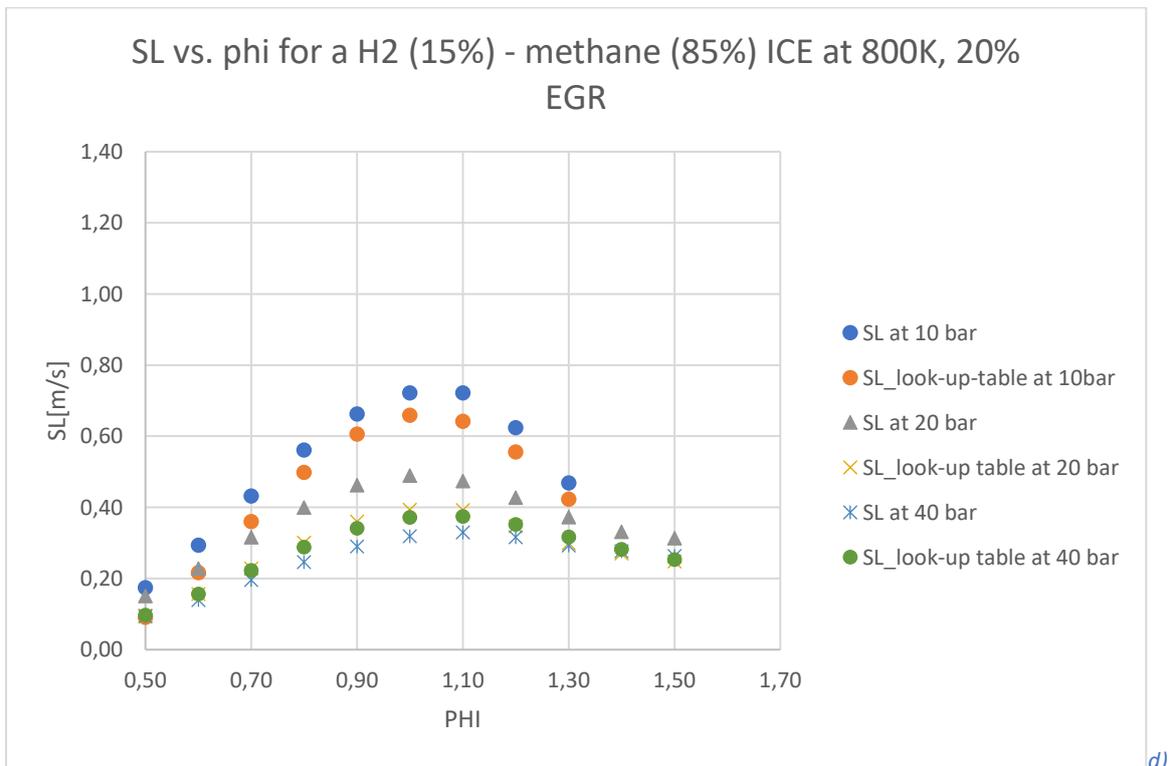
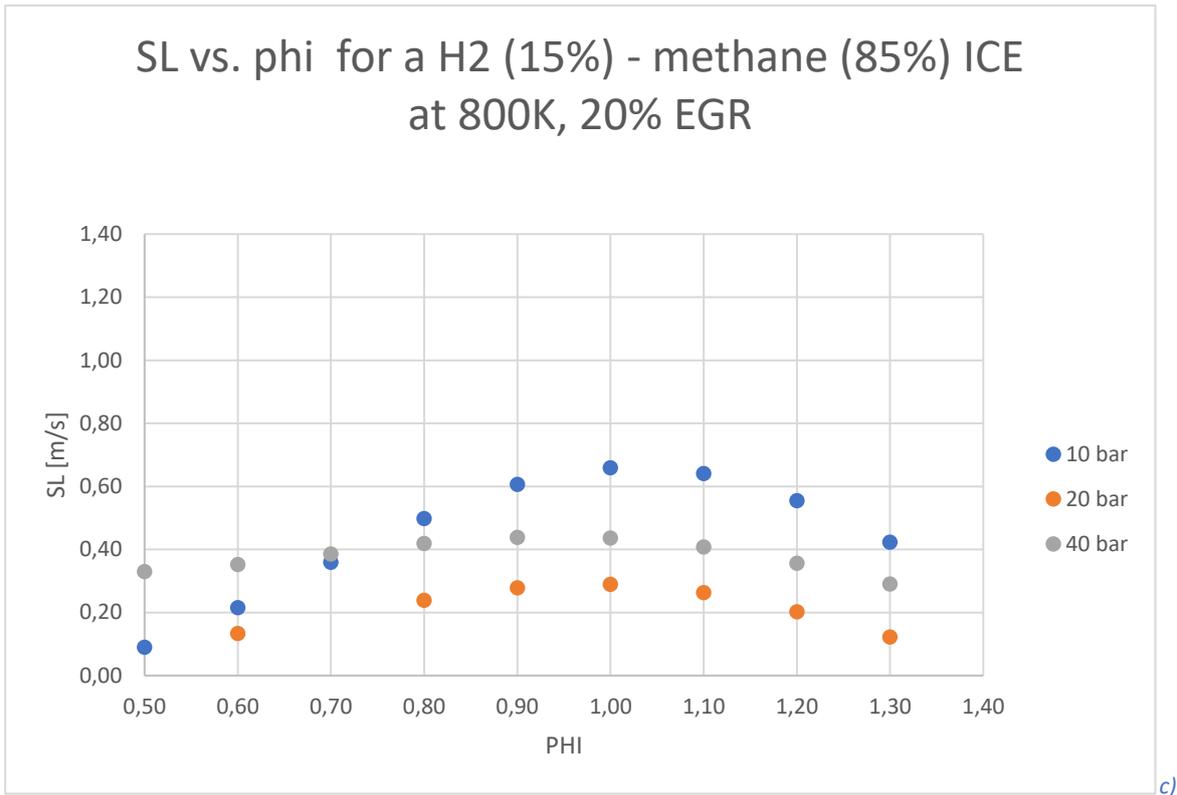


Figure 4.1.3, Laminar burning speed vs. equivalence ratio at 800 K: a) methane-hydrogen mixture with 0% EGR, b) comparison laminar burning speed of model vs. look-up table laminar burning speed at 0%EGR, c) methane-hydrogen mixture with 20% EGR, d) comparison laminar burning speed of model vs. look-up table laminar burning speed at 20%EGR

4.2 Laminar burning speed vs. pressure with Rational Quadratic GPR method

In this section, the analysis of the pressure effect on the laminar burning speed will be treated. It will be plotted in function of different level of pressure, considering, as in the previous section, the two different mixtures for this purpose. As consequence of the result reached in the section 4.1, the pressure growth tends to mitigate the propagation of flame. The curves, shown in the Figure 4.2.1, have a horizontal asymptotic trend when a certain pressure level is exceeded.

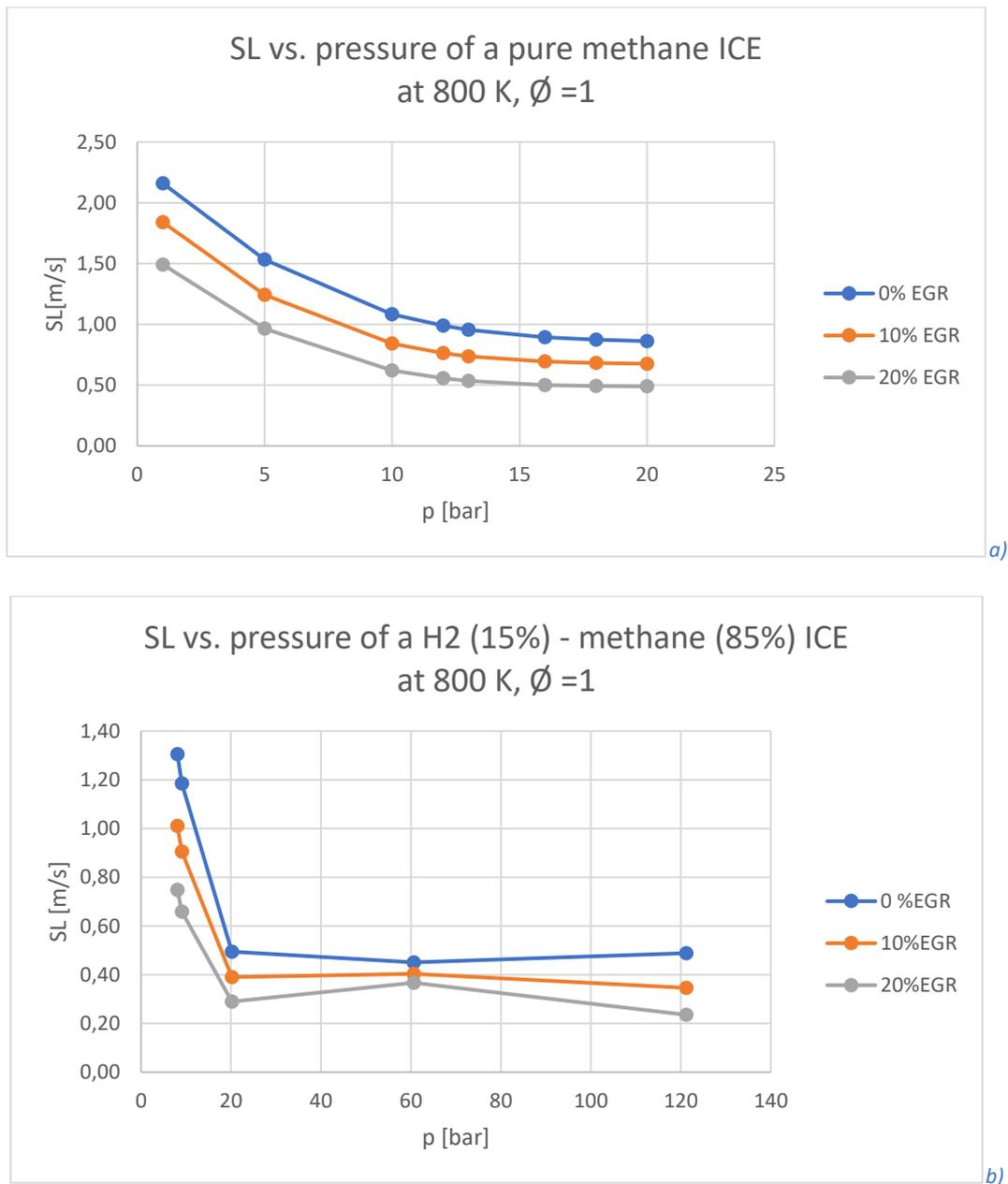


Figure 4.2.1, Laminar burning speed vs. pressure at 800 K: a) pure methane mixture with 0%,10%, 20% EGR; b) methane-hydrogen mixture with 0%,10%,20%EGR

4.3 Laminar burning speed vs. Temperature with Rational Quadratic GPR method

In this section, the analysis of the temperature effect on the laminar burning speed will be treated. The Figure 4.3.1 shows an increasing of laminar burning speed because of the unburned gas temperature rising. This variation of flame propagation is more evident at high temperature events.

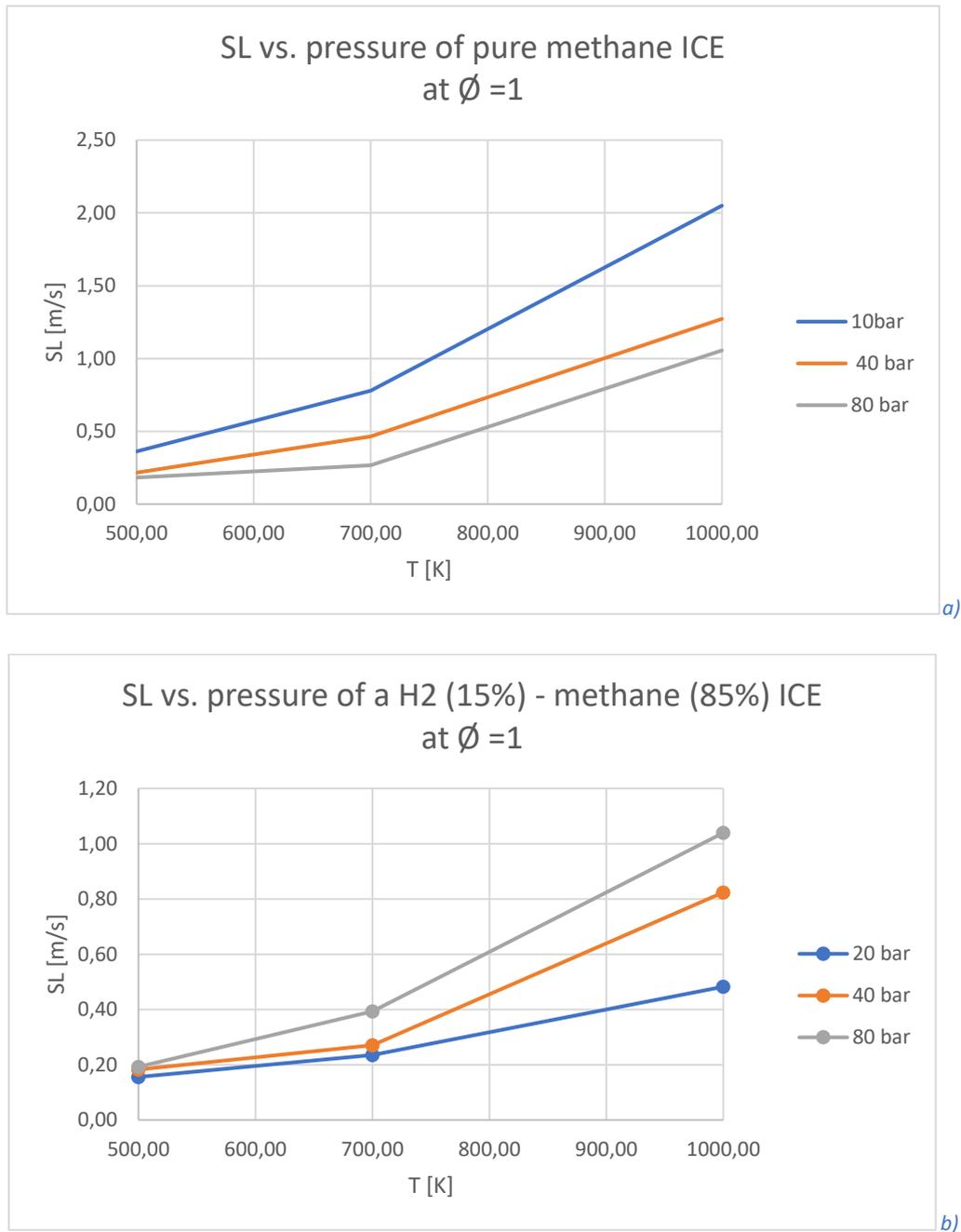


Figure 4.3.1, Laminar burning speed vs. Temperature at $\phi=1$: a) pure methane mixture at 20 bar, 40 bar, 80 bar; b) methane-hydrogen mixture at 20 bar, 40 bar, 80 bar

4.4 Laminar burning speed vs. EGR with Rational Quadratic GPR method

In this section, the analysis of the EGR effect on the laminar burning speed will be treated. The Figure 4.4.1 shows an exponential correlation and with a high level of residual in the combustion chamber, the laminar flame speed decreases.

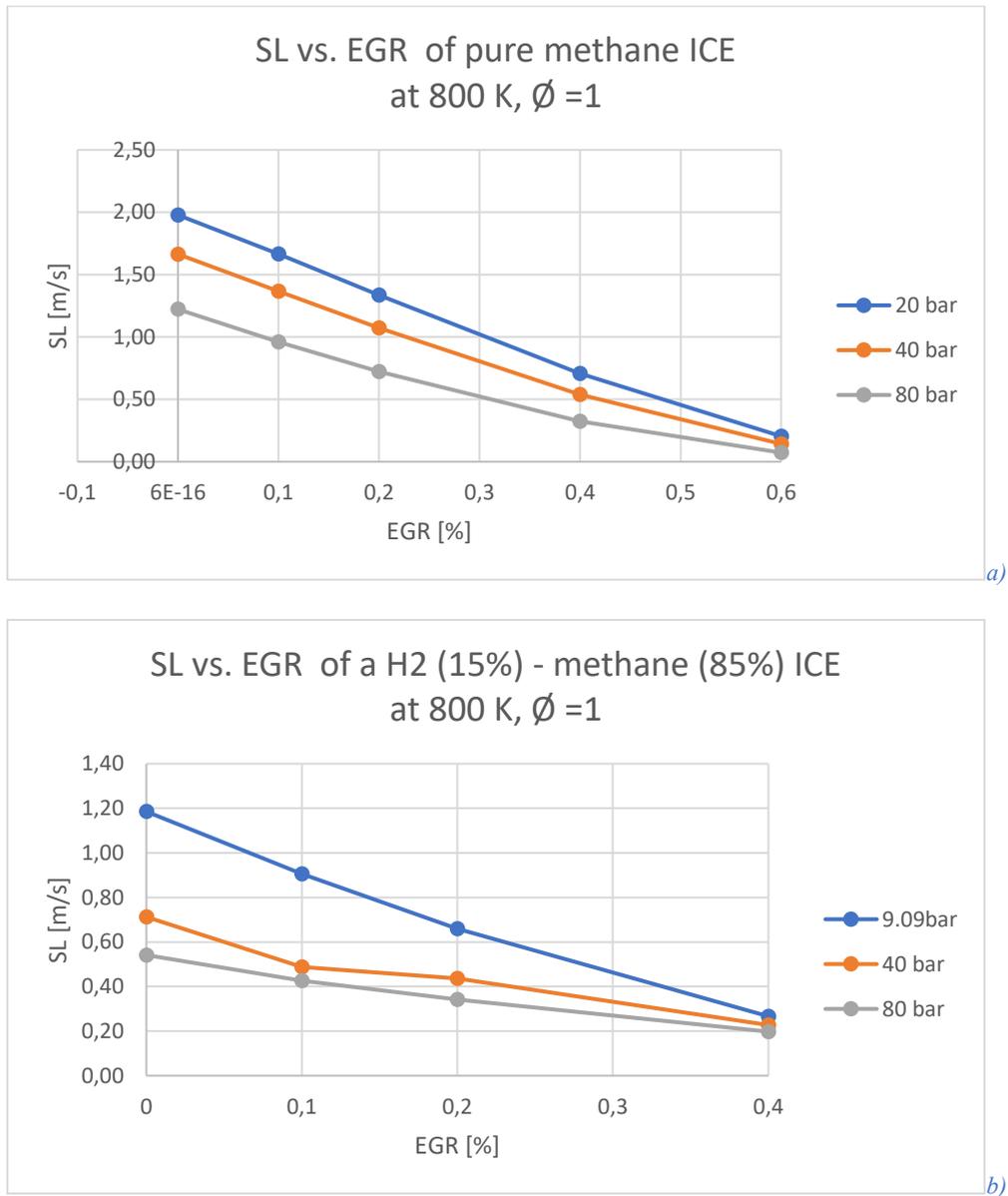
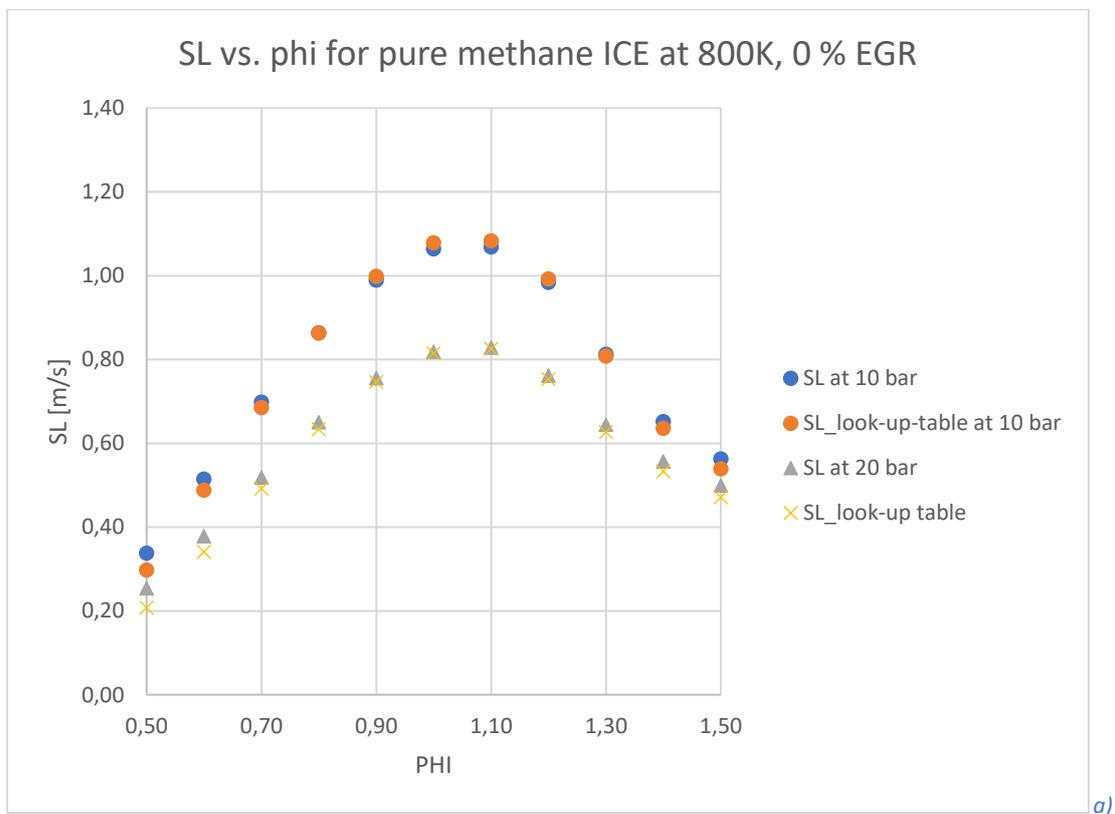


Figure 4.4.1, Laminar burning speed vs. EGR at $\phi=1,800K$: a) pure methane mixture at 9.09 bar, 40 bar, 80 bar; b) methane-hydrogen mixture at 9.09 bar, 40 bar, 80 bar

4.5 Laminar burning speed vs. Fuel-air Equivalence ratio with ANN

In this paragraph, in contrast to the 4.1, the plots refer to the laminar burning speed predicted through the ANN algorithm. As consequence, all the following sections will discuss results obtained with this last method. In addition to the two-mixture considered in the previous four paragraphs, there will be another one with a level of Hydrogen equal to 25%. Previously, during the exploiting of Rational Quadratic GPR method for the prediction of S_L , this third mixture is not taken into account because the results do not have a physical meaning; in fact, the values of laminar speed were negative and not acceptable. Despite this, the trend of laminar burning speed versus temperature, pressure, equivalence ratio and EGR rate has been conserved correctly. Maybe, in a future analysis, it will be possible to better understand why this last mixture generates this prediction errors. By the way, in the Figure there will be a comparison among curves that are moving in function of pressure level.



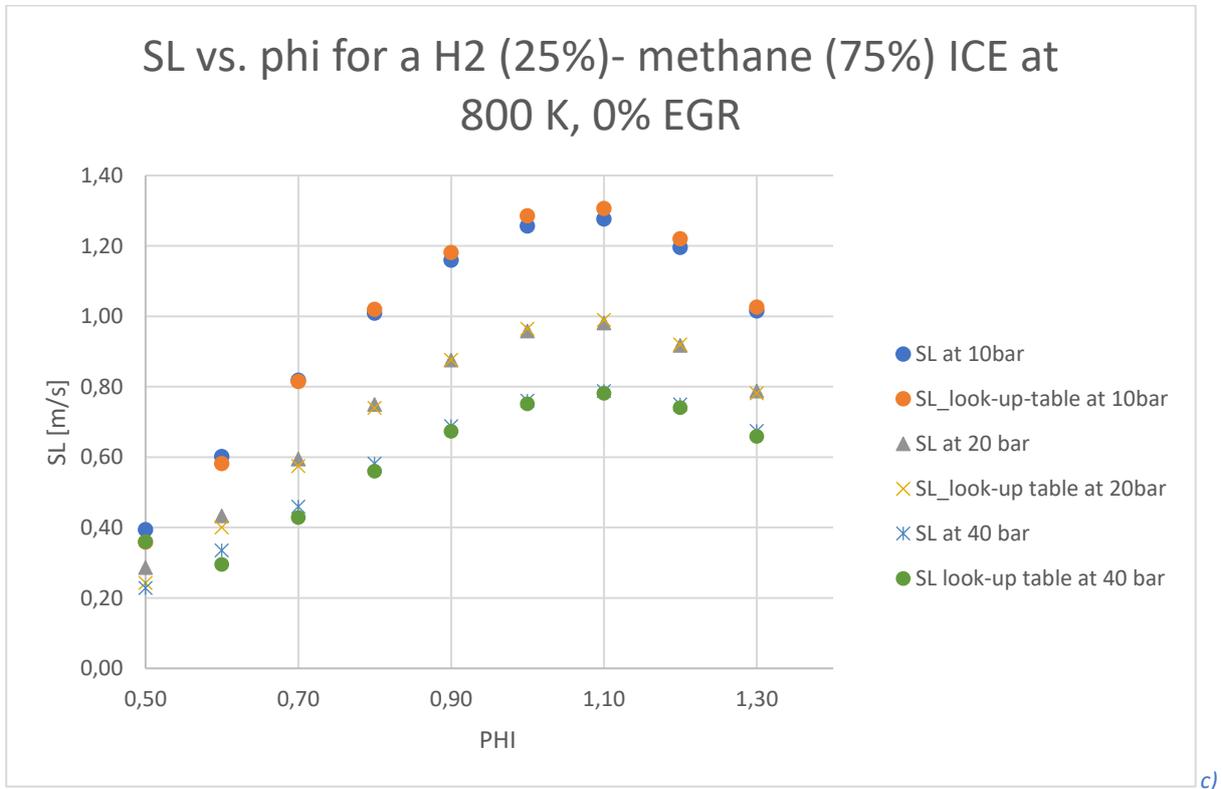
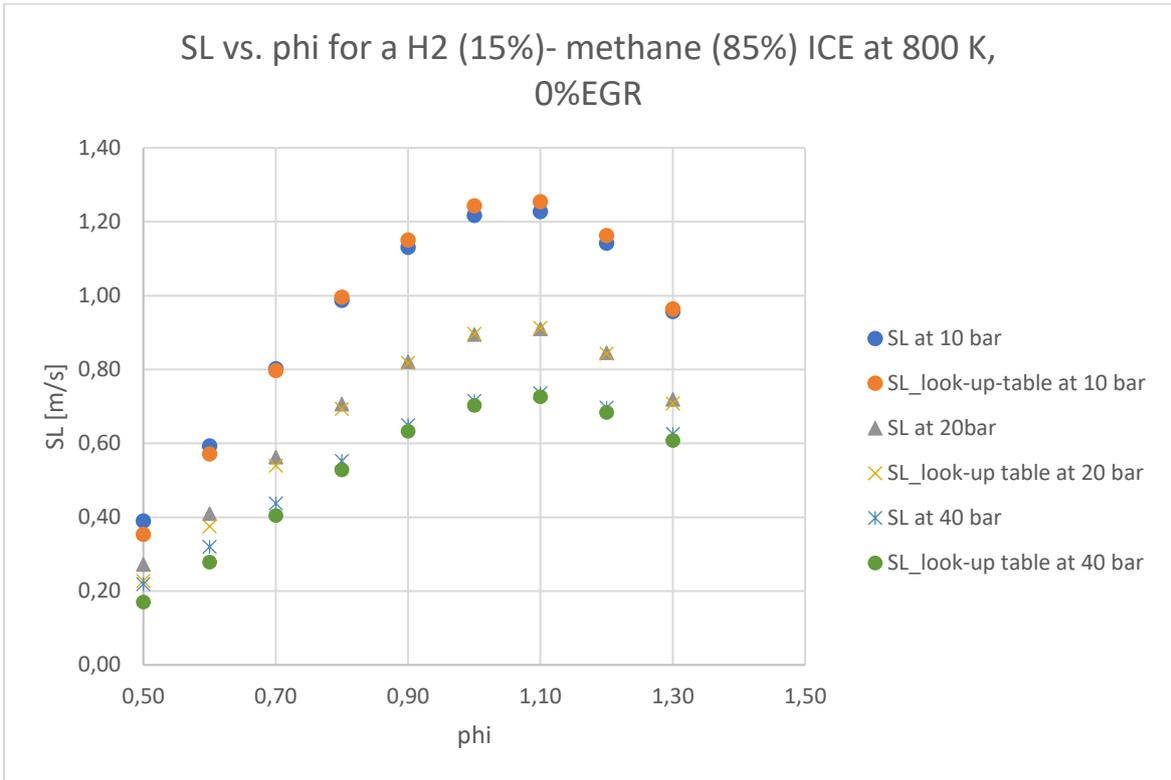
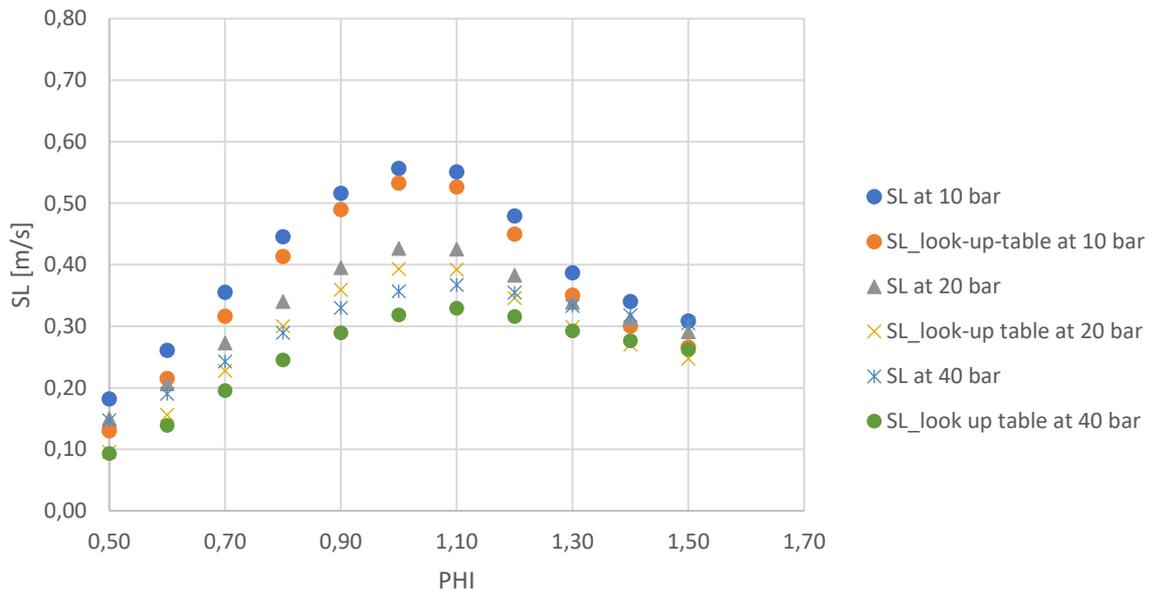


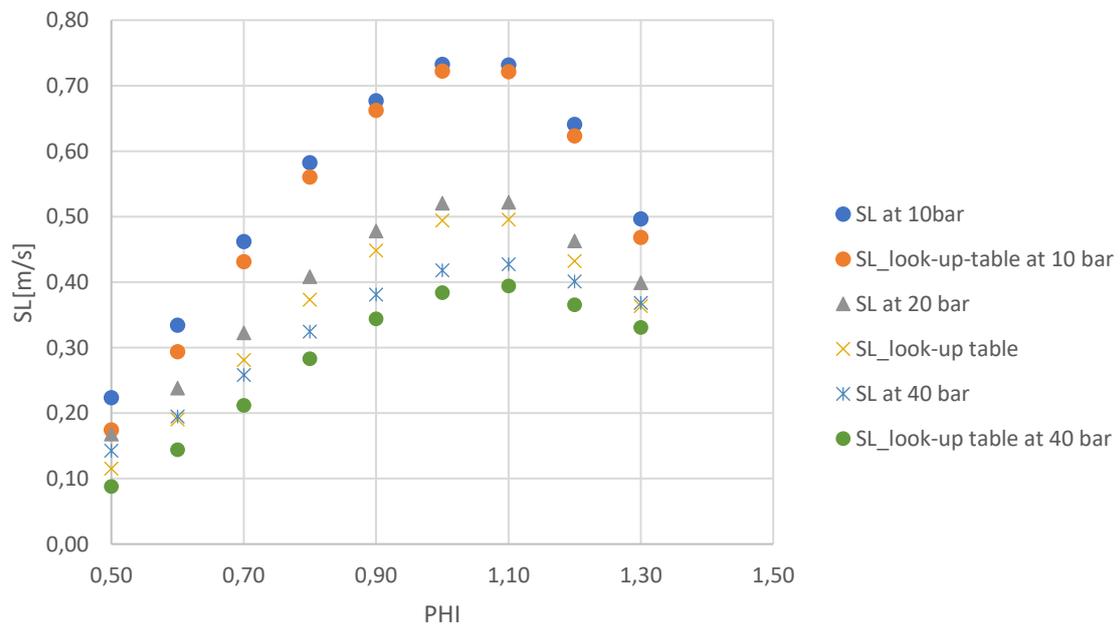
Figure 4.5.1, Laminar burning speed vs. equivalence ratio at 800K, 0% EGR: a) for a pure methane mixture; b) for a mixture with 85% CH₄ and 15% H₂; c) for a mixture with 75% CH₄ and 25% H₂

SL vs. phi for pure methane ICE at 800K, 20 % EGR



a)

SL vs. phi for a H2 (15%) - methane (85%) ICE at 800K, 20%EGR



b)

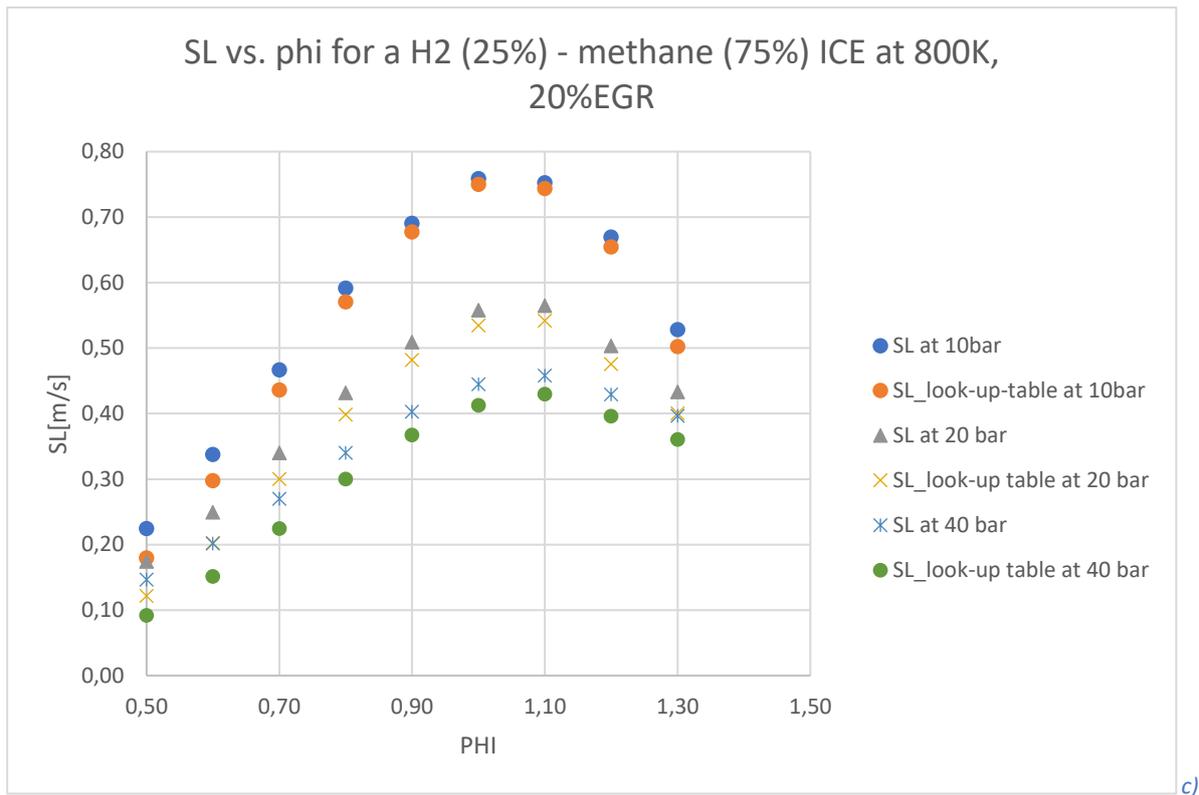


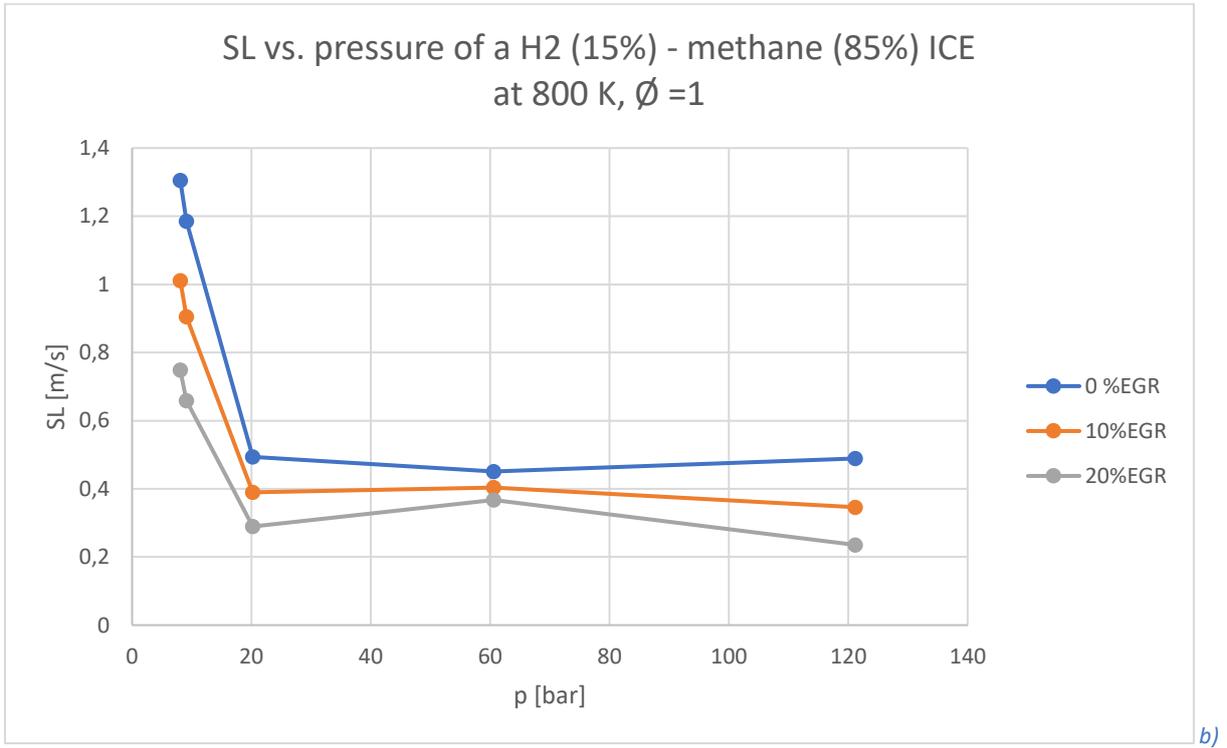
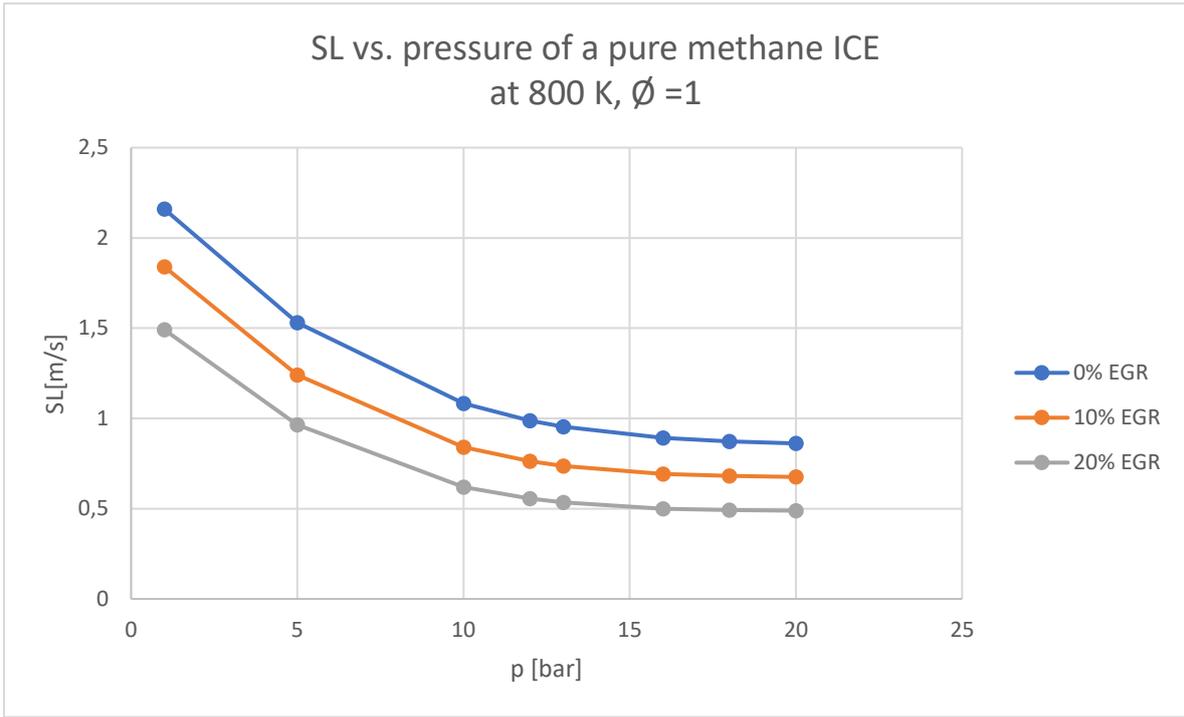
Figure 4.5.2, Laminar burning speed vs. equivalence ratio at 800K, 20% EGR: a) for a pure methane mixture; b) for a mixture with 85% CH₄ and 15% H₂; c) for a mixture with 75% CH₄ and 25% H₂

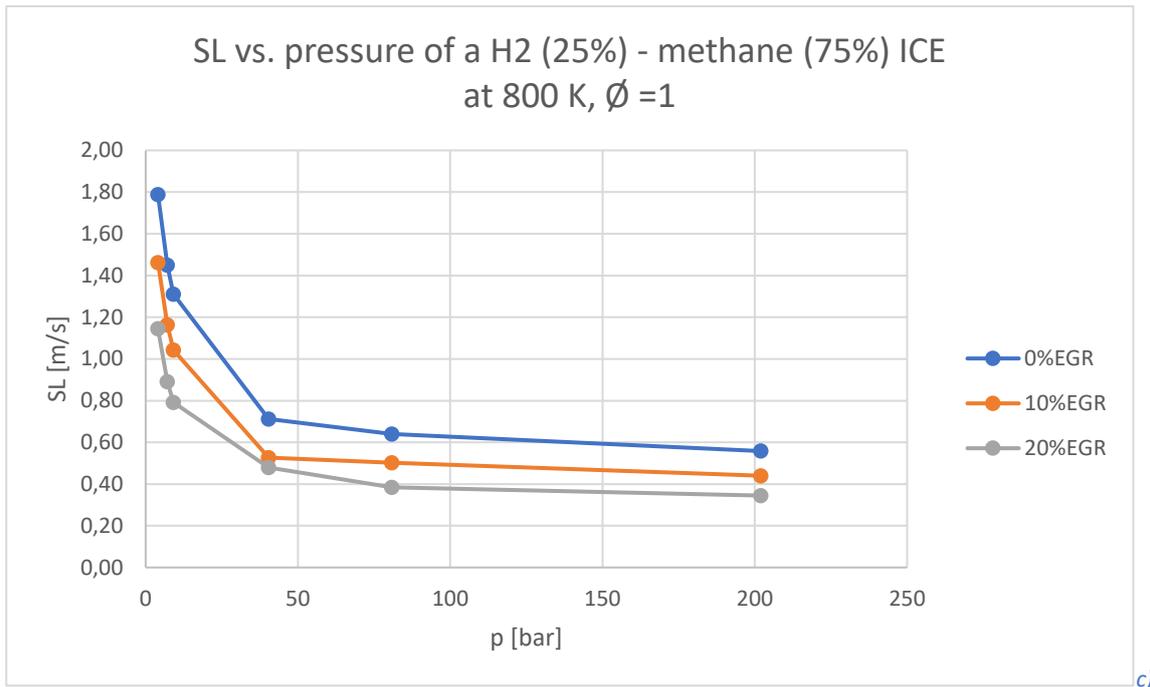
4.6 Laminar burning speed vs. pressure with ANN

In this section, the analysis of the pressure effect on the laminar burning speed will be treated. Different levels of pressure and EGR rates, for the three different mixtures, are considered for this purpose. The curves, shown in the Figure 4.6.1, have a horizontal asymptotic trend when a certain pressure level is exceeded.

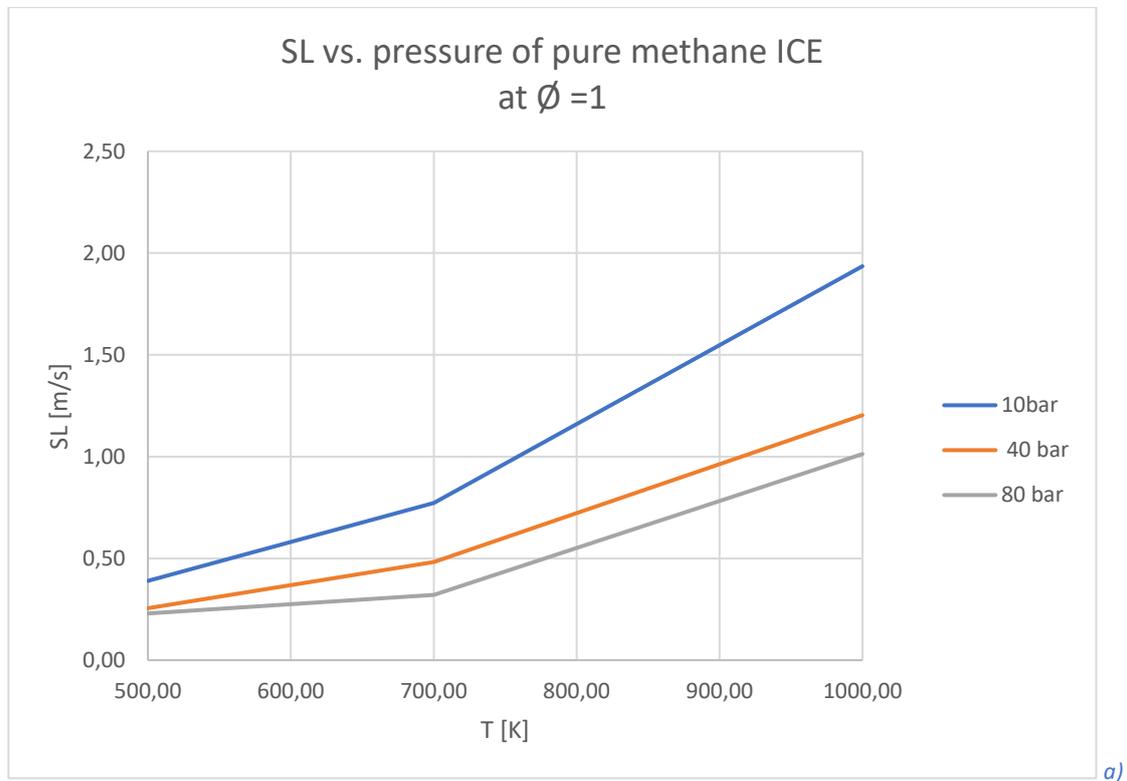
4.7 Laminar burning speed vs. Temperature with ANN

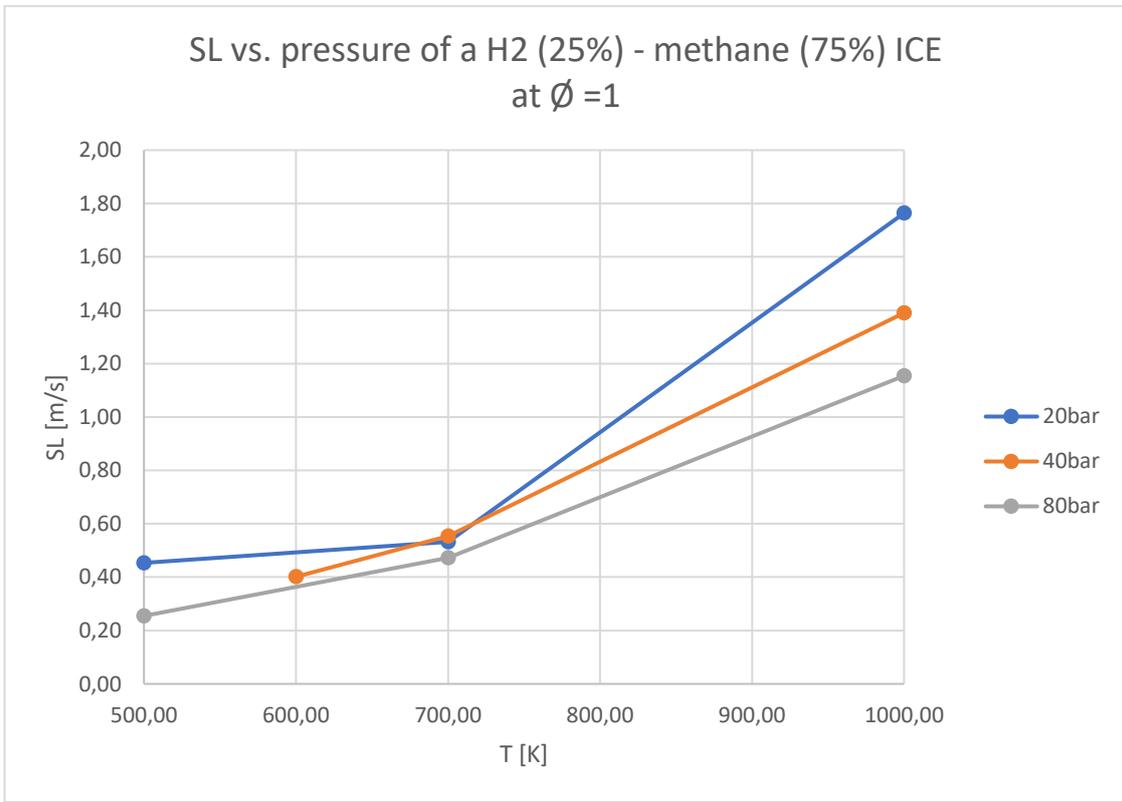
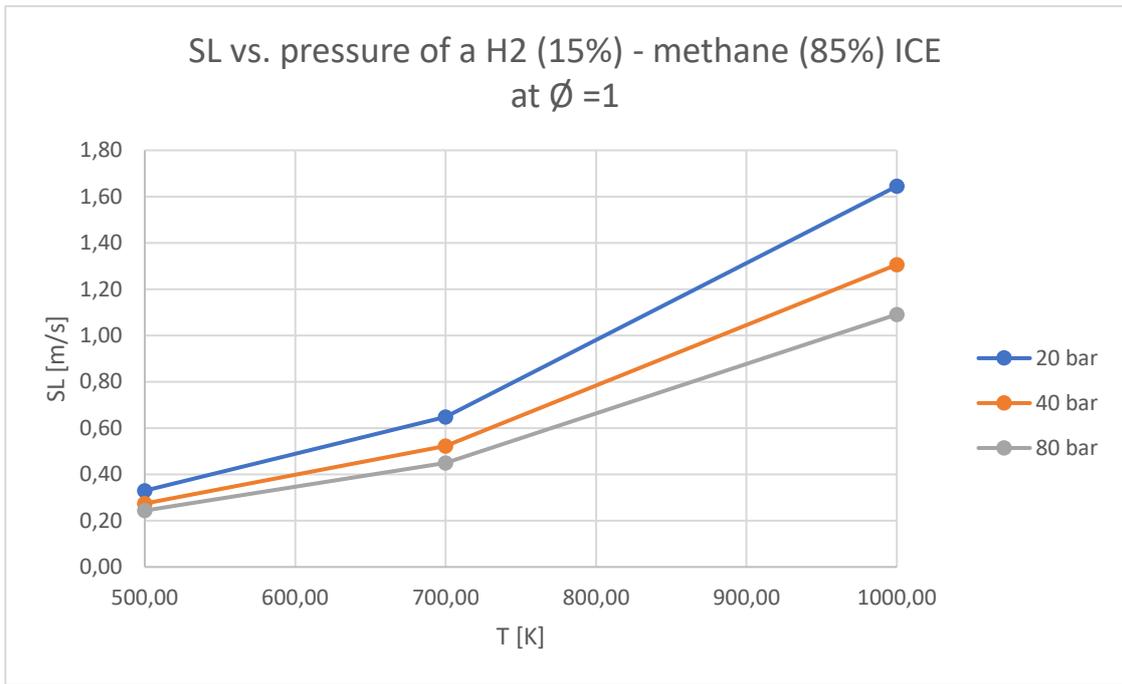
In this section, the analysis of the Temperature effect on the laminar burning speed will be treated. The Figure 4.7.1 shows an increasing of laminar burning speed because of the unburned gas temperature rising. This variation of flame propagation is more evident at high temperature events.





4.7.1, Laminar burning speed vs. pressure at 800 K: a) pure methane mixture with 0%,10%, 20% EGR; b) ; b) for a mixture with 85% CH₄ and 15% H₂ with 0%,10%,20%EGR ; c) for a mixture with 75% CH₄ and 25% H₂ with 0%,10%,20%EGR

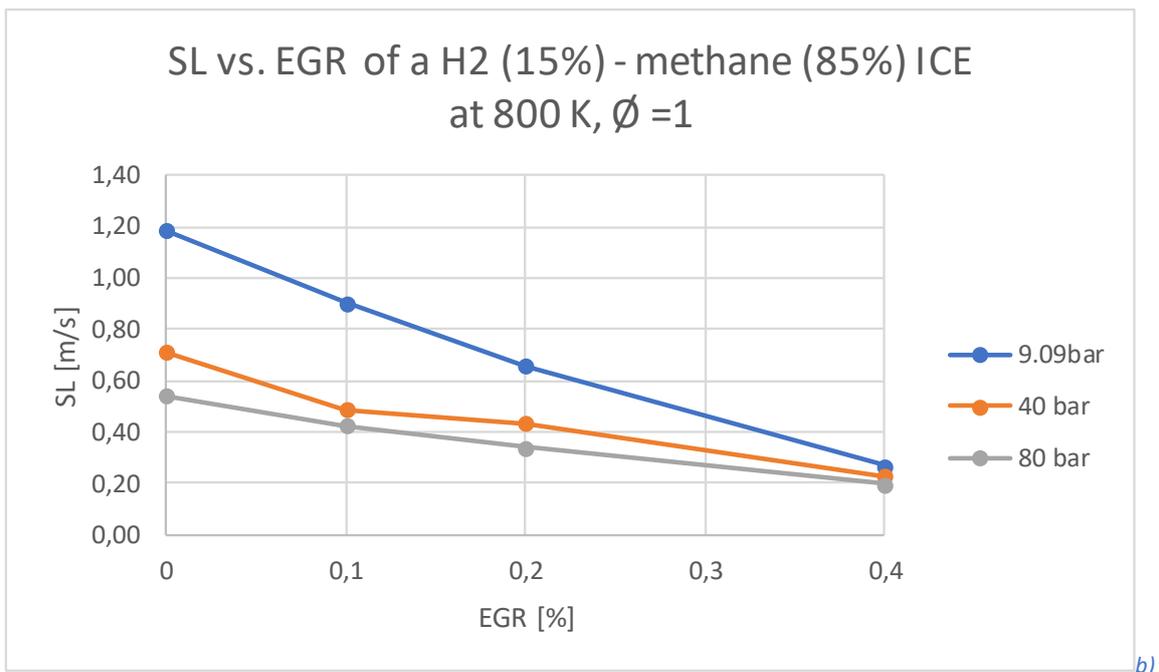
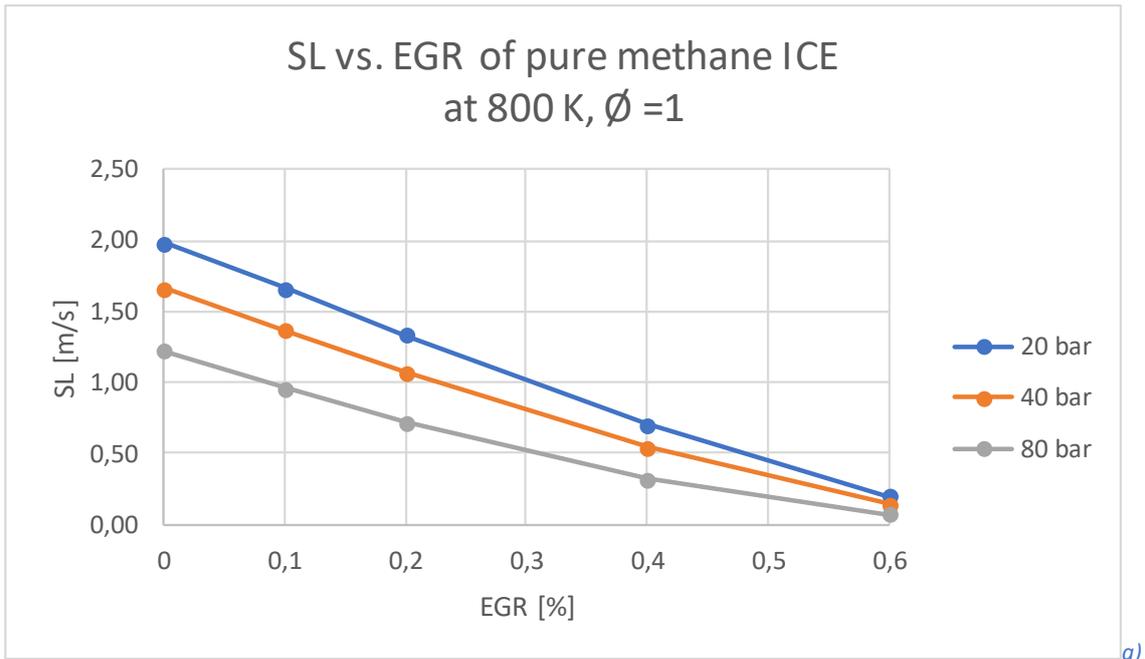




4.7.1, Laminar burning speed vs. Temperature at $\phi=1$: a) pure methane mixture at 20 bar, 40 bar, 80 bar; b) mixture with 85% CH₄ and 15% H₂ at 20 bar, 40 bar, 80 bar; c) mixture with 75% CH₄ and 25% H₂ at 20 bar, 40 bar, 80 bar

4.8 Laminar burning speed vs. EGR with ANN

In this section, the analysis of the EGR effect on the laminar burning speed will be treated. The Figure 4.8.1 shows an exponential correlation between EGR rate and S_L ; with a high level of residual in the combustion chamber, the laminar flame speed tends to decrease.



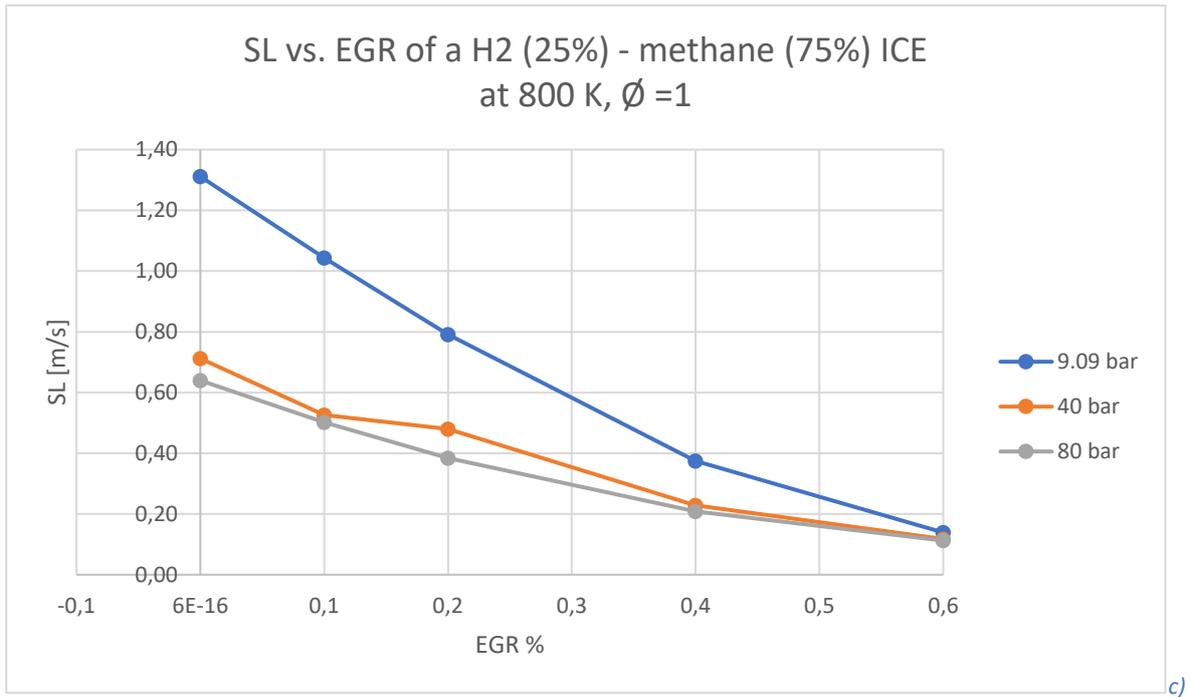


Figure 4.8.1, Laminar burning speed vs. EGR at $\phi=1$, 800K : a) pure methane mixture at 9.09 bar, 40 bar, 80 bar; b) mixture with 85% CH_4 and 15% H_2 at 9.09 bar, 40 bar, 80 bar; c) mixture with 75% CH_4 and 25% H_2 at 9.09 bar, 40 bar, 80 bar

4.9 Laminar burning speed vs. crank angle with Rational Quadratic GPR method

In this section, there will be a new laminar burning speed comparison between:

- SL obtained through the Huang et al. correlation, seen in the equation (12), exploiting both the input data and the obtained one of DARS Multi-zone model
- SL obtained through the Machine Learning model applied on DARS Multi-zone data

The input data are relative to the ensemble average 0, so the “average” behavior of the engine is considered instead to evaluate the relative one of each single cylinder. These are summarized in the Table 4.1:

Θ [deg]	T_u [K]	p_u [bar]	EGR rate[%]	phi[-]	X_{CH_4} [%]	X_{H_2} [%]
300°÷420°	300÷820	2.8 ÷53	3÷12	0.6÷1.4	>75%	0÷25%

Table 4.9.1, Input data used for the prediction

and the two working points at which the comparison is done are:

- 2000rpm x 600 kPa
- 3000 rpm x 800 kPa

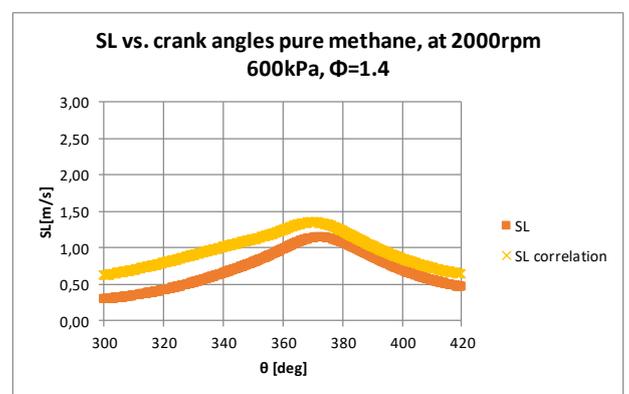
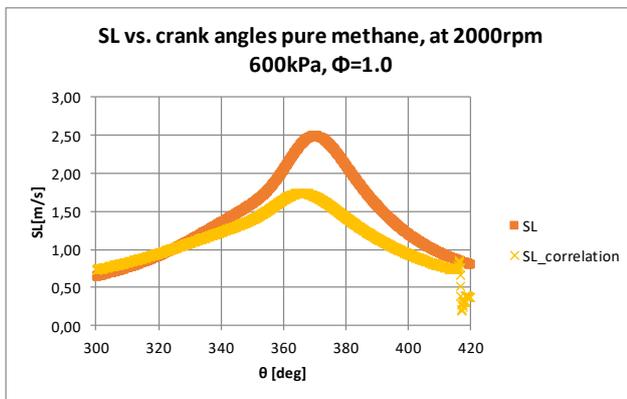
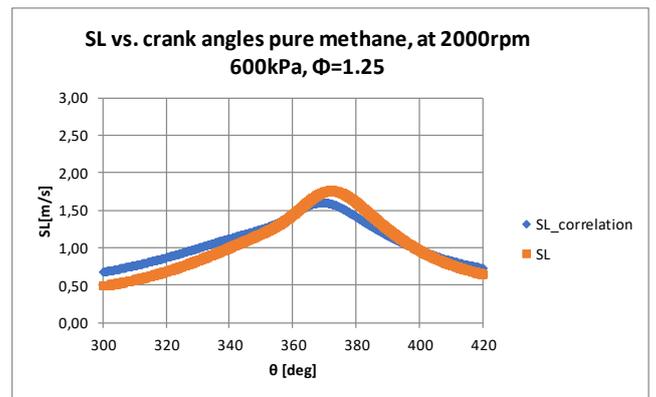
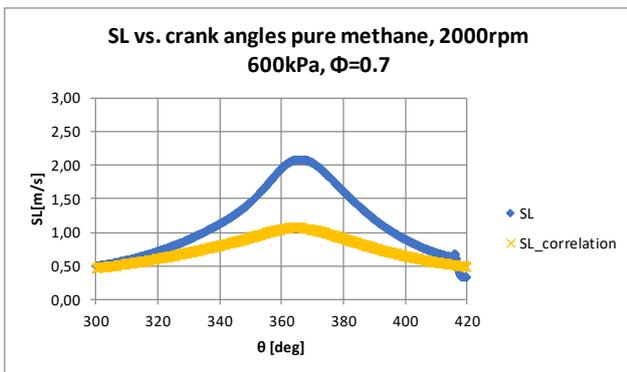


Figure 4.9.1., Laminar burning speed vs. Crank at 2000 rpm x 600 kPa for pure methane mixture comparing SL predicted through machine learning model and SL obtained from Huang et al. correlation

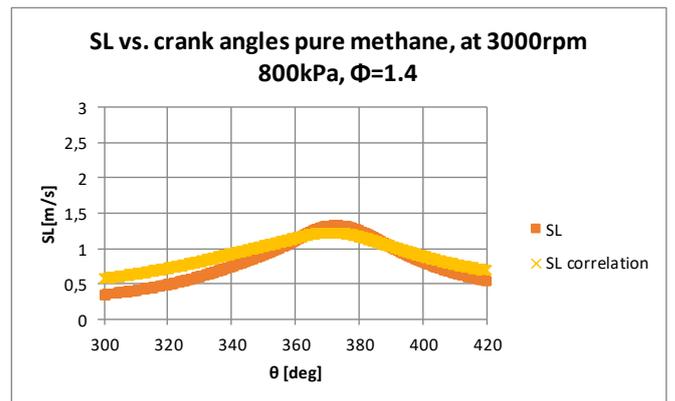
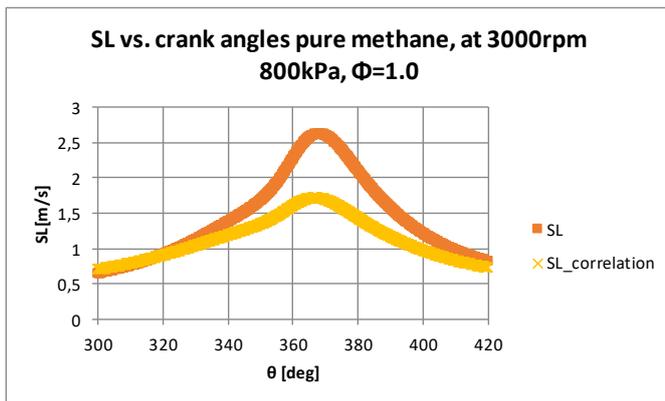
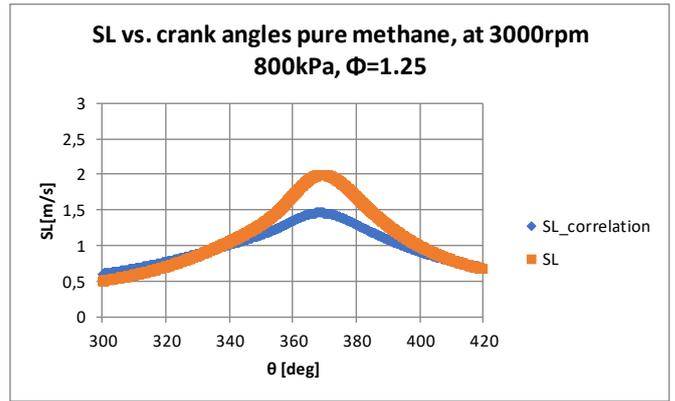
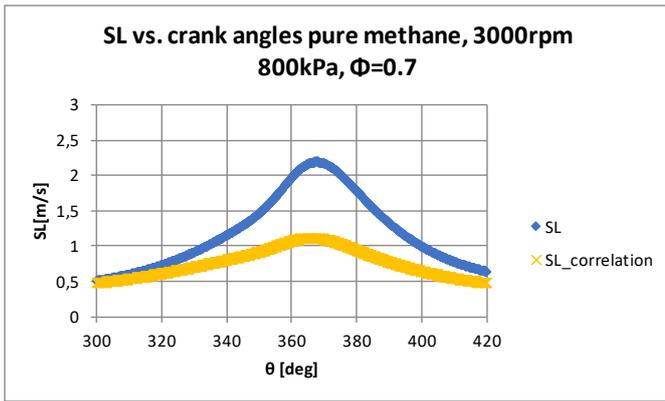


Figure 4.9.2., Laminar burning speed vs. Crank at 3000 rpm x 600 kPa for pure methane mixture comparing SL predicted through machine learning model and SL obtained from Huang et al. correlation

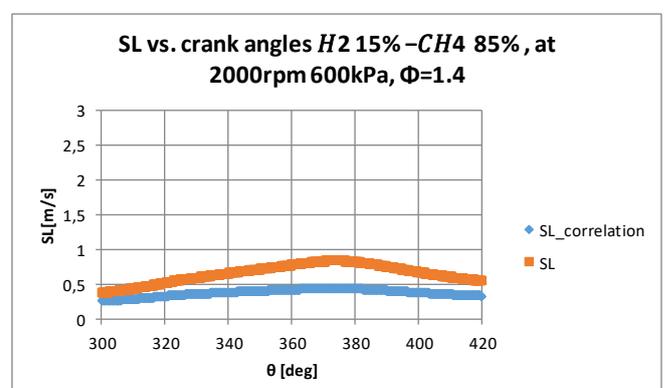
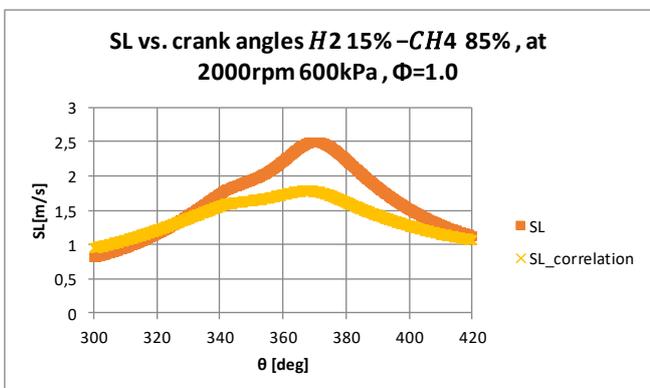
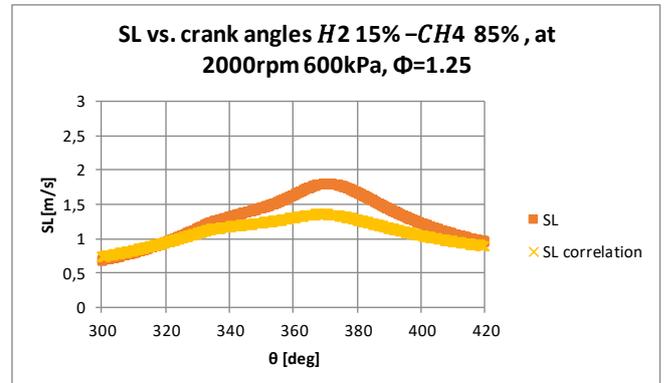
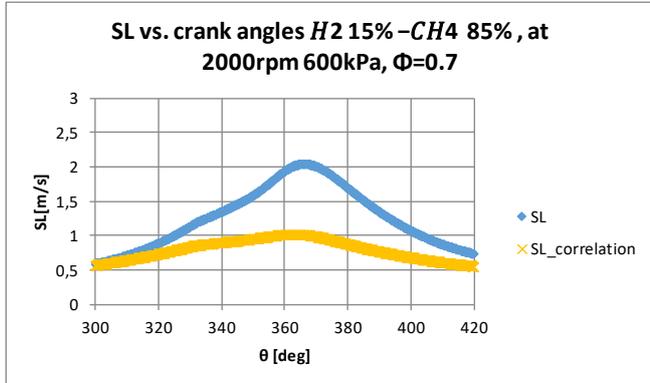


Figure 4.9.3., Laminar burning speed vs. Crank at 2000 rpm x 600 kPa for HCNG15 comparing SL predicted through machine learning model and SL obtained from Huang et al. correlation

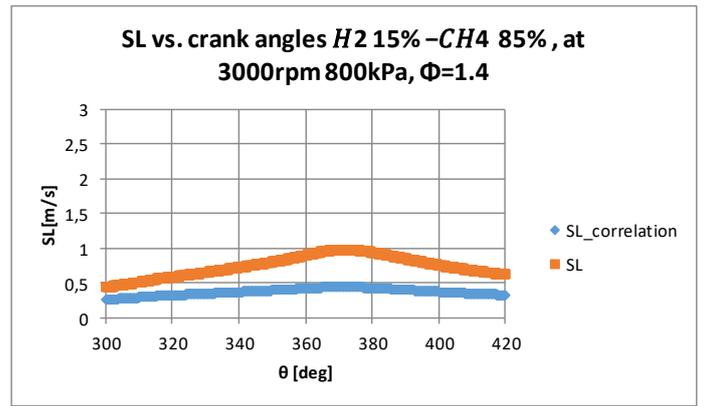
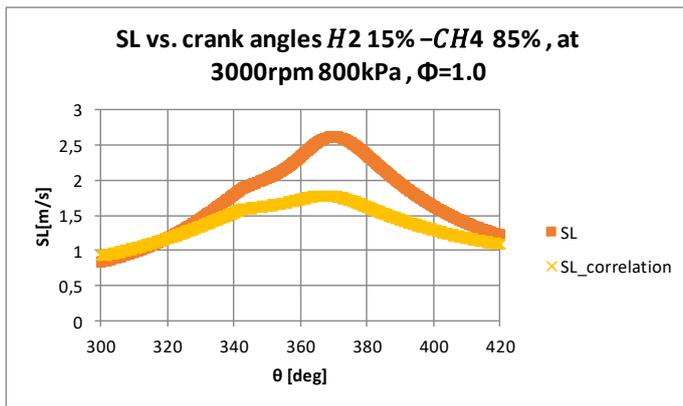
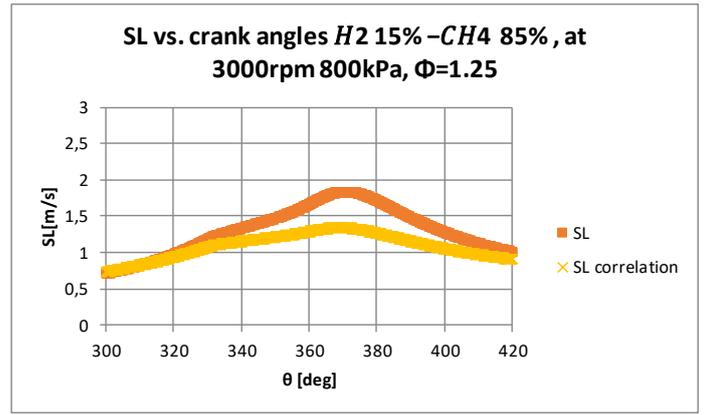
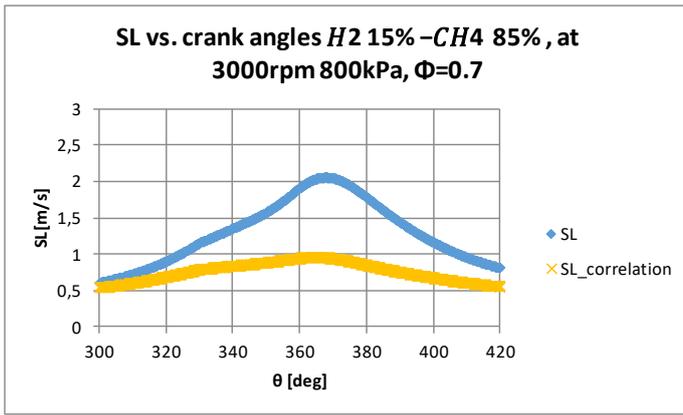


Figure 4.9.4., Laminar burning speed vs. Crank at 3000 rpm x 600 kPa for HCNG15 comparing SL predicted through machine learning model and SL obtained from Huang et al. correlation

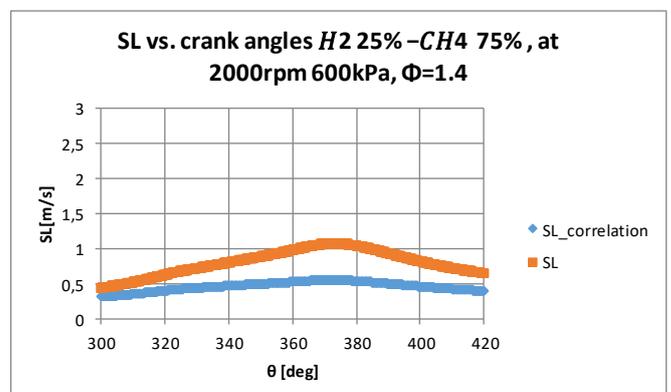
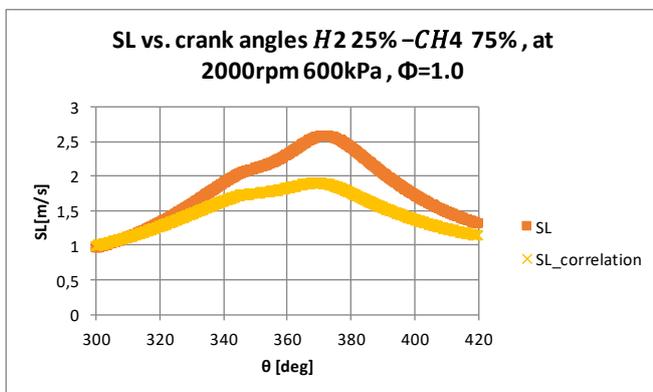
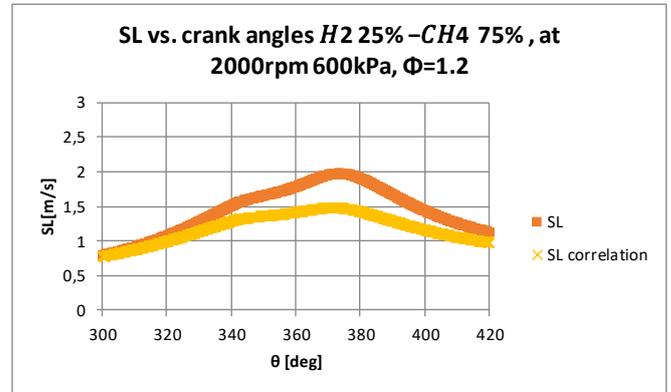
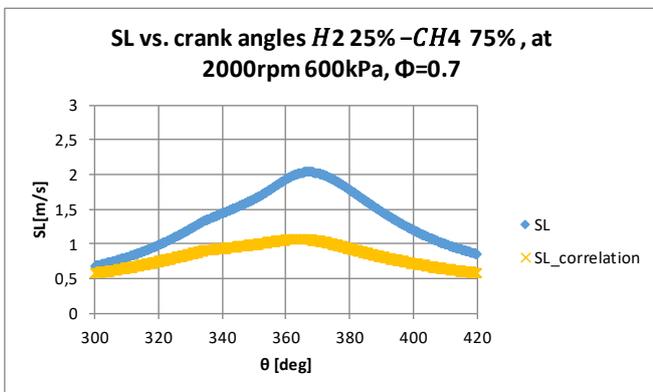


Figure 4.9.5., Laminar burning speed vs. Crank at 2000 rpm x 600 kPa for HCNG25 comparing SL predicted through machine learning model and SL obtained from Huang et al. correlation

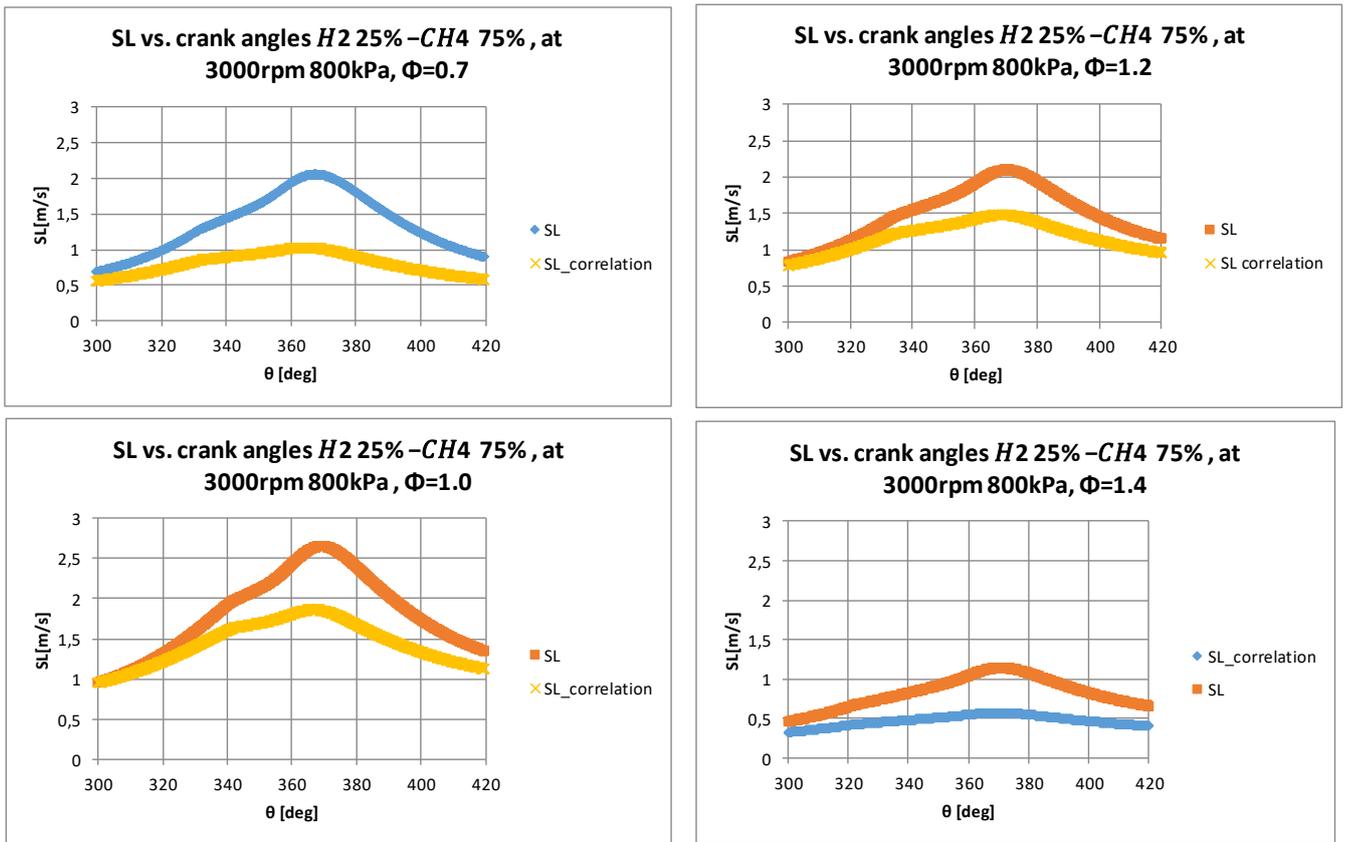


Figure 4.9.6., Laminar burning speed vs. Crank at 2000 rpm x 600 kPa for HCNG25 comparing SL predicted through machine learning model and SL obtained from Huang et al. correlation

From the Figure 4.9.6. it is possible to highlight how the laminar burning speed, forecasted by means of Machine Learning model, increases with amount of hydrogen concentration in the mixture. Moreover, all that in the Section 1.3 has been addressed from theoretical viewpoint of hydrogen contribution on methane mixture is confirmed.

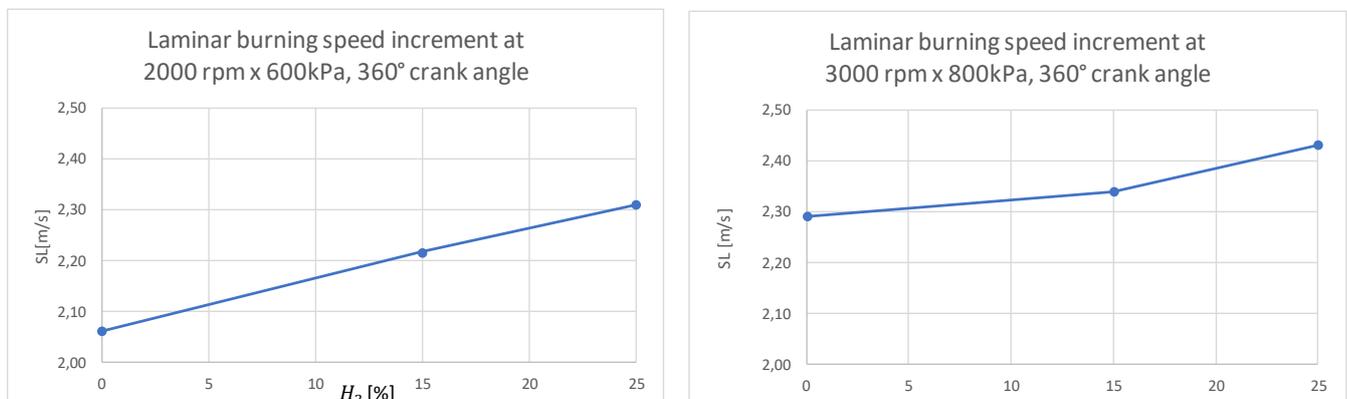


Figure 4.9.7., Laminar burning speed increment w.r.t. hydrogen rising concentration

2000rpmx600kPa		
H2[%]	SL[m/s]	SL increment
0	2,06	-
15	2,22	7,0%
25	2,31	10,8%

3000rpmx800kPa		
H2[%]	SL[m/s]	SL increment
0	2,29	-
15	2,34	2,1%
25	2,43	5,8%

Table 4.9.2, Laminar burning speed increment w.r.t. the hydrogen concentration

5. Conclusion and future development

The present work thesis provides a laminar burning speed predictive model obtained through supervised Machine learning. In the specific, the Regression Learner Matlab App is used to train the regression model and forecast the laminar burning velocity. The used input data, organized as a look-up-table, are extrapolated by means of the GRI 3.0 mechanism on DARS for different pressure, unburned gas temperature, fuel-air equivalence ratio, EGR rate, methane and hydrogen concentration in mixture. These are used in the app to select the validation scheme, train the model and assess the best regression model. As consequence, this latter can generate predicted responses for each new dataset if they are organized with the same number of input argument given during the training activity processed in the app. After that, the prediction capability of this model is checked making another comparison, in correspondence of two specific working points, between the laminar burning speed obtained by the application of machine learning algorithm on DARS Multi-zone data and the one obtained from DARS Multi-zone model applying the Huang and al. correlation for the laminar burning velocity computation. In addition, it is analyzed also the possibility to design an Artificial Neural Network, by means of Matlab Fitting App, extending it at the same two case previously highlighted for the supervised machine learning. The main results can be summarized as follows:

1. The regression model designed thanks to the machine learning algorithms, has not only a certain mathematical meaning, in fact its goodness of fit is around 0.9, but also a physical one and this is confirmed by results shown in the figures of Chapter 4. Indeed, as it is expected the laminar burning speed trend is bell-shaped reaching the maximum in the slight lean condition, decreases with rising of unburned gas pressure with a horizontal asymptotic trend and has an exponential correlation with the unburned gas temperature. Moreover, it is always possible to highlight from figures of same chapter the positive contribution of hydrogen in the mixture treated in the Chapter 1.
2. The discontinuity points present in the look-up-table laminar burning speeds generated by GRI 3.0 mechanism on DARS disappear in the ones predicted by supervised machine learning as shown in the Chapter 4.

3. When the predictive model is managing the DARS Multi-zone data, the major differences in terms of laminar burning velocities are detected in the field of leaner mixture, where the hydrogen high burning speed improves the pure methane behavior.
4. The choice of programming language for machine learning is quite important; due to some limits of Matlab app in terms of code “flexibility” and limited integration with other languages, in the future plans it is expected to use Python. This could also give the possibility to better exploit the ANN tool, since with the Fitting Matlab App there was only the opportunity to execute a curve fitting.

Nomenclature

deg = degree

CNG = Compressed natural gas

SI = Spark Ignition

DISI = Direct Injection Spark Ignition

ICE = Internal Combustion Engine

HCNG = Hydrogen and CNG blend

EGR = Exhaust gas recirculation

CO = Carbon monoxide

CO_2 = Carbon dioxide

HC = unburned hydrocarbon

NO_x = Nitrogen oxide level at engine exhaust

bmep = brake mean effective pressure

n = engine speed

p = pressure

T = temperature

p_u = unburned mixture pressure

T_u = unburned mixture temperature

V_u = unburned mixture volume

p_o = room pressure

T_o = room temperature

S_L = laminar burning speed

$S_{L,o}$ = laminar burning speed at room condition

Z, W, η , ξ and σ = Gulder's correlation coefficients

α = influence exponent

β = influence exponent

X_{CH_4} = methane volume fraction in the mixture

X_{H_2} = hydrogen volume fraction in the mixture

phi = Φ = fuel-air equivalence ratio

h = enthalpy

ρ = density

θ = crank angle

x_b = mass fraction burned

TDC = top dead center

BDC = bottom dead center

m = mass

q = heat flux transfer

\tilde{h} = convective heat transfer coefficient

CFD = Computational fluid dynamics

ANN = artificial neural network

GPR = Gaussian process regression

RMSE = root mean square error

σ_f = standard deviation

ε = error term

R_2 = the goodness of fit

Subscripts

b = burned

u = unburned

i = summation index

res = residual gas

Bibliography

- [1] Amirante R, Distaso E, Tamburrano P, Reitz RD. *Laminar Flame Speed Correlations for Methane, Ethane, Propane and their Mixtures, and Natural Gas and Gasoline for Spark-Ignition Engine Simulations*. Int J Engine Res 2017: In press.
- [2] Gülder ÖL. *Correlations of laminar combustion data for alternative SI engine fuels*. SAE Tech Pap 841000 1984.
- [3] R.Amirante, E.Distaso, P.Tamburrano, R.D.Reitz. *Analytical Correlations for Modeling the Laminar Flame Speed of Natural Gas Surrogate Mixtures*
- [4] Patricia Dirrenberger, Hervé Le Gall, Roda Bounaceur, Olivier Herbinet, Pierre-Alexandre Glaude, et al. *Measurements of Laminar Flame Velocity for Components of Natural Gas*. Energy and Fuels, American Chemical Society, 2011, 25 (9), pp.3875-3884.
- [5] Y-F Liu, B Liu, L Liu, K Zeng, and Z-H Huang. *Combustion characteristics and particulate emission in a natural-gas direct-injection engine: effects of the injection timing and the spark timing*. State Key Laboratory of Multiphase Flow in Power Engineering, Xi'an Jiaotong University, Xi'an, People's Republic of China. April 2010. DOI: 10.1243/09544070JAUTO1532
- [6] Akhilendra Pratap Singh, Anuj Pal, Avinash Kumar Agarwal. *Comparative particulate characteristics of hydrogen, CNG, HCNG, gasoline and diesel fueled engines*. Engine Research Laboratory, Department of Mechanical Engineering, Indian Institute of Technology Kanpur, Kanpur 208016, India. August 2016.

- [7] Henry J. Curran. *DETAILED CHEMICAL KINETIC MODELING; IS THERE LIFE AFTER GRI-MECHANISM 3.0?* National University of Ireland, Galway, Ireland
- [8] Shahab Mohaghegh. *VIRTUAL INTELLIGENCE AND ITS APPLICATIONS IN PETROLEUM ENGINEERING*. West Virginia University, 2010.
- [9] Saurabh Karsoliya. *Approximating Number of Hidden layer neurons in Multiple Hidden Layer BPNN Architecture*. International Journal of Engineering Trends and Technology- Volume 3 Issue 6- 2012
- [10] Lutz Prechelt, *Early Stopping – But When?*
- [11] C. E. Rasmussen, *Gaussian processes for machine learning*
- [12] Catania, A. E., Misul, D., Mittica, A., and Spessa, E., 2003, “A Refined Two-Zone Heat Release Model for Combustion Analysis in SI Engines,”
- [i] www.machinelearningmastery.com
- [ii] www.towardsdatascience.com
- [iii] www.analyticsvidhya.com
- [iv] www.mathworks.com