

Politecnico di Torino

Department of Mechanical and Aerospace Engineering

Master Degree in Automotive Engineering



Master Thesis

Development and validation of a Mean-Value Model, in GT-Power environment, for the real-time simulation of a 2.3L light-duty Diesel engine

Supervisors

Prof. Roberto Finesso

Prof. Stefano D'Ambrosio

Candidate

Marco Agostini

July 2020

and validation

Value Model,



*Se insisti e resisti
raggiungi e conquisti*

environment,

me simulation

ity Diesel engine

The target of this Master Thesis is to develop and validate the Real-Time ID model of the FIA Diesel engine provided by FPT Industrial. The engine is a prototype for light-duty application also used for studying PCCI combustion process. The Real-Time model is instead developed with GT-Power, the simulation software by Gamma Technologies LLC.

Nowadays, about the 55% of the world oil demand is due to the overall transport system, whereas the road transport accounts for about 43% of the demand. This intense exploitation of internal combustion engines causes the increase of air pollution due to the concentration of solids, liquids or gases residuals that have negative impact on the surrounding environment and people. For this reason, fuel-efficiency standards for new vehicle sales are expected to broaden continuously. The car-maker are then forced to increase efficiency and reliability of the engines through engine improvements, new technologies, refined after-treatment systems and mostly electrification. Therefore the preliminary design and control of the engines represents the very efficient solution to save time and money. The main objective for electronic Diesel engine control systems is to provide the required torque with minimal fuel consumption under the constraint of matching the exhaust gas and noise emission regulation. At early stages, the development of the controllers goes at equal pace with the development of the virtual engine model. To this purpose, it is required that engine models run as fast as possible to implement the operation of the controllers. Again, this Thesis aims to create the Real-Time, virtual engine model which will be then used to perform the tests together with the controllers.

This project has been carried out on a FPT Diesel engine EURO VI b, in line 4-cylinder, turbocharged by a variable geometry turbine and provided of Common Rail injection system. The experimental data and the provided tests refer to 103 stationary engine operative points that cover the overall engine map.

First, the Thesis runs through the working principle of both the physical and virtual environments. Thus, the first two chapters refer to the compression ignition engine, its combustion process, the kind of pollutions, the injection technologies, and the actual FPT FIA engine. On the other hand, the third chapter deals with the working principle of the simulation software.

All these information merge into the fourth chapter, in which the real engine is reproduced in the virtual environment as close as possible to reality. However, the more the faithful between the real engine and the model, the higher the computational time required to perform any kind of test in the following design stage. With this in view, the chapter compares the performance of the Detailed engine, which must comply with the experimental data (the data measured on the real engine), with those of the Fast-Running model, which in turns must be compliant with the result of the Detailed model. The fifth chapter introduces the engine model to the controllers, their functioning, principles and successive test phases.

The sixth chapter represent the core of the activity. Here, the Fast-Running model have been further developed to obtain an even faster model, that is, the Mean Value Model. The development of this model passes through the Design of Experiment, the training of the Neural Network and the calibrations to match the Mean Value model results with those of the Detailed and Fast-Running model.

Eventually, the last chapter evolves the Mean Value Model into the Real-Time model, thus reducing the overall Factor of Real Time, which is the ultimate parameter that states the response of the model concerning computational time.

ABS
STR
ACT

1	Compression Ignition Engines	1	2	FPT FIA CI Engine	61
	1.1 Introduction	2		2.1 Engine Characteristics	62
	1.2 Standard Combustion	4			
	1.3 J.E Dec Model	10			
	1.3.1 Describing DecModel	17			
	1.3.2 Diffusion Flame development	18			
	1.3.3 Time-history graph	20			
	1.4 Non-conventional combustion	22			
	1.4.1 HCCI	22			
	1.4.2 PCCI	24			
	1.4.3 Kamimoto Bae diagram	29			
	1.5 Pollutant emissions	32	3	GT-Power Simulation Software	73
	1.5.1 Nitrogen Oxides - NO _x	32		3.1 GT-Power software	74
	1.5.2 Particulate Matter - PM	42		3.2 Combustion Models	76
	1.5.3 Unburned Hydrocarbons - HC	48		3.2.1 Non-predictive model	78
	1.5.4 Carbon Monoxide - CO	50		3.2.2 Semi-predictive model	78
				3.2.3 Predictive model	79
	1.6 Injection Technologies	50		3.3 DI-Pulse Combustion model	79
	1.6.1 Solenoid CR injectors	51		3.4 Burn Rate Analysis	82
	1.6.2 NOD, NCD, ITL	53		3.4.1 CPOA	83
	1.6.3 Multiple injection events	57		3.4.2 TPA	84

CON
TEN
TS

4	FIA Engine Models in GT-Power	87	6	The Mean value model	133
	4.1 Detailed Engine model	88		6.1 Design of Experiment - DoE	135
	4.2 The Controllers	98		6.1.1 Latin Hypercube	136
	4.2.1 Mean Effective Pressure	98		6.1.2 Settings of the DoE	137
	4.2.2 Variable Geometry Turbocharger	98		6.2 Training of the Neural Network	144
	4.2.3 High Pressure EGR	100		6.3 From FR to MV Model	148
	4.3 Fast-Running engine model	101		6.4 Comparison between DETM, FRM and MVM	152
	4.4 Comparison between DETM vs FRM	106		6.5 Integration of MVM in HiL testing	163
5	Design of Control Algorithm	119	7	The Real-Time Model	169
	5.1 Control Theory	120		7.1 Development of the RT Model	170
	5.1.1 Open-loop control	121		7.2 Comparison between DETM, FR, MV and RT Models	171
	5.1.2 Closed-loop control	122			
	5.2 Combustion Process control	122	8	Conclusion	179
	5.2.1 Pressure-based control	123			
	5.2.2 Model-based control	124			
	5.2.3 Map-based control	126			
	5.3 Control testing procedures	129			
	5.3.1 Model in the Loop	129			
	5.3.2 Hardware in the Loop	130			
	5.3.3 Rapid Prototyping	130			

CON
TEN
TS

FIG
UR
ES

110	Intake Manifold Pressure, DETM vs FRM - Figure 4.31
110	Intake Manifold Temperature, DETM vs FRM - Figure 4.32
111	Exhaust Manifold Pressure, DETM vs FRM - Figure 4.33
111	Exhaust Manifold Temperature, DETM vs FRM - Figure 4.34
112	Turbine Average Speed, DETM vs FRM - Figure 4.35
112	Turbine Outlet Temperature, DETM vs FRM - Figure 4.36
113	VGT Controller (Boost Actuator), DETM vs FRM - Figure 4.37
113	Compressor Outlet Pressure, DETM vs FRM - Figure 4.38
114	Intercooler Outlet Pressure, DETM vs FRM - Figure 4.39
114	Intercooler Outlet Temperature, DETM vs FRM - Figure 4.40
115	EGR Rate, DETM vs FRM - Figure 4.41
115	EGR Valve Mass Flow Rate, DETM vs FRM - Figure 4.42
116	EGR Controller; DETM vs FRM - Figure 4.43
116	EGR Cooler Outlet Temperature, DETM vs FRM - Figure 4.44
120	Scheme of a controller principle - Figure 5.1
121	Scheme of an open-loop controller - Figure 5.2
121	Scheme of a feed-forward controller - Figure 5.3
122	Scheme of a closed-loop controller - Figure 5.4
123	SOI correction for pressure-based MFB50 control - Figure 5.5
124	SOI correction for model-based MFB50 control - Figure 5.6
130	Hardware in the Loop set-up - Figure 5.7
131	Rapid Prototyping set-up for pressure-based controller - Figure 5.8
131	Rapid Prototyping set-up for model-based controller - Figure 5.9
134	Scheme of the variable in a generic process - Figure 6.1

159	VGT Controller (Boost Actuator), DETM vs FRM vs MVM - Figure 6.28
159	Compressor Outlet Pressure, DETM vs FRM vs MVM - Figure 6.29
160	Intercooler Outlet Pressure, DETM vs FRM vs MVM - Figure 6.30
160	Intercooler Outlet Temperature, DETM vs FRM vs MVM - Figure 6.31
161	EGR Rate, DETM vs FRM vs MVM - Figure 6.32
161	EGR Valve Mass Flow Rate, DETM vs FRM vs MVM - Figure 6.33
162	EGR Controller, DETM vs FRM vs MVM - Figure 6.34
162	EGR Cooler Outlet Temperature, DETM vs FRM vs MVM - Figure 6.35
163	NI-PXIe-1082 from National Instruments - Figure 6.36
164	Veristand execution mechanism - Figure 6.37
165	ETAS ES910 Rapid Prototyping Device - Figure 6.38
166	Hardware connections in the HiL testing procedure - Figure 6.39
170	Real-Time Model conversion parts - Figure 7.1
172	Factor of Real Time, DETM vs FRM vs MVM vs RTM - Figure 7.2
172	Injected Average Mass Flow Rate, DETM vs FRM vs MVM vs RTM - Figure 7.3
173	Brake Torque, DETM vs FRM vs MVM vs RTM - Figure 7.4
173	BMEP, DETM vs FRM vs MVM vs RTM - Figure 7.5
174	Intake Manifold Pressure, DETM vs FRM vs MVM vs RTM - Figure 7.6
174	Intake Manifold Temperature, DETM vs FRM vs MVM vs RTM - Figure 7.7
175	Exhaust Manifold Pressure, DETM vs FRM vs MVM vs RTM - Figure 7.8
175	Exhaust Manifold Temperature, DETM vs FRM vs MVM vs RTM - Figure 7.9
176	Intercooler Out Pressure, DETM vs FRM vs MVM vs RTM - Figure 7.10
176	Intercooler Out Temperature, DETM vs FRM vs MVM vs RTM - Figure 7.11
177	VGT Controller, DETM vs FRM vs MVM vs RTM - Figure 7.12
177	EGR Actuator, DETM vs FRM vs MVM vs RTM - Figure 7.13

TA
BL
ES



1

COMPRESSION IGNITION ENGINES

Introduction

Standard Combustion

J.E. Dec Model

Non-standard Combustion

Pollutant Emission

Injection Technologies

1.1 Introduction

The Compression Ignition (CI) engines are commonly named *Diesel engines* since they refer to the Diesel thermodynamic cycle. Also, CI engines differ from Spark Ignition (SI) engines because of the injection and the lack of the ignition system.

Despite CI engines have no ignition system, they undergo to loads which are much higher compared to the loads of SI engines, causing heavier components to be mounted. The higher weight of the parts forming the crank drive, and the criticalities of the combustion process (injection and combustion timing), obstacle the achievement of high rotational speed of the crank shaft (around 5000 rpm). Specific powers (CV/liter) of CI engines are lower compared to that of SI engines.

Indeed, SI engines admit *Gasoline Direct Injection* (GDI) and *Port Fuel Injection* (PFI), that is, the injection occurs in the intake runner just ahead the *Top Dead Center* (TDC). This is possible since the gasoline is poorly reactive, whereas the Diesel is highly reactive, meaning that CI engines do not require an ignition system.

The fuel of CI engines are mainly Diesel and Biodiesel which are characterized by long and flexible hydrocarbons chains. These characteristics permits to avoid an external energy supplier to trigger the combustion process (as for the spark in SI engines).

Because of the high reactivity of the Diesel fuel, the combustion occurs differently since it is not required to pre-mix air and fuel (otherwise the mixture would self-ignite spontaneously), but rather the fuel is injected in the form of liquid particles directly in the combustion chamber at

the end of the compression phase. Therefore, the engine only intake fresh air and gases from the Exhaust Gas Recirculating (EGR) system.

The fuel is then injected around the TDC through high pressure injectors (around 1000 bar) at speed in the order of 100 m/s. The fuel particles clash against high-density air molecules. The aerodynamic interactions between the fuel droplets and the air cause the atomization of the fuel jet. Here the fuel droplets reach dimensions which are in the order of *nanometers*. The droplets of atomized fuel are surrounded by compressed hot air (around 900 K) that trigger the spontaneous combustion of the fuel without the need of an external apparatus.

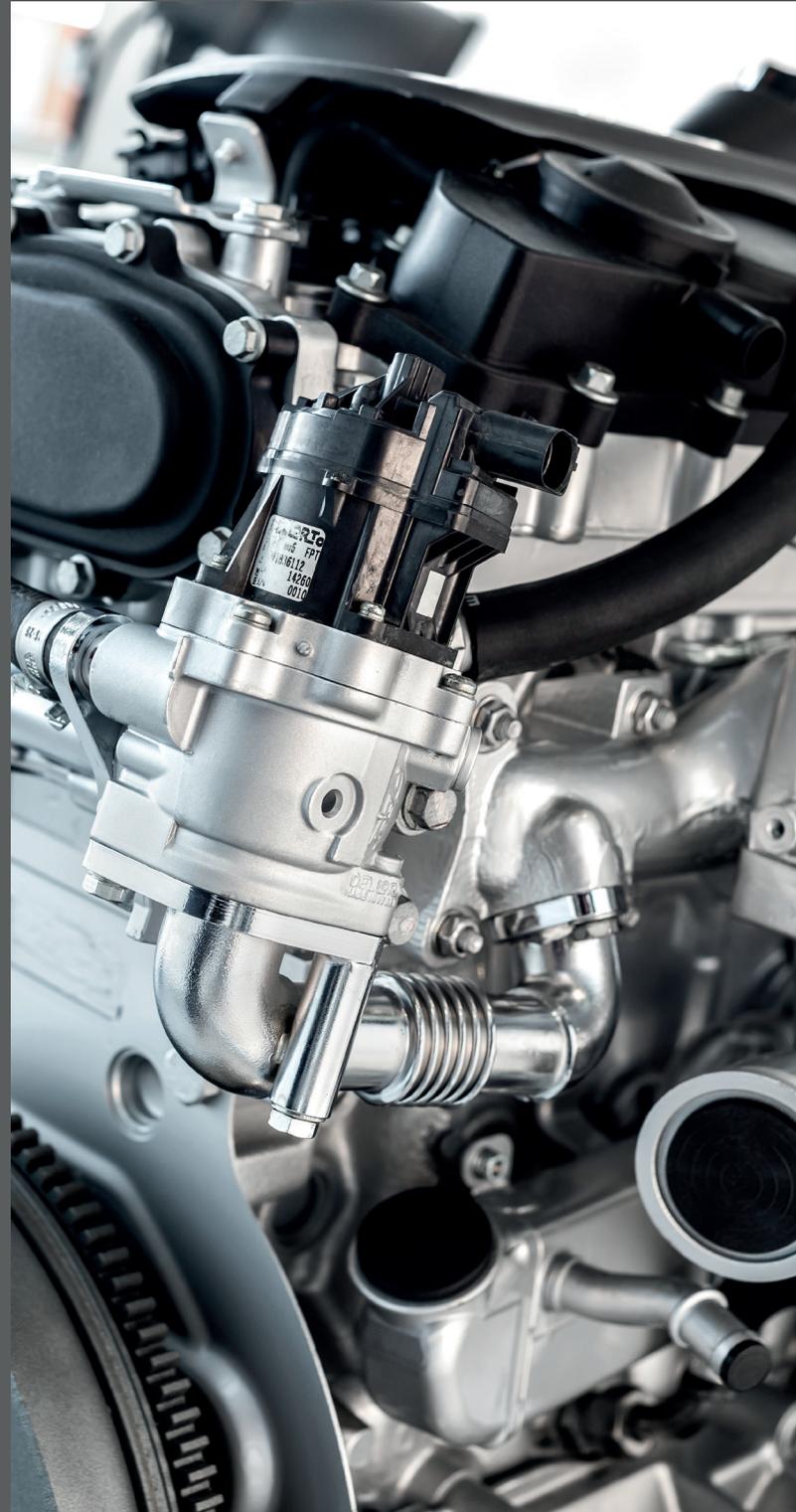
Since the evaporation time of the droplets is directly proportional to the nominal diameter, it is then required to have droplets as small as possible that evaporate as quick as possible because the fuel is highly reactive by itself, meaning that the time slot available for the combustion is relatively short. To balance this characteristic the fuel, the injection advance is reduced, however, the reduction of droplets diameters refers to the injection pressure that is ultimately the cost of the system.

In CI engines, the fuel injection directly in the combustion chamber determine an *non-homogeneous* mixture since the value of the equivalence ratio ϕ differs depending in the considered region of the mixture. Actually, the combustion process occurs in the regions where the air/fuel ratio is stoichiometric, but due to the high reactivity of the fuel, the process also occurs in the further zones. Still, the region which are rich in fuel are highly responsible for the generation of pollutant carbonous molecules known as *soot*.

Because of the nature of the combustion in CI engines, the process is not capable of adapting to the engine speed variations. However, the control of the engine load is easier since it only refers to the quantity of injected fuel. This allows to avoid a throttling system (butterfly valve) that worsens the overall efficiency of the engine.

In addition, the combustion process in CI engines avoids any kind of abnormal conditions that constraint the geometry of the engine. This permits to obtain engines with very large bore that are suitable in marine applications. Also, the combustion in CI engines does not affect limitations over the Compression Ratio (CR). Hence, a higher CR can be used in the CI engine, improving its fuel conversion efficiency relative to SI engines. Moreover, since injection timing is used to control combustion timing, the delay period between the start of injection and start of combustion must be kept short. Thus, the spontaneous ignition characteristics of the fuel-air mixture must be held within a specific range. The high compression ratios exploited in Diesel engines force to use of robust piston that generate high inertial forces (frictions). Then CI engines undergo higher pressures and temperatures resulting in higher weights, relative to SI engines.

Another peculiar feature of CI engines is the noise that is mainly audible in low load and low temperature conditions. Combustion is noisy because the delay in between the fuel injection and the start of combustion (*ignition delay*) leads to a fuel accumulation that simultaneously burns. The abrupt pressure raise into the combustion chamber during the premixed burning phase produces stresses throughout the engine structure that are translated into vibrations which in turn irradiates noise.



1.2 Standard Combustion

The fundamental difference between the spark-ignition (SI) engine and the compression-ignition (CI) engine relies on the type of fuel involved in the combustion process. Indeed, the combustion does not develop at constant volume, as for Otto cycle, nor at constant pressure, as for Diesel cycle.

An operating cycle of an engine can be represented by plotting the evolution of the in-cylinder pressure (p) against its volume (V) during the cycle. A typical p - V diagram of a four-stroke CI engine is shown in Figure 1.1.

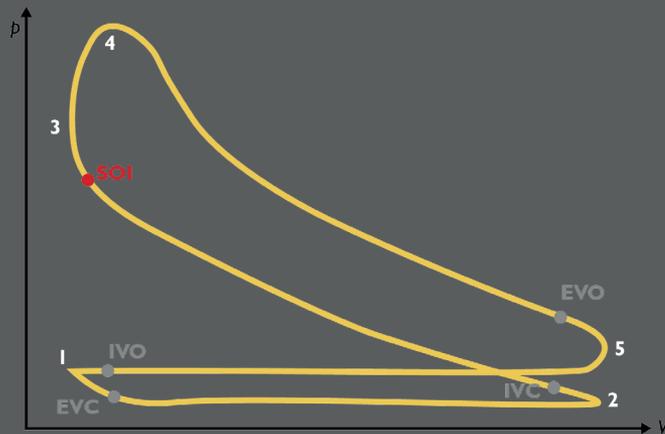


Figure 1.1 p - V chart of a CI engine

Fuel is injected by the fuel-injection system into the engine cylinder toward the end of the compression stroke, just before the desired Start of Combustion (SOI).

The liquid fuel, usually injected at high velocity as one or more jets through small orifices or nozzles in the injector tip, atomizes into small drops and penetrates into the combustion chamber. The fuel vaporizes and mixes with the high-temperature, high-pressure cylinder air. Since both the pressure and temperature of the cylinder contents at the

time of injection are very high, some chemical reaction begins as soon as the first droplet of the injected fuel enters the cylinder. As a consequence of the temperature value, (above the fuel ignition point), spontaneous ignition of portions of the already-mixed fuel and air occurs after a delay period of a few crank angle degrees. The cylinder pressure increases as combustion of fuel-air mixture occurs. However, this chemical reaction starts so slowly that the visible flame occurs only after a period of time called *delay period*. Combustion in CI engine is a matter of local condition of the charge and only assisted, not dependent, on the spread of the flame.

The CI combustion can be divided into IV main phases:

- I. Ignition Delay
- II. Premixed Combustion
- III. Mixing Controlled (Diffusive) Combustion
- IV. Late Combustion

The combustion in CI engines is analyzed through the AHRR - Apparent Heat Release Rate [J/deg] - which is the energy index per unit time expressed as function of the crank angle. Thus, the AHRR is the rate of release of the heat during combustion. Actually the combustion releases chemical energy, not heat, but the effect of such energy release, is an increase of temperature and pressure in the combustion chamber. For this reason, the term *apparent* refers to a sort of an equivalent heat even though the combustion is not properly a heat transfer process. The Apparent Heat Release Analysis is the technique to study combustion according to two main contributions:

$$AHRR = \frac{1}{k-1} V dp + \frac{k}{k-1} p dV, \quad k = \frac{c_p}{c_v}$$

where V is the volume of the combustion chamber and p represents the in-cylinder pressure measured through pressure transducers placed into the glow plug. Since we are referring to the apparent rate, the heat transfer towards the cylinder wall, head and heat of evaporation are neglected. The goal of AHRR analysis is to understand how this chemical energy is released during the combustion.

The graph deals with the time evolution of the apparent heat release, expressed as the degrees of crank angle after start of injection (ASI) based on the AHRR mathematical model.

At the very beginning, when the fuel evaporates, such evaporation of the liquid fuel determines a cooling of the charge represented as a negative heat release. The combustion actually starts when the heat release rate increases without any changes in its trend.

Through the graph of Figure 1.2 [3], it is possible to locate the main phases of the combustion evolution. The origin of the graph is placed at start of injection (SOI), which however does not coincide with Top Dead Center, but takes place few degrees in advance.

Combustion in Diesel engine occurs in a form of a plume that is starting from the injector tip. In general, the first part of this plume is made up of liquid which starts penetrating the combustion chamber while atomization is occurring.

This atomization determines the mixture between fuel and hot air. Autoignition of the rich premixed charge occurs for the suitable value of the equivalence ratio ϕ .

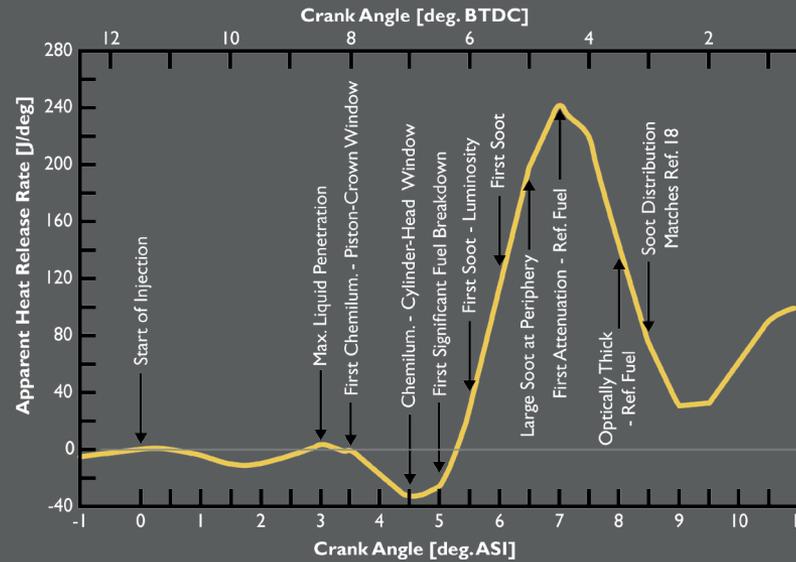


Figure 1.2 - Apparent Heat Release Analysis of a CI engine

I. Ignition Delay

Referring to Figure 1.2 [3], the injection occurs at *Start of Injection* (SOI). Its importance is mainly due to its influence on the subsequent rapid combustion, whose pressure increase rate is related to the mass of premixed air/fuel-vapor charge formed during the delay.

First of all, the definition of the delay period is based on the experimental techniques available. The start of the delay is usually taken as the time when the injection begins, measured by the lift of the injector needle. The end, instead, marked by the Start of Combustion (SOC), can be determined in the following ways:

- I. detection on the cylinder pressure diagram of the moment, when the pressure increase due to combustion divide the pressure curve from that without injection of a fixed percentage (typically of 1%);

2. determination of the first appearance of the flame inside the combustion chamber by means of a flame luminosity detector (for example a *photodiode*);
3. measure of the change in composition of the cylinder charge, produced by the combustion of the first fuel element.

During the ignition delay both physical and chemical processes take place. The physical actions change the aggregation state of the fuel molecules, while the chemical actions modify the fluid nature. Schematically, the physical processes are:

1. the *atomization* of the liquid fuel jet, forming smaller and smaller droplets, carried away by the air in turbulent motion,
2. the *vaporization* of the liquid droplets because of the heat received by the surrounding hot air;
3. the *diffusion* and mixing of the fuel vapors with air, to form a fuel-air mixture able to auto ignite.

These phenomena are mainly influenced by the droplet size distribution in the fuel spray (depending on: injection pressure, injector hole diameter and geometry, fuel viscosity and volatility, etc...), the air turbulent motion (depending on: combustion chamber and manifold geometry, engine speed, etc.) and the temperature and pressure of the cylinder air (depending on: compression ratio, supercharging ratio, cooling system characteristics, etc...).

The chemical processes can be summarized as

1. thermal *cracking* of large hydrocarbon molecules in elements of smaller molecular weight;

2. *partial oxidation* reactions with the formation of unstable intermediate oxygenated products (peroxides, aldehydes, ketones, etc...);
3. start of *chain reactions* able to bring to the auto-ignition the fuel/air mixture.

These chemical processes depend on dimension and structure of the fuel molecule, which cause it to react more or less easily with the oxygen, as well as the cylinder charge temperature and pressure and the physical processes that control the mixing of the fuel with the air.

However, the chemical component of the delay is mainly influenced by the ignition characteristics of the fuel, summarized by its cetane number, which is determined by comparing the ignition delay of the fuel with that of primary reference fuel mixtures in a standardized engine test.

It is worth mentioning that the physical and chemical processes, listed above, occur in at the same time and in an interconnected way. Therefore, only from a schematic point of view it is useful to distinguish the physical from the chemical aspects of the phenomena.

The physical component of the process takes on a major relevance compared to the chemical because Diesel engines work with high reactive fuels. This actually means that chemical reactions occur very rapidly, whereas the atomization, vaporization and mixing of fuel molecules with air molecules requires longer time. However, the chemical reactions surely start only after the mixing of the fuel vapors with the air. Therefore the first part of the ignition delay is mainly controlled by the physical processes, which produce the combustible mixture, while the second part is mainly influenced by the chemical reactions that bring

to its autoignition. Also, during the ignition delay it is possible to note the *accumulation* phenomenon. This because high-pressure fuel particles are injected into the combustion chamber before SOC. In the AHRR graph of Figure 1.2^[3], the negative value of the Heat Release located before the breakdown is caused by the fuel evaporation that subtracts energy from the environment to trigger the chemical process. These fuel accumulation will burn in the successive phase causing the typical abrupt change in pressure within the chamber.

II. Premixed Combustion

Once both physical and chemical delays are exhausted, the first fuel fractions start to release chemical energy through oxidation reactions.

This is translated as a rapid change in pressure inside the combustion chamber, as shown in Figure 1.3^[5]. The moment in which the fuel starts to release chemical energy is defined *autoignition*.

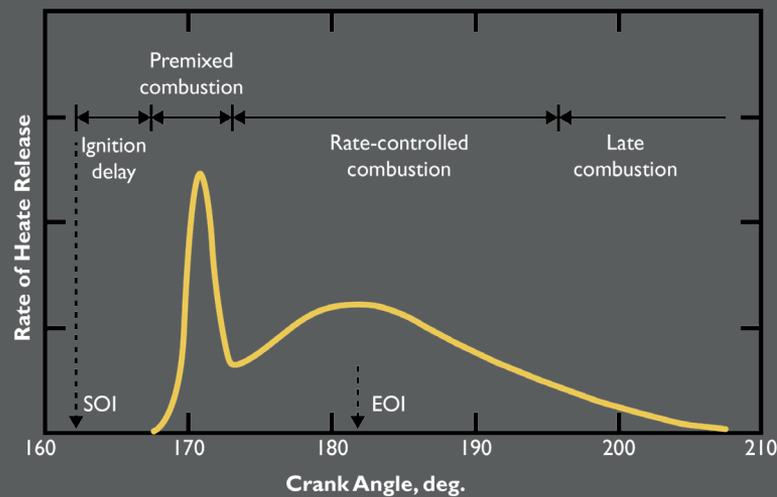


Figure 1.3 - Combustion phases in the AHRR

Thus, the simultaneous burn of the fuel stresses all the engine components inducing vibrations and noise.

Also, this phase is the main responsible for the generation of nitrogen oxide NO_x emissions. Indeed, the fuel tends to burn in a poor-oxygen environment.

This condition leads to an increase of pressure and temperature which are favorable parameters for the successive formation of NO_x in the following phase, that is, the mixing-controlled combustion. The NO_x formation is enhanced by high temperature ($T > 1850$ K). Thus, during the premixed combustion, the NO_x formation is broken by a poor oxygen availability despite favorable high temperatures.

As a consequence, one obtains two contradictory effects:

- it is convenient to *reduce* the pressure peak in the HRR to dampen down engine noise and the successive formation of NO_x ,
- it is favorable to keep the pressure peak at the highest possible level in order to improve the thermo-fluid dynamic efficiency.

The premixed combustion is also named *rapid combustion* or *uncontrolled combustion* because the fuel accumulation leads to the rapid pressure variation. The rate of pressure rise directly depends on fuel quantity at the end of the ignition delay, on the turbulence effect and on the jet atomization. Once this fuel premixed with air before ignition has burned, the combustion process slows down. The rate at which the energy release proceeds is determined by liquid fuel atomization, vaporization, oxidation reactions, but mainly by the mixing of fuel vapor with air.

III. Mixing Controlled Combustion

During the rapid combustion of the fuel accumulated during the ignition delay, only a fraction of the fuel is burnt (around 10%). Hereinafter, the mixed controlled combustion is starting. In this phase, the rate of combustion depends on the speed at which the fuel, that is continuously injected, evaporates spreading among air molecules.

Furthermore, this phase can be monitored by controlling the injection rate which is the angular interval of the hydraulic duration of injection. As confirmed by Figure 1.3^[3], the rate of heat release experiences a peak of lower intensity, compared to the peak during the rapid combustion. In the mixing controlled combustion, the temperature reaches the highest value around 3000 K. Here, the fuel must oxidize as much as possible despite the increasing of exhaust gases. Indeed, the unoxidized fuel originates soot and unburned hydrocarbons (HC), such as polyaromatic hydrocarbons (PAH). To enhance the mixture, swirl motion within the combustion chamber becomes crucial. Indeed, the rotational motion of the fluid promotes the fuel-air mixture.

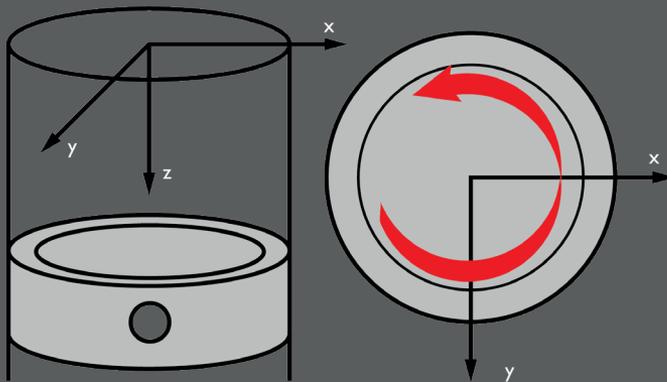


Figure 1.4 - Swirl motion in the combustion chamber

Thus swirl motion is needed to balance the high reactivity of the diesel fuel. Its ability to burn in rich environments allows hydrocarbons molecules to pursue de-hydrogenation processes, passing from a H/C ratio from 1.85 to 0.1. Subsequent to this behavior, it results that carbonaceous chains remain unoxidized since the hydrogen reacted with few oxygen molecules.

Such carbonaceous chains tend to accumulate up to form solid particles with millions of carbon atoms. These agglomerations are conveyed by the exhausted gas generating soot at engine exhaust.

However, soot may continue to oxidize as soon new oxygen is available along the pipes provided the temperature is sufficiently high.

IV. Late Combustion

The late combustion represents the final phase of the combustion process. This phase is characterized by the stop of the injection despite exhaust valve are still closed. Indeed, fuel particles may still burn.

Chemical reactions start diminishing gradually as time goes by. Also the heat release trend is dampen down till reaching the end of combustion. This phase allows to oxidize soot particles that have not found oxygen molecules.

Useful work is low since the piston has almost completed the expansion stroke. For this reason, it is crucial that the late combustion should not take long time, otherwise the engine efficiency is worsen. To increase the efficiency it is required to focus the heat release in the round of the Top Dead Center (TDC).

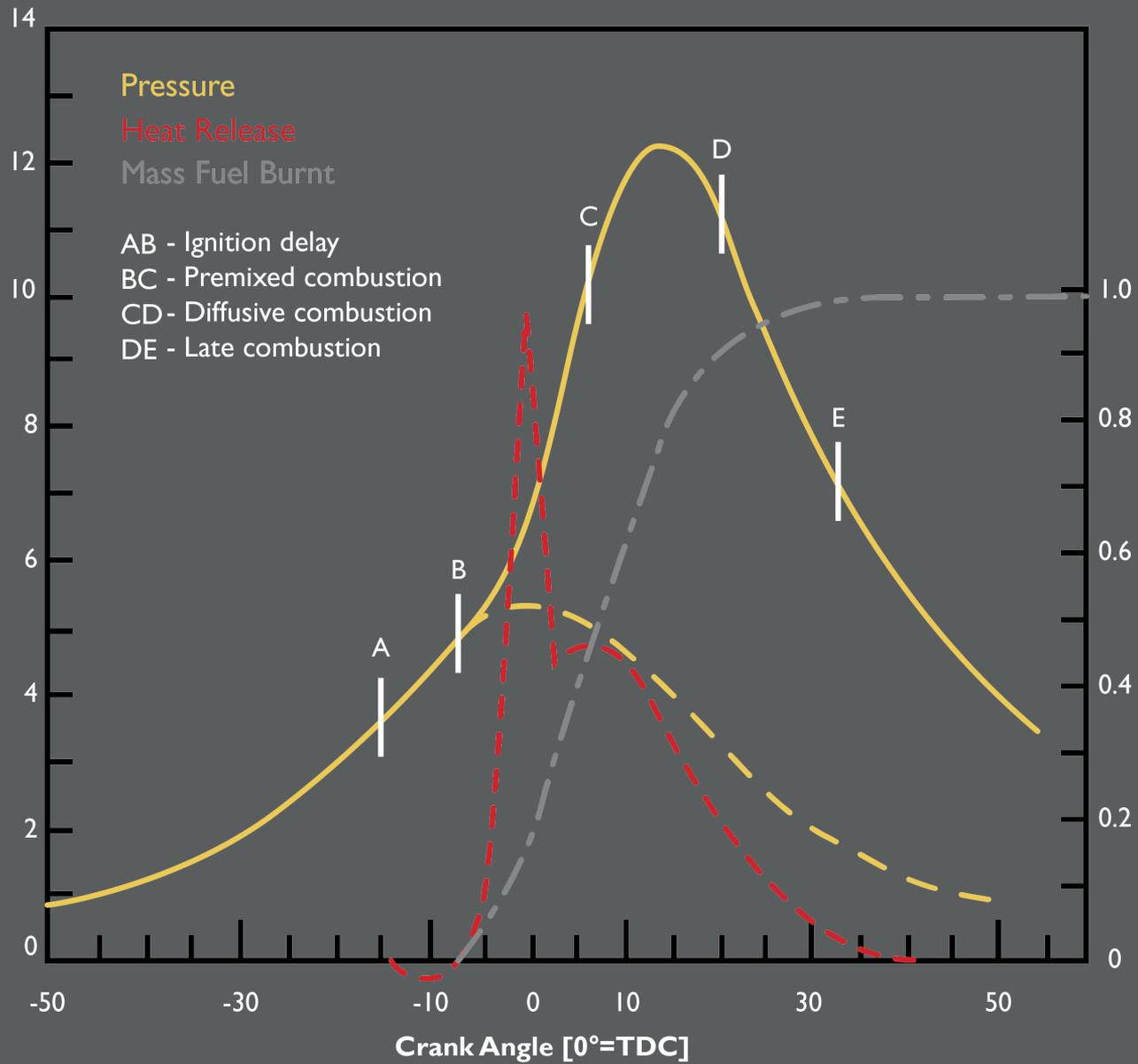


Figure 1.5 - Combustion phases in CI engines ^[3]

1.3 J.E. Dec Model

The development of advanced laser-based diagnostics has provided a means for making detailed *in-situ* measurements of the processes occurring inside of a reacting diesel fuel jet. This more complete understanding led to a conceptual model of diesel combustion that differs significantly from the old description and offers new insight into the controlling physics of a combusting diesel fuel jet. The experimentations on an optically accessible engine greatly contributed to have a better understanding of the phenomena. The optical-access engine is represented in Figure 1.6 [6]. Basically, it is a modified engine, in which, instead of the piston top, a mirror is introduced, in order to have an optical access of the combustion chamber. In addition, an exhaust valve has been removed and replaced by another window. A laser sheet is sent to highlight the spray and pictures are taken to understand how the fuel spray is entering, penetrating and evaporating.

Figure 1.7 [7] shows that before the application of laser-sheet imaging, it was assumed that the diesel combustion process had a fuel spray with a fuel rich liquid core and a fuel distribution that dropped off in a Gaussian-like manner with increasing radius, with three important characteristics:

- the liquid phase penetrated well out from the injector with fuel droplets being present up to or within the combustion zone,
- after the premixed burn, combustion occurred solely in a diffusion flame and was confined to the peripheral region of the jet,
- soot formation occurred mainly in the shell-like region around the jet periphery.

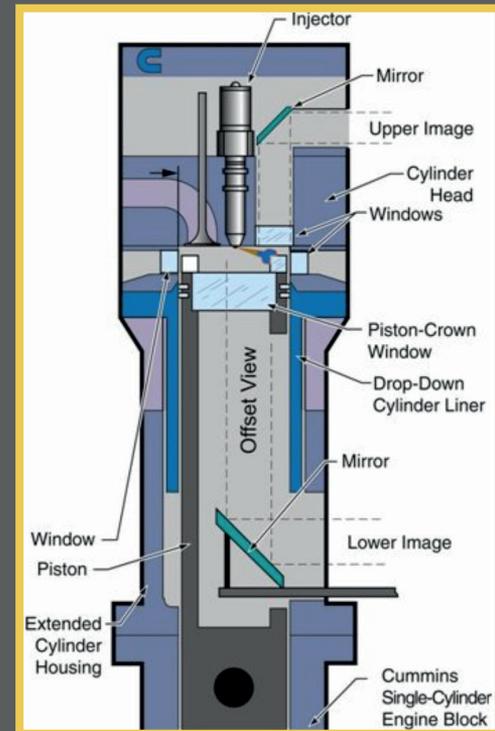


Figure 1.6 - Cummins N-14 Optical-Access Engine

Engine base type.....	Cummins N-14, DI Diesel
Number of cylinder.....	1
Cycle.....	.4-stroke
Number of intake valves.....	2
Number of exhaust valves.....	1
Combustion chamber.....	Quiescent, direct injection
Bore.....	139.7 mm (5.5 in)
Stroke.....	152.4 mm (6.0 in)
Chamber diameter.....	.97.8 mm (3.85 in)
Displacement.....	2.34 liters (142 in ³)
Connecting rod length.....	304.8 mm (12.0 in)
Piston pin offset.....	None
Compression ratio.....	11.2 : 1

Table 1.1 - Specification of the Optical-Access Engine

Autoignition and the premixed burn were thought to occur in regions that were nearly stoichiometric, primarily around the jet periphery (concentration of fuel close to the stoichiometric value).

In addition, ignition was thought to occur at only a few points followed by a rapid spread of the flame around the jet periphery. The concept idea was to have zones of varying fuel-air mixture from center to edge of the spray.

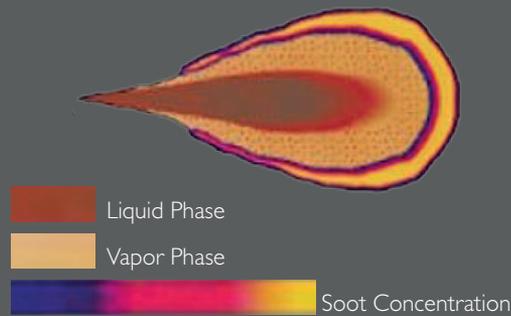


Figure 1.7 - Old model combustion plume

The spray-flame theory from which the original diesel combustion description was derived dealing with fuel vaporization, mixing, and combustion zones, and it was not specific as to the location of the soot formation.

Above all, this model was not able to represent soot formation, which is occurring when an hydrocarbon based fuel is burnt (*not* heated-up) under extremely rich conditions.

Since soot formation results from fuel pyrolysis at temperatures above about 1300 K, mixing with the hot in-cylinder air (900-1000 K) is not sufficient to induce soot formation. Consequently, it is required to extract energy from the heat coming from the combustion process. Therefore,

during the quasi-steady portion of Diesel combustion, it was generally assumed that soot would form on the fuel-rich side of the diffusion flame, where the temperatures were sufficiently high.

The initial premixed burn was not considered to be an important source of soot production since it was thought to occur in regions that were nearly stoichiometric, primarily around the jet periphery.

In the last 20 year, the development of the new model was made possible through new knowledge and equipment, such as the optical-access engine previously presented.

The new theoretical model, known as **J.E. Dec model**, is based on experimental studies performed on the engine optically accessible through laser techniques. This has allowed to visualize and analyze the liquid and vapor fuel jet, the air/vapor mixture and to better understand the soot formation mechanism.

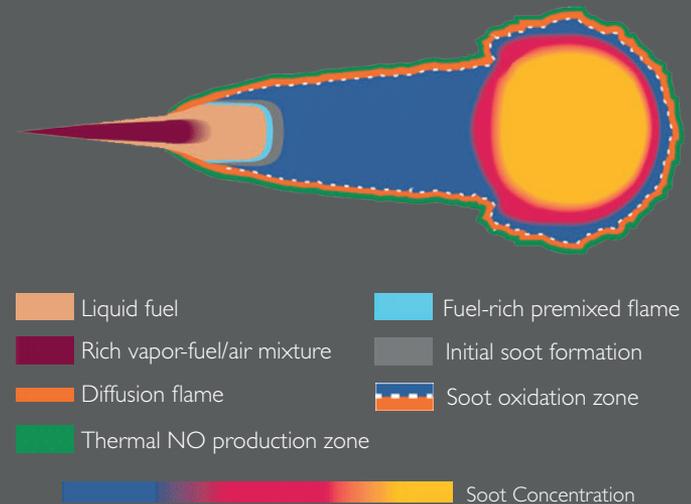


Figure 1.8 - J.E. Dec combustion model ^[7]

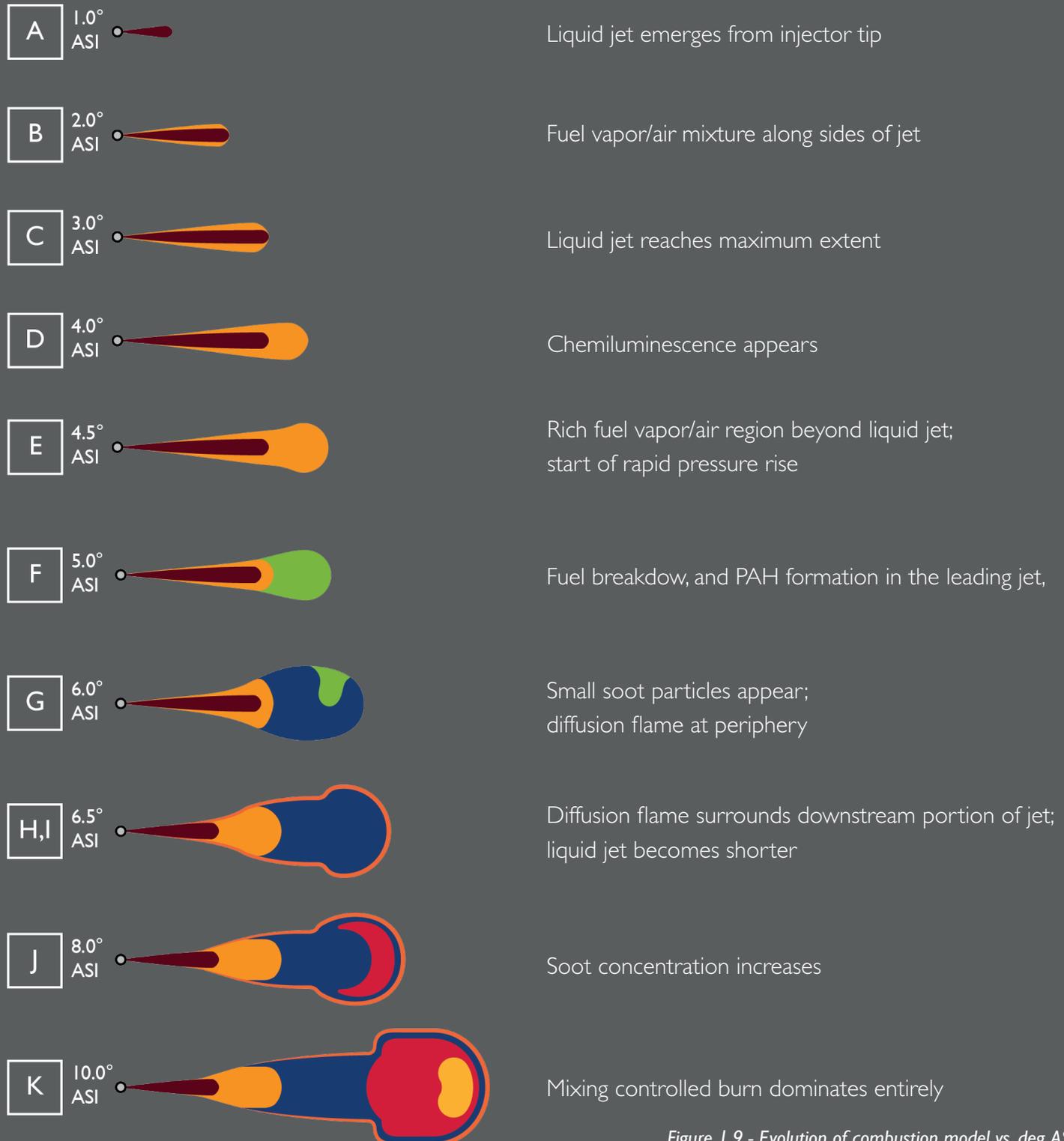


Figure 1.9 - Evolution of combustion model vs. deg.ASI

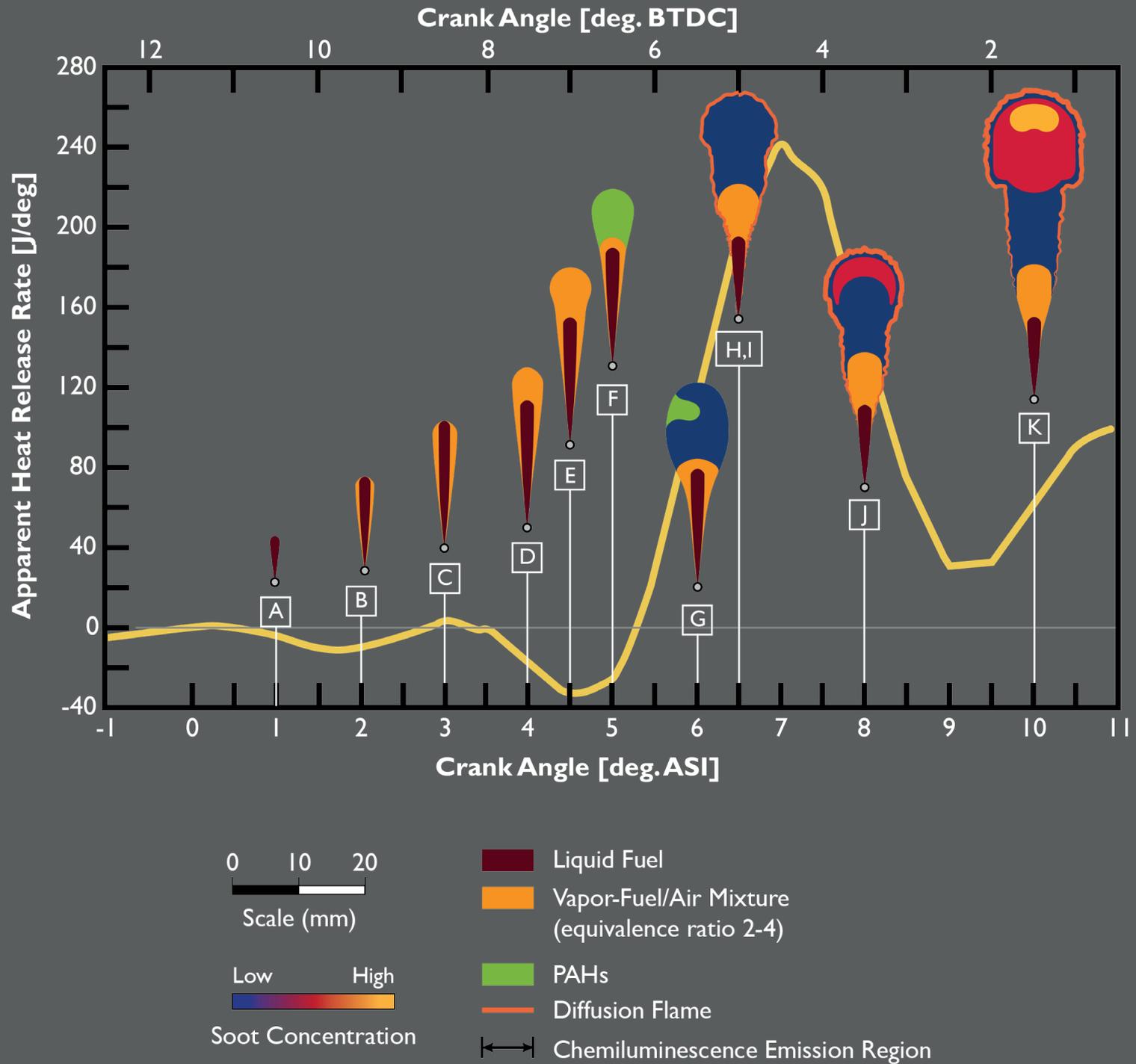


Figure 1.10 - Evolution of combustion model vs. AHRR

Figure 1.9^[3] and 1.10 describe the evolution of the J.E. Dec model right after the start of injection (Figure 1.9) and related to the Apparent Heat Release Rate (Figure 1.10).

Since the Diesel engine is provided of direct injection, then the nozzle of the injector is physically inside the combustion chamber. For this reason it is reasonable to distinguish the start of injection (SOI), which is a command, and the instants after the injection (ASI), that is, the moments in which the needle of the injector lifts up along its stroke.

Still referring to Figure 1.9^[3], consider now a qualitative trend of the combustion plume.

Phase A - 1° ASI



One degree of delay after the fuel injection, it is possible to observe a loose liquid jet made of small fuel droplets.

Phase B - 2° ASI



Starting from two degrees after start of injection, the first air molecules are visible.

Phase C - 3° ASI



Compressed hot air droplets, at a temperature around 600 K, start mixing with fuel droplets. Because of this high temperature, the fuel droplets start evaporating, thus creating a mixture of air and fuel vapor around the liquid jet. The overall plume composed of liquid and vapor parts is named *spray*.

Since the nozzle is still injecting, then the liquid part of the jet is increasing. Such droplets are produced by the fuel atomization due to the pressure drop downstream the nozzle. The average droplet diameter is much lower compared to the nozzle holes that are in the order of 100 μm .

The maximum extent of the liquid jet, defined as the maximum distance from the nozzle tip to the tip of the liquid part is named *penetration of the liquid length*.

The higher the temperature in the combustion chamber, the higher the evaporation effect, the lower the maximum penetration of the liquid length.

Phase D - 4° ASI



In this phase it is possible to observe a slight brightness related to the energy released by radicals, which are generated by the exothermic reactions between fuel and oxygen.

This process is defined as *chemiluminescence*, since these radicals emit visible photons.

The air and fuel vapor mixture becomes bigger and takes place ahead the liquid part. The continuous evaporation of fuel vapor triggers the premixed combustion. Such behavior is verified in Figure 1.10 by the negative contribution of the AHRR, indeed the evaporation process subtracts energy (heat) from the environment. As a result, the evaporation reduces the temperature making the heat release rate *apparently* negative.

The successive rise in pressure is due to chemiluminescence process which is directly related to a higher rate of chemical energy release.

Phase E - 4.5° ASI



The mixture of air and fuel vapor is in rich condition. The equivalence ratio ϕ ranges between 2 and 6, meaning that the fuel component in the mixture is up to six times higher than the optimal value. In Diesel engines, this rich mixture faces an environment which is plenty of oxygen. Indeed, the combustion chamber is a non-homogeneous environment because, locally, one has a highly rich mixture, but the surrounding is full of oxygen.

Since the mixture is highly non-uniform, the analysis of the average properties (temperature, pressure, air-fuel ratio) may not be suitable to describe pollutant formation mechanism.

Autoignition causes the so-called pre-mixed burning, which is a combustion relatively cold (1600 K). In the pre-mixed burning, the energy release from the autoignition of the pre-mixed charge suddenly raises up the temperature which in turn, propagates the combustion to a larger extent causing the rapid heat release. The main products of the pre-mixed burning are:

- CO carbon monoxide,
- HC fuel fragments,
- C_2H_2 , C_2H_4 , C_3H_3 Soot precursors (PAH).

Also CO_2 , and H_2O are present in a minor extent since the high equivalence ratio ($\phi = 2 \div 6$) does not allow the fully oxidation of the hydrocarbons. The high equivalence ratio is due the low time required to mix air with fuel.

Phase F - 5° ASI

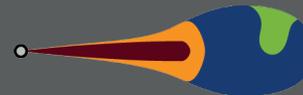


Phase F is the very early stage of the combustion. The fuel vapors start to chemically react with the surrounded oxygen at a slower rate compared to that of the combustion. The product of these preliminary reactions is composed of radicals (Polycyclic Aromatic Hydrocarbons - PAH) producing a very small rate of chemical energy release-

Once these radicals are created and sufficiently in high number (green color), the probability that the fuel itself reacts with the radicals is very high and when this occurs, more radicals are produced.

Thus, the combustion is made up of several reactions between fuel and radicals occurring at the same time. This primary combustion, which is the first stage of a Diesel engine, is defined pre-mixed combustion.

Phase G - 6° ASI



Starting from residual PAHs, the region which are at low temperature (around 1600 K), as well as in absence of oxygen, promote the formation of soot.

These carbonaceous molecules are responsible for the formation of particulate matter.

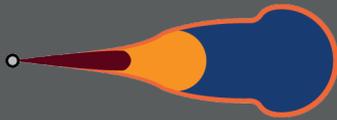
Because the formation of soot is once again an exothermic reaction, then the interfaced region between soot and oxygen all around the plume is named *diffusion flame*.



Diesel combustion is a complex, turbulent, three-dimensional, multiphase process that occurs in a high-temperature and high-pressure environment. As a result, prior to the relatively recent advent of advanced laser diagnostics, detailed measurements of the events occurring within a reacting Diesel fuel jet were not possible.

- SAE 970873

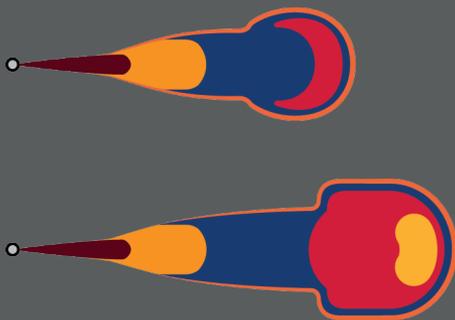
Phase H,I - 6.5° ASI



The liquid part of the jet tends to reduce whereas the internal part keep growing because the nozzle is still injecting. When the reactants become products, the temperature increases promoting further evaporation.

Also, the spray is surrounded by the diffusive flame which introduces the diffusive combustion. As seen, the premixed combustion occurs at high equivalence ratio, meaning that the fuel is still available for oxidation in the second combustion (i.e. diffusive). The products of the rich, premixed combustion lay within the jet, while the fresh oxygen is surrounding the plume. The diffusion flame is resultant of a local diffusive motion that interface the reactants with the oxygen.

Phase J,K - 8°-10° ASI



The jet main structure remains unchanged, still, it keeps growing by increasing the soot concentration. It is worth noting that the highest soot concentration is located ahead the jet because in the diffusion flame the tempe-

perature is higher (around 2700 K). The final composition of the combustion plume is composed of a small liquid part, a mixture rich of air and fuel vapors, residuals from premixed combustion (blue color) and finally the diffusive flame surrounds all the plume.

Through the Dec model, it was possible to understand that carbonaceous molecules are generated within the plume during premixed combustion, rather than around the diffusion flame. However, nitrogen oxides (NO_x) are formed in the diffusion flame, where temperatures are higher and oxygen is fully available. Finally, unburned hydrocarbons (HC) are promoted in the final part of the plume due to over-mixture phenomenon, as well as in locally-rich regions due to under-mixture phenomenon.

1.3.1 Parameters to describe Dec Model

Considering the combustion plume structure till the end of injection reported in Figure 1.11^[8], the combustion occurs in two main steps: premixed burning after auto-ignition and diffusive burning which carries most of the chemical energy. The main parameters needed to describe the plume structure are:

- L penetration of liquid fuel;
- S penetration of the spray;
- H lift-off length.

The liquid fuel penetration L is the maximum distance at which liquid droplets are measured with respect to the origin of the injector tip.

Also, the spray penetration S is defined as the last position, always measured compared to the injector tip, where fuel vapor traces are still appreciated.

Concerning the lift-off length H , it is the distance from the injector tip to the start of the diffusion flame. The longer the lift off length H , the more the opportunities for air entrainment after autoignition.

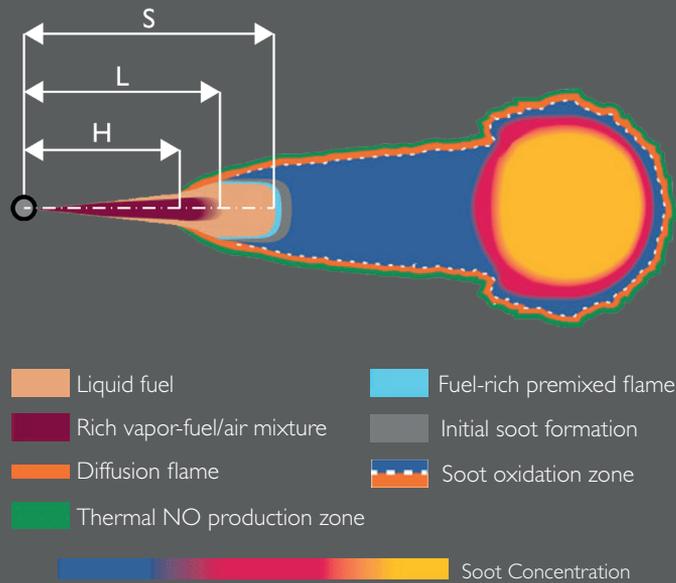


Figure 1.11 - Parameters of the Diesel combustion plume

1.3.2 Diffusion flame development

Pictures in Figure 1.13^[9] represent the evolution of the diffusion flame, identified by the natural luminous emission of the soot, as function of the crank angle after start of injection (CA deg ASI). The diffusion flame propagates through the combustion chamber until reaching the bowl rim represented by the white line. As the flame overcomes the bowl rim, the flame enters the squish region.

The standard case in all Diesel engines is that injection target is inside the bowl. Such point is defined as the intersection between the prolongation of the nozzle hole and the bowl rim. Injectors in CI engines target a point that is inside the bowl since the main purpose of the bowl is to realize a volume inside the combustion chamber that host the combustion.

The dimension of the squish region is dependent on the degree of the squish motion, that is, the motion that forces the air entering the bowl in compression (direct squish) or quitting the bowl in expansion (reverse squish). Squish is mainly caused by the difference of local pressure raised by the different compression ratio between the bowl region and the squish region.

Direct squish is an organized motion that promoting turbulence during combustion. On the other hand, reverse squish depends on the extracted products since it can lead soot to an environment full of oxygen (positive effect) or it can extract the burning flame out of the bowl causing its extinguishment (negative effect).

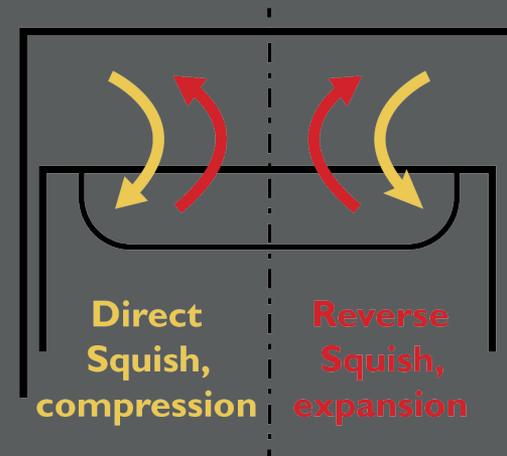


Figure 1.12 - Squish motion in the chamber

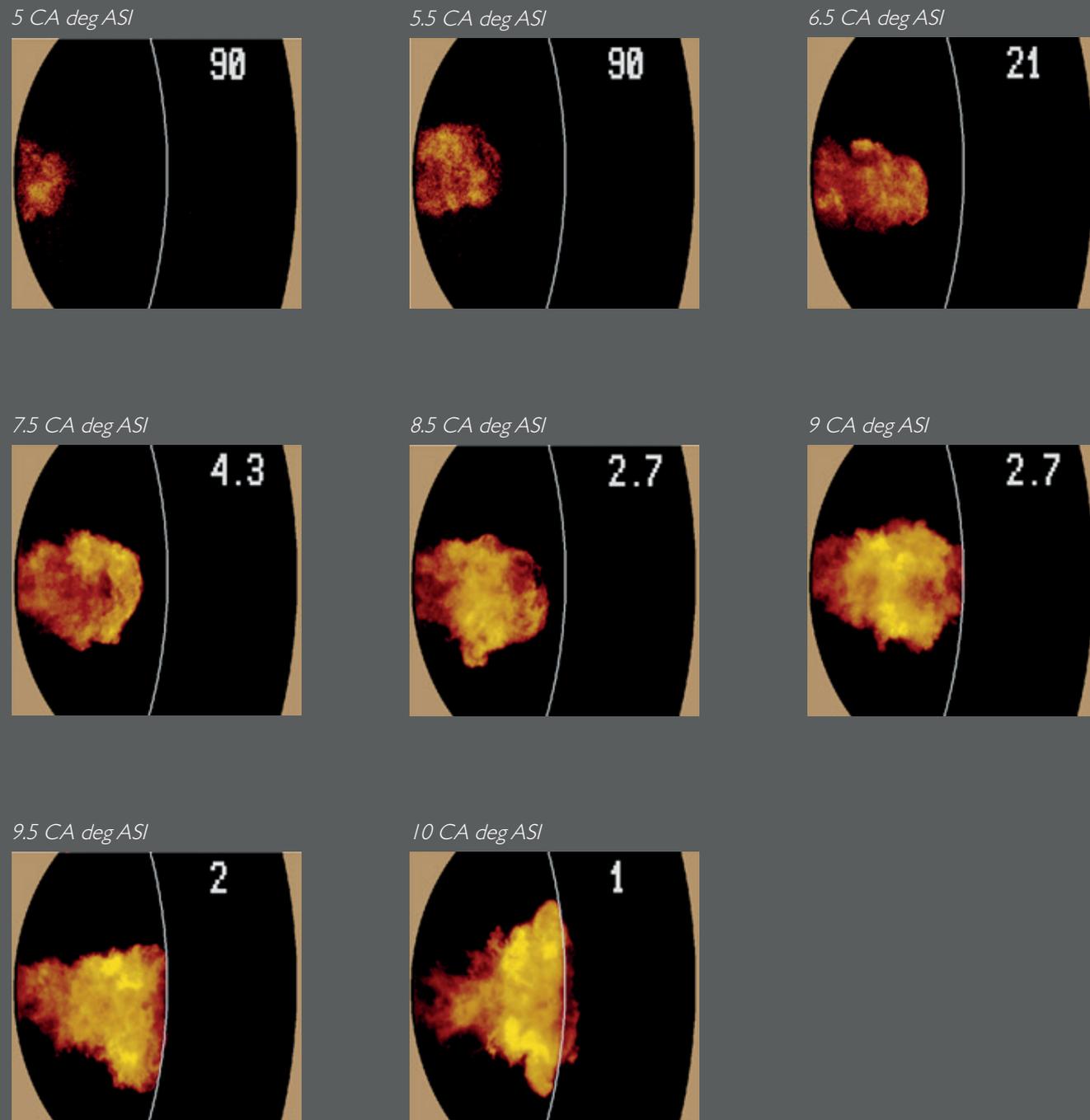


Figure 1.13 - Diffusion flame development ^[9]

1.3.3 Time-History Graph

The time-history graph of Figure 1.14 ^[10] plots the evolution of several parameters such as temperature, cumulative heat release, concentrations of oxygen, soot and NO_x as function of the relative time. Referring to the combustion plume, the relative time (scaled to 100) is the time needed by the droplets to travel the complete plume.

The particle temperature when, when it enters the liquid jet, is about few hundreds of Kelvin (~ 300, 400 K) because there is no combustion, but the injector tip is still in contact with the hot environment of the combustion chamber.

Along the process, the particle is going to heat up due to the mixing with hot air (~ 800 K, depending on the CR). Thus, the compression and the mixture raise the particle temperature till 850 K. The slope of temperature increases as preliminary chemical reactions with oxygen occurs. Such reactions are then responsible for the autoignition. Indeed the products of these reactions are represented by radicals. The time needed to trigger these reactions is identified as the chemical ignition delay.

When the equivalence ratio reaches the autoignition demand ($\phi \sim 2 \div 4$), the temperature raises up to 1600 K. However, this values is quite low since autoignition burns in fuel-rich condition and the fuel acts as a dilutant.

The further sudden increase in temperature may be understood from the equivalence ratio trend which, in the graph, is expressed as O₂ concentration.

When entering the combustion chamber, the oxygen concentration is zero (the plume is pure fuel). Then the O₂

concentration starts increasing because the fuel is mixed with air during final part of compression. When the reference single particle burns, it becomes burned gas and the chemical composition of this burned gas is formed of fuel, CO, hydrogen and very few carbon dioxide.

Consequently, after burning, the oxygen concentration turns back to zero and stuck until the temperature remains constant.

The particle then enters the diffusion flame, and within the flame it mixes with oxygen reaching a higher temperature since the diffusive combustion occurs at a level of equivalence ratio ϕ closer to unity.

However, in this condition, the local value of ϕ is determined by diffusion. After the diffusive combustion, the temperature decreases due to the effect of the piston expansion. In terms of oxygen, its concentration keeps increasing because also during expansion, the particle can entrain additional oxygen.

The cumulative HR stands for the chemical energy released by the single particle and not the apparent HRR (since the cumulative heat is a portion of the apparent). The chemical energy is released partially during the premixed burning and the major part during the diffusive burning.

By looking at the graph, the energy release proportions are in the range of 20% of energy in the premixed and 80% in the diffusive burning depending on how the combustion is structured. For example, the opposite case is reached for passenger car at partial load.

As seen, NO_x concentration raises up in the diffusive flame and part of the nitrogen oxide can still be formed after

the diffusion flame in the burned gas. As a matter of fact, burned gases can mix with hot air, thus with oxygen, reaching the thermal mechanism to produce NO_x.

Concerning soot and unburned HC, the hydrocarbon fragments and the polyaromatic carbons resulting from the premixed burning (fuel-rich combustion) are prone to create soot. If this portion of matter arrives to the diffusive burning, the flame oxidizes the matter cutting off the soot concentration.

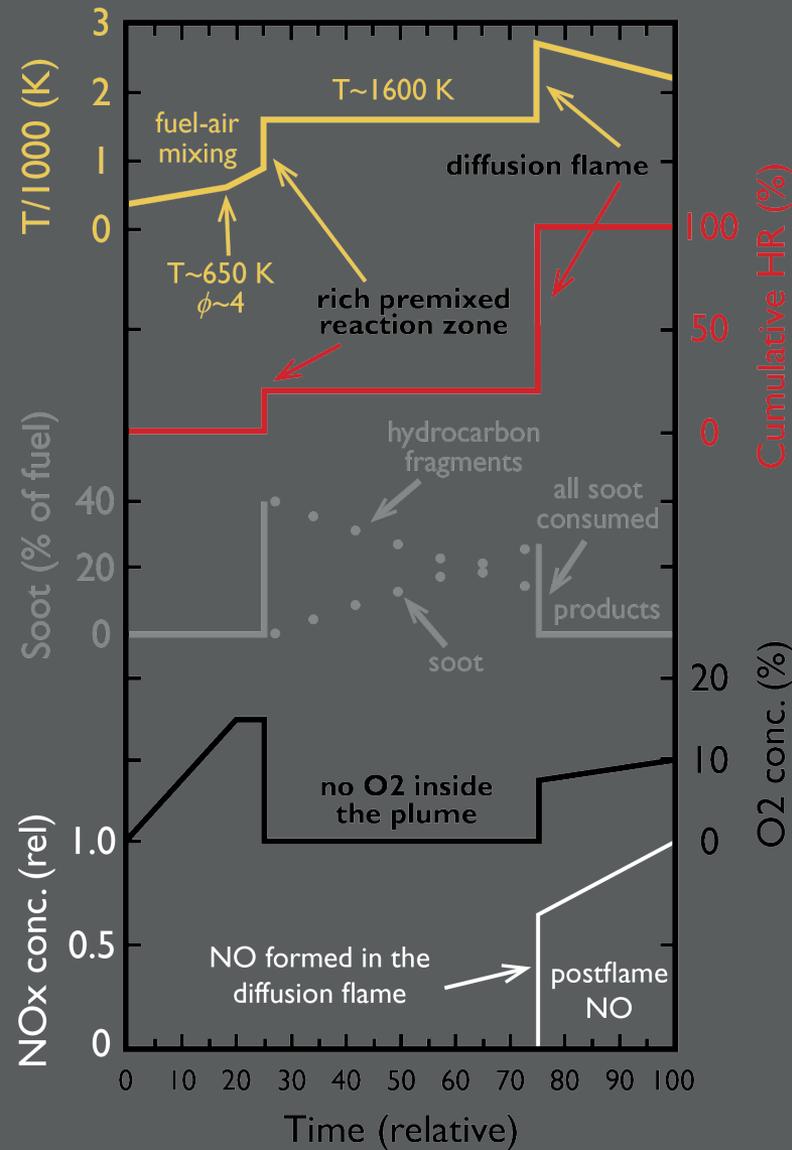


Figure 1.14 - Time-History graph of the CI combustion

1.4 Non-Conventional Combustion

Advanced combustion techniques aim to simultaneously reduce both NO_x and particulate matter (PM) without affecting performances and engine efficiency. Among all advanced combustion techniques, it is worth mentioning the Low Temperature Combustion (LTC). Principle advantages concerns the lower temperature peaks of a diluted mixture during combustion.

The main strategies to achieve a LTC can be divided into two main categories:

- HCCI - Homogeneous Charge Compression Ignition
- PCCI - Premixed Charge Compression Ignition

In turn, these two categories may be further divided depending on the injection time which may be *early injection* or *late injection*.

Both of the strategies are common by high degrees of EGR which allows to reduce temperature peaks, and thus slowing down the NO_x formation mechanism.

1.4.1 HCCI Combustion

The HCCI combustion, also known as CAI - *Controlled Auto-Ignition*, occurs without an external energy supply, as for any CI engine.

However, in a traditional process, the combustion occurs in a highly non-homogeneous condition, since the fuel is injected just before the desired combustion timing; consequently the mixture formation and combustion are bounded. Differently from the conventional combustion,

in the HCCI process, the fuel vapors are well mixed with air within the cylinder before the compression stroke. As a result, one obtains a homogeneous charge even in the lean side of the chamber by pre-forming the charge outside the cylinder through *port fuel injection*, or directly in the chamber through *direct fuel injection* during intake stroke or at the start of the compression stroke (very anticipated injection). In this way, the resultant mixture is homogeneous and generally poor. Such mixture allows to decrease the peak temperature, and reduce NO_x and PM emissions up to 98%. On the contrary, HC and CO tend to increase.

Despite this solution provides high levels of efficiency (mixture autoignition occurs according to an isochoric process, that in the ideal Otto cycle presents the highest efficiency, provided the same geometric CR), the HCCI strategy is not suitable for real applications. Suppose to design an HCCI engine and run it at a desired speed. By doubling the speed, the compression stroke will have half the time for being travelled.

Thus, the time required for make the mixture homogeneous has noticeably reduced for the same temperature and pressure condition. The autoignition then will not take place in the proximity of the TDC, but rather in the expansion stroke, destroying any advantage of the HCCI purpose, that is, isochoric combustion and high efficiency.

Consider now the case in which the engine speed has halved. Here, the combustion will take place early to the TDC, namely, during the compression stroke. The abrupt pressure peak, given by combustion and compression effect, will ultimately damage the engine, causing high stresses and affecting the process efficiency.

For these reasons, HCCI engines are not suitable for automotive applications, since these engines require to run at a high range of different speeds.

HCCI engines are not even suitable for stationary conditions (constant engine speed, for example, electric generators), because the different load conditions may affect the fuel-air mixture ratio, degrading the NO_x emissions. Indeed, the load variation has a similar effect to the speed variation. Different loads require different fuel quantities, the air/fuel ratio increases, combustion temperature increases, as well as NO_x emissions.

Figure 1.15 [3] shows on the left-hand side a traditional CI process, with the Diesel model combustion plume. Instead, the right-hand side shows a homogeneous charge obtained with very early, direct fuel injection. In HCCI, the combustion is not local, as for CI engines, but rather it takes place with multiple triggers. Ideally, the HCCI combustion provides that all the charge burns simultaneously.

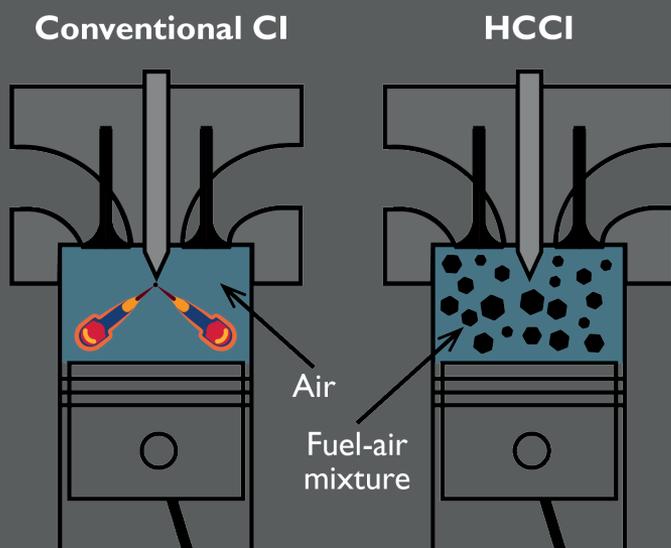


Figure 1.15 - Conventional CI vs HCCI

In conclusion, the main advantages of HCCI combustion can be summarized in the following terms:

- high thermodynamic efficiencies, due to a very fast combustion process, which approaches the ideal instantaneous combustion, and to the use of a very lean mixture;
- the engine operates un-throttled, spending a low pumping work and thus keeping high efficiencies also at partial loads,
- a premixed charge is burned (as in traditional SI engines), so there is no emission of particulate while a very low amount of NO_x is formed, since the burning charge is highly diluted with residual gases, able to reduce the maximum combustion temperatures.

On the other side the HCCI combustion process has still to face the following problems:

- charge mass increases accordingly to engine load. Then high heat releases and pressure rise rates occur, as in knocking combustion. So HCCI combustion can be safely obtained only at low loads,
- some products of incomplete combustion (CO and HC) are produced, since partial burning can occur at the boundaries of the small volumes of fresh charge, but they can be controlled by a proper after-treatment action,
- it is more difficult than in traditional CI engine to obtain a reliable control of the ignition timing and of the development of the combustion process, to optimize engine performance, fuel consumption, pollutant and noise emissions.

1.4.2 PCCI Combustion

The Premixed Charge Compression Ignition strategy is an alternative to the HCCI technique. Indeed, in the PCCI, the mixture is strongly lean, still not completely homogeneous. This implies that the combustion timing depends on the EGR variation and injection timing.

The experimental studies aim to increase the operative range of these advanced combustions operating on parameters such as fuel composition, injection timing and injected fuel mass flow rate. The PCCI process has been studied over a turbocharged single-cylinder Diesel engine with early direct injection provided by a high-pressure Common Rail system.

With respect to the conventional combustion, a low-temperature type features: longer liquid penetration, an extended ignition delay that allows more premixing of the fuel, a more distinct and temporally extended two-stage ignition, reduced and altered soot formation regions and an increased overmixing leading to incomplete combustion.

Late emission targets, also require the installation of after-treatment systems with higher costs, durability issues, fuel consumption penalties and packaging constraints. The improvements to in-cylinder strategies to further reduce the engine-out emissions and thereby rely less on the After Treatment System (ATS), with all the subsequent advantages, are then of great interest.

In diesel engines, the primary contribution to NO_x is given by the *Zeldovich thermal mechanism* whose production rates have an exponential dependency on the temperature. Low temperatures slow down the soot formation, but they hinder its oxidation, resulting in an increased en-

gine-out emission, as EGR increases. Hence, rather than using such high degree of EGR rates, to deal with soot emissions, the idea is to pursue an extensive pre-combustion mixing with the aim to limit the fuel-rich regions from which soot arises.

For conventional conditions, fuel is injected slightly before TDC and thus ignites immediately afterwards with the injection still ongoing; a *negative ignition dwell time*, which is the time from the end of injection to ignition, is therefore marked; as a result, only a small portion of the fuel is premixed prior to initiation of combustion.

For PCCI conditions, the ignition dwell is positive (still short enough to ensure that the jet structure remains intact up to the moment of ignition) meaning that the fuel undergoes to a premixing phase before burning.

To obtain this result, it is necessary to have high EGR rates and shorter injection durations, which make this strategy suitable to low-load operations only, considering also the dangerously steep pressure rises and noise that would be obtained at high fueling rates.

PCCI strategies further divide into *early* and *late* injection timing (with respect to TDC):

- in the early injection, the gases into which the fuel is injected are either initially cooler and less dense because of less compression,
- in the late injection, the gases into which the fuel is injected become cooler during and after the injection because of expansion.

By combining this, with the delaying effects of EGR on ignition chemistry, a relevant premixing can be obtained.

Note that the advantages in terms of NO_x and soot emissions provided by LTC strategies (HCCI, PCCI) always come at the price of increased CO, unburned HC, engine noise and fuel consumption. Whether these parameters exceed regulation, then PCCI combustion is not suitable for automotive applications.

Figure 1.16 [3] points out the initial penetration of liquid and vapor fuel for an early PCCI condition at a load of 4 bar (IMEP) in a heavy-duty engine. This is possible because Diesel fuel absorbs ultraviolet light very strongly. As a result, it is possible to clash a laser sheet with the fuel jet

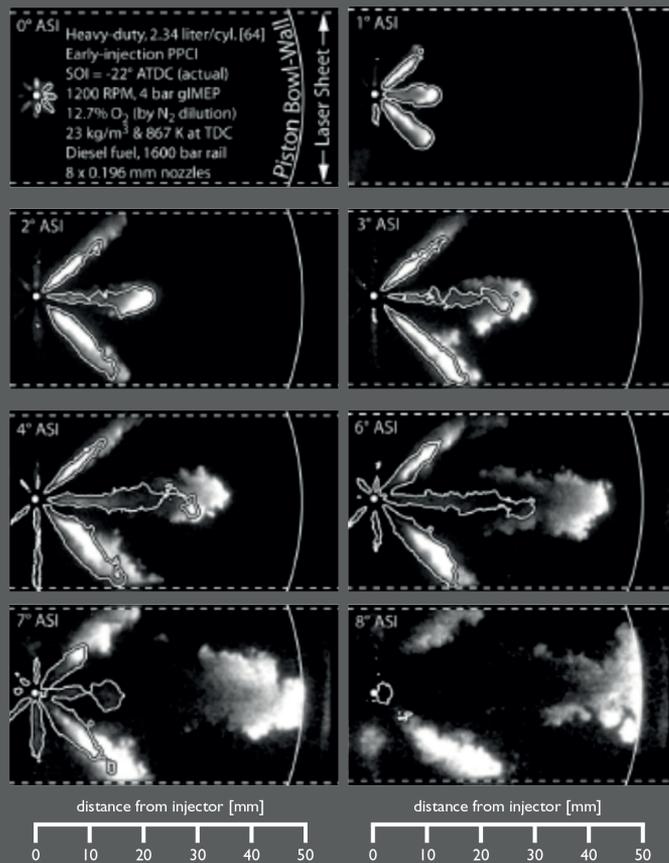


Figure 1.16 - Initial liquid/vapor fuel penetration for an early PCCI combustion in HD application

under investigation. At 3° ASI, the laser sheet is appreciably attenuated by the fuel at the tip of the spray till fading into a complete lack of the signal before end of injection (EOI).

Therefore, the intensity of the fluorescence emission should not be interpreted as an indication of the local fuel concentration but rather as a simple preliminary view of the temporal and spatial history of the leading edge of the jet.

In addition to the fluorescence imaging, a simultaneous visualization of liquid fuel by Lorenz-Mie scattering technique is provided in the same picture. However, the laser sheet used for this task is at lower frequency. This low-frequency laser allows to balance the lack of signal of the ultraviolet light.

The overall image thus provides a first idea of the relative spatial distribution of liquid and vapor fuel: the fluorescence signal arising from within the contours may be due to both liquid and vapor fuel while the one outside them shall come mainly from the vapor phase.

The initial spray penetration displayed in Figure 1.16 [3] is similar to that of a conventional Diesel combustion apart from the liquid length which is longer. Also, the jet velocity is higher because of the early timing and the higher injection pressures (1600 bar vs. 600 bar).

By 3.0° ASI the liquid length is established and the vapor continues penetrating on its own; due the signal attenuation (fuel absorbs the laser ultraviolet light), fluorescence from the upstream regions of the horizontal jet under examination is weak but the corresponding areas on the adjacent jets is still able to receive some direct light, indicating that fuel is present for the whole sequence duration.

From 4.0° ASI to 6.0° ASI, the liquid length remains approximately constant at the maximum value permitted by these conditions while at 5.0° ASI the injection rate peaks to end near 7.0° ASI after a ramp-down of about 2.0° , which is approximately $300 \mu\text{s}$ at the speed of 1200 RPM used for the experiment, during which the liquid penetration shortens up to disappear at 8.0° ASI.

PCCI Combustion Model

In view of the above, the model exploited for the analysis of PCCI combustion is an evolution of the Dec model. The combustion plume has been created starting from a Partially Premixed Compression Ignition (PPCI) which is another combustion which falls into the LTC processes.

The combustion plume evolution, plotted in Figure 1.17 ^[1], is valid for low load conditions and for heavy-duty engines, indeed, one of the main assumptions is that the jet does not interfere with the mechanical components of the engine as well as the fluids within the combustion chamber.

Referring to scattering analysis in Figure 1.16 ^[3], for the first 5° ASI, the fuel jet travels with no obstacles, mainly enhanced by the low density of the combustion chamber.

After the fuel injection peak, the decrease of fuel velocity creates a dragging wave (dotted vertical line in Figure 1.17). Note that such wave is created around 5° ASI and travels all the jet, promoting the evaporation and thus reducing the liquid penetration of the jet.

In this plume, chemiluminescence only appears after 6° ASI (indicated with the double arrow). In the following, the fuel tends to become looser while the tip starts burning.

Such combustion promotes the formation of the diffusion flame. Still, around this region, we have fuel which is burning at the previous intermediate step.

Rich area generate soot particles which may be oxidized in the following phases. The several combustion steps repeat for the whole volume of the combustion chamber, starting from the outside region, up to the injection point.

As seen, the combustion takes place differently than the standard one, and so the Heat Release Rate is different too. PCCI combustion occurs in a low-temperature, low-density chamber.

The additional dilution effect of the EGR delays the pre-reactions, thus obtaining a highly mixed charge compared to conventional CI engines.

Heat Release in PCCI combustion

As already mentioned and depicted in Figure 1.5, the Apparent Heat Release Rate of the conventional combustion may be divided into the combustion phases: ignition delay, premixed combustion and mixing controlled combustion.

The first negative peak is caused by the evaporation effect which subtract heat from the environment, followed by a step rise of the pressure as soon as the combustion starts (SOC).

The first positive peak is associated to the premixed combustion, triggered by the chemical kinetic, whereas the second positive peak is associated to the diffusive combustion controlled by mixing conditions.

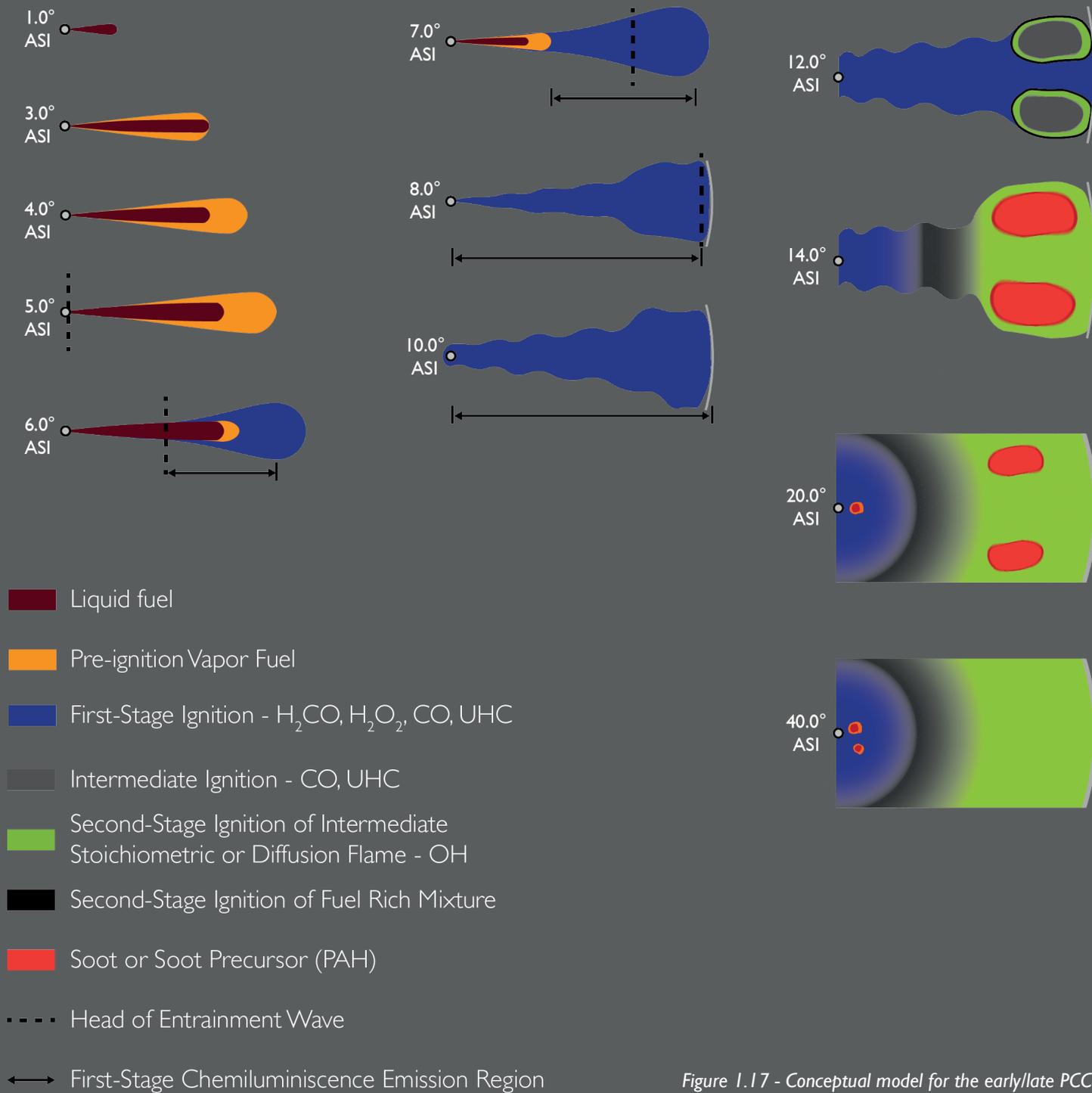


Figure 1.17 - Conceptual model for the early/late PCCI Diesel combustion in HD application

The heat release for a PCCI strategy admits an additional phase. Consider the case in which the injection is anticipated (*early injection*) in a low-temperature, low-density chamber. In a first stage, the trend is approximable to a conventional behavior. As usual, the combustion occurs when the AHRR reaches the null value after the negative peak. The first increase corresponds to *Low Temperature Heat Release* (LTHR).

After the LTHER region, the trend is reducing up to reach the second peak in the *second stage of ignition*, which has an amplitude three times higher compared to the conventional process. This highlights the load constraints for a LTC due to knocking and noise.

In turn, in a *late injection* PCCI, the premixed burn spike is less prominent, but the heat released during the mixing-controlled phase is again not negligible.

Consequently, for both early and late PCCI conditions, the AHRR preceding the mixing-controlled portion may be roughly divided into four different parts, each corresponding to a distinct phase of the chemical kinetics of ignition:

- First-stage pre-ignition,
- First stage-ignition,
- Second-stage pre-ignition,
- Second-stage ignition

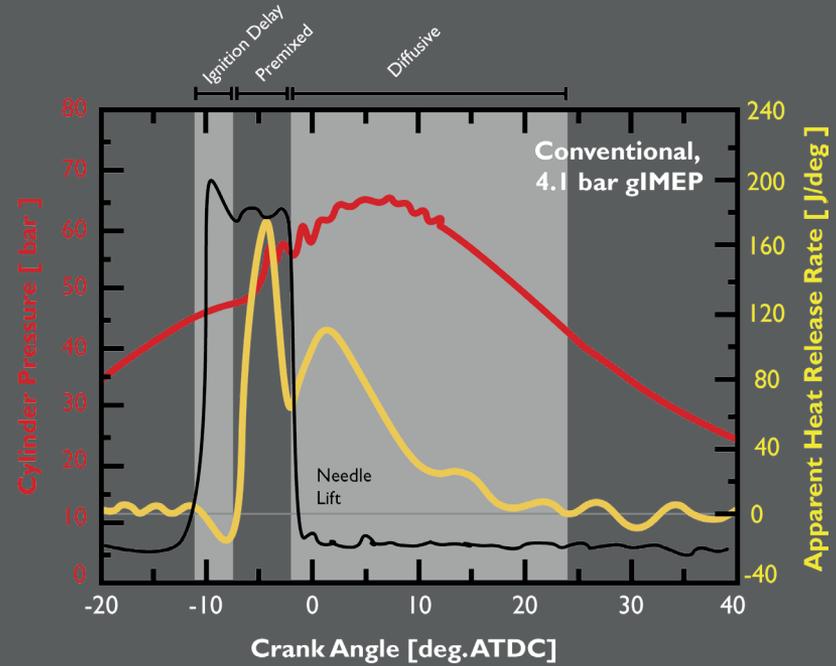


Figure 1.18 - Typical AHRR of a conventional Diesel Combustion ^[11]

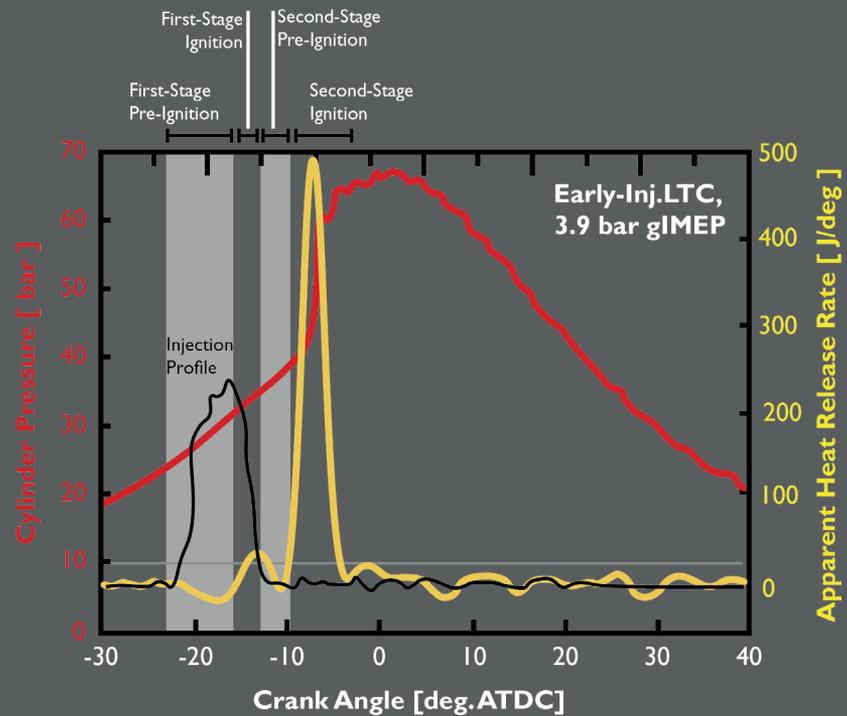


Figure 1.19 - Typical AHRR of an early PCCI Diesel Combustion ^[11]

1.4.3 The Kamimoto-Bae Diagram

The Kamimoto-Bae diagram in Figure 1.21 ^[13], reports on the vertical axis the equivalence ratio ϕ and on the horizontal axis the temperature T , indeed the diagram is also named ϕ - T diagram. To better understand the logic behind the diagram, it is required to refer to the model of the combustion chamber represented in Figure 1.20. By dividing the useful space into a grid, one creates a specific mesh, each cell having a value of ϕ and T . Stretching to the limit case this consideration (real case), every point could have different values of ϕ and T . Consequently, the values of the equivalence ratio ϕ and the temperature T are both function of the time but also of the specific region of the combustion chamber. For this reason, ϕ and T are local properties. In addition to these values of ϕ and T , it is possible to get the average quantity for a fixed time instant.

Trough the Kamimoto-Bae diagram, it is possible to represents the time-history of the average quantities of ϕ and T in the single-zone chamber. For example, considering a passenger car at partial load, the overall equivalence ratio results to be very lean depending on how low is the load. So the average value of ϕ is less than unity. However, the local value of ϕ at the premixed burning may be very rich even though the load is still low.

To get the average value of ϕ and T is required to perform a CFD calculation. As matter of fact the output of a CFD calculation is the cloud of points around the average line. The Kamimoto-Bae diagram is very useful concerning pollutant emissions. The main pollutants from Diesel engines are soot and NOx. In the diagram, there are present two peninsula that refer to the soot and NOx formation. These peninsula are estimated with a computational ap-

proach decoupled from the combustion, so it is possible to run a very detailed model for soot and NOx formation in which a specific value of ϕ and T is taken every iteration. Merging all these information in the Kamimoto-Bae diagram, it is possible to approximate the pollutant emission according to the combustion path. Depending on how premixed and diffusive burning are managed, one can enter or skip the soot production or NOx regions. Note that all different strategies result in a trade-off between soot and NOx formation. For example, by increasing the mixing, and so reaching a premixed burning whose equivalence ratio is lower than 2, probably the path will avoid the region of soot production but instead it will deeply enter the region of NOx formation and vice versa.

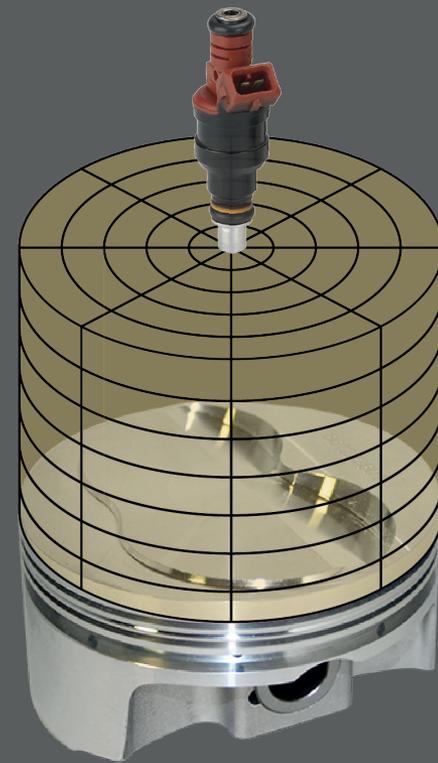


Figure 1.20 - Schematic example of a mesh in a combustion chamber

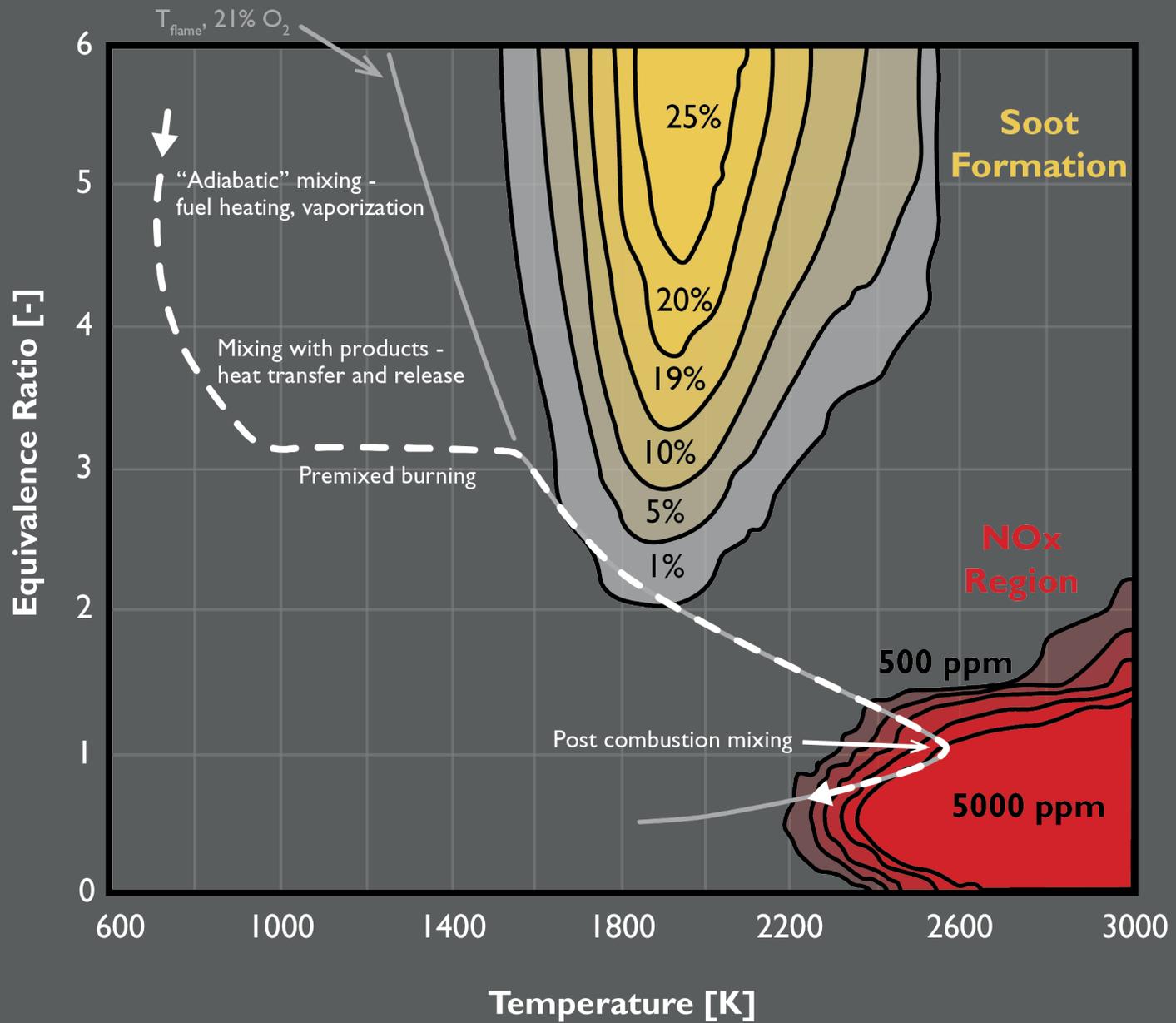


Figure 1.21 - Diesel combustion process through the Kamimoto-Bae diagram

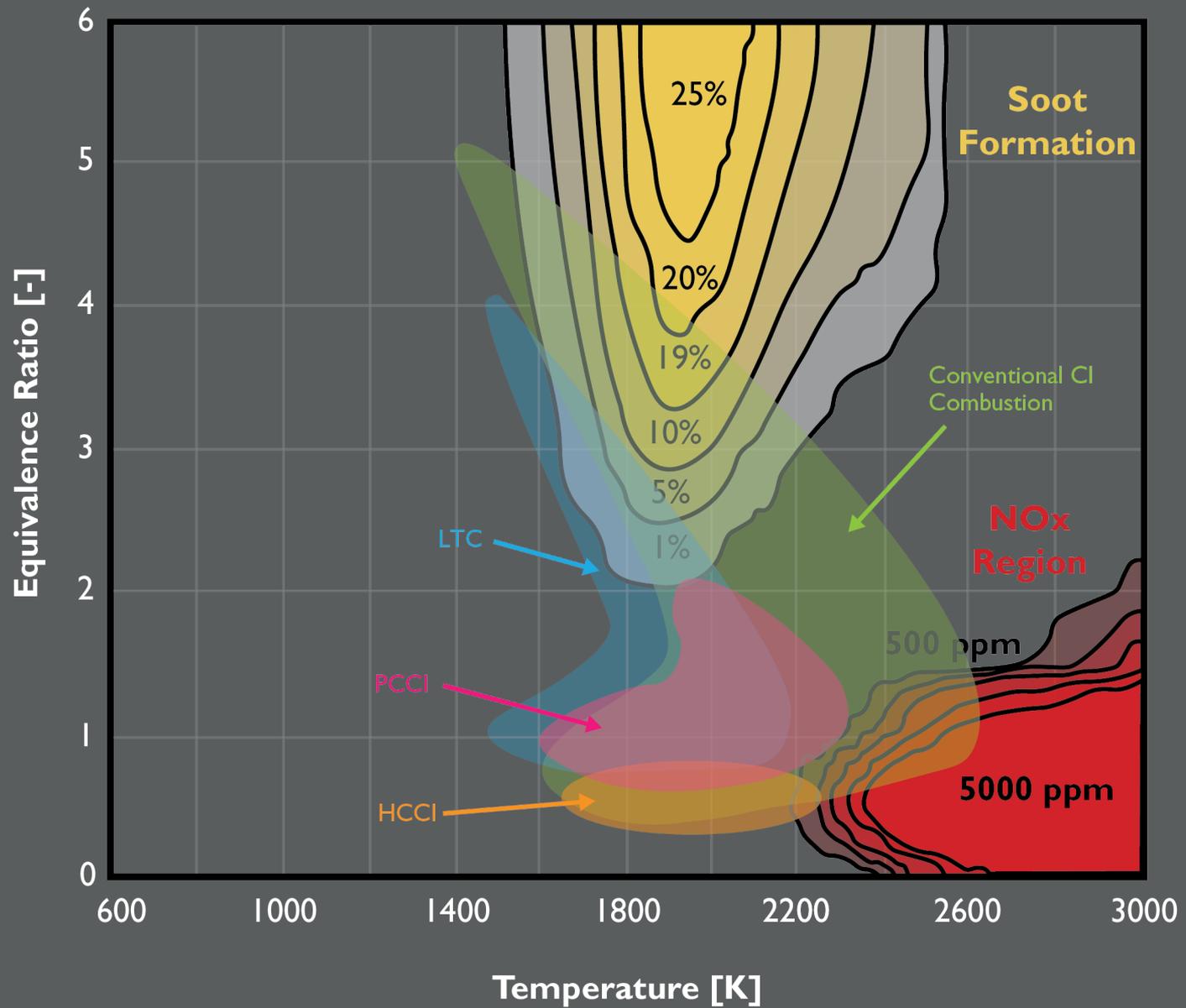


Figure 1.22 - Conventional combustion, LTC, PCCI and HCCI regions are highlighted in the Kamimoto-Bae diagram ^[3]

1.5 Pollutant Emissions

The main drawback of Internal Combustion Engines (ICE) consists in pollutant emissions. The road transport represents the major contribution to the atmospheric pollution. For this reason, nowadays the trend points towards alternative solutions such as electric vehicles (EVs) and hybrid electric vehicles (HEVs) which guarantee lower engine-out emissions. Modern ICEs rely on advanced After-Treatment Systems (ATS) to cut-off pollutant emissions that are dangerous both for human health and the environment. The knowledge of the specific pollutants and their formation mechanism then represents the most useful strategy to counteract engine-out emissions.

Since SI and CI engines operate with different fuels and combustion processes, then the pollutants are slightly different too.

In CI engines, pollutants formation mechanisms are highly dependent on fuel distribution within the combustion chamber and the mixture with air and residual gases (EGR). As a consequence of the Diesel engine working principle, the charge in the chamber is highly non-homogeneous causing local highly-rich and highly-poor regions.

The presence of these non-homogeneous regions are the main responsible for the particulate matter (PM) formation. Also, the great availability of oxygen causes peaks temperature that enhance the formation of nitrogen oxides (NO_x). The unburned hydrocarbons (HC) are one more pollutant species which is emitted by CI engines. Still, in this engine typology, the carbon monoxide (CO) does not get high relevance because the high oxygen availability allows the oxidation of CO into CO₂.

To recap, the main pollutants out of CI engines are:

- NO_x Nitrogen Oxides,
- PM Particulate Matter;
- HC Unburned Hydrocarbons.

1.5.1 Nitrogen Oxides - NO_x

The Nitrogen oxides NO_x emitted by CI engines are mainly composed of NO and NO₂. These are primary pollutants that directly react in atmosphere with HC and CO causing, in particular climate condition, the *photochemical smog*.

The photochemical smog is a process in which NO₂ is broken by UV solar radiation into NO and oxygen, highly reactive:



Afterwards, O reacts with O₂ to form ozone O₃, an highly harmful gas:



O₃ can then react with NO, producing NO₂ and O₂:



The abovementioned process would then reach an equilibrium, thus limiting O₃ concentration.

However, unburned hydrocarbons HC can break the equilibrium, leading to an increase in ozone and PeroxyAcyl Nitrates (PANs), which are powerful respiratory and eye irritants.

During the combustion in CI engines, the high temperatures break up the nitrogen and oxygen present in air into N and O. In turns, these elements are combined into NO and NO₂. Among all NO_x formation mechanisms, the dominant effect is the called *thermal mechanism*, described by the Extended Zeldovich Model (EZM) :



The oxygen atoms O in the exhaust gases burnt behind the front flame promote the nitrogen molecular decomposition into nitrogen atoms - reaction (1).

These, by reacting with molecular oxygen O₂ and with OH-radicals, enhance the formation of NO - reactions (2) and (3).

The activation energy of reaction (1) is about 319 kJ/mol and represents the main limit to the reaction velocity in most cases concerning ICEs. Such energy is supplied by the combustion process ($T > 1850 \text{ K}$).

About NO formation, this thermal mechanism becomes crucial especially in low-temperature, lean mixtures.

However, also in rich mixture, this mechanism is relevant since the formation rate of NO through EZM may be significantly reduce by reaction (3).

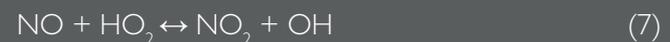
On the other hand, the NO decomposition mainly relies on the thermic path through N₂O, rather than EZM. Still, this does not have any visible effect over NO exhaust emission because reactions (1) and (3) are much slower compared to expansion and exhaust stroke, thus, the NO rate freezes. Indeed, the reaction velocity constant depends on temperature, accordingly to the following law:

$$k = aT^b e^{-\frac{A}{RT}}$$

- a, b, A constants characteristics of the analyzed reaction, independent from any other factor;
- e natural logarithm base
- R universal gas constant
- T absolute reference temperature

Differently from SI engines, in which the NO_x are represented by NO up to 99%, the CI engines produce around 70% of NO and 30% of NO₂.

From a chemical perspective, this major difference is explained by the following reactions:



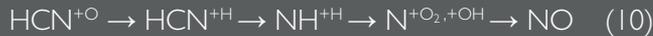
Reaction (7) shows that the NO may be converted into NO₂. In turns, reaction (8) explains that the NO₂ transforms back into NO by reacting with combustion products.

In CI engines, reaction (7) always takes place, whereas reaction (8) does not complete because temperatures are not sufficiently high.

These last two reactions are of paramount importance concerning the analysis of NO_2 because they represent the main input for some ATS, such as the Selective Catalyst Reduction (SCR).

Consider Figure 1.23^[4]. The NO does not start from a null value due to EGR degrees but also due to the *prompt formation mechanism*. This mechanism is extremely fast and reaction producing NO do not require oxygen but radicals and hydrocarbons in which temperature has small relevance.

Thus, the reactions of the prompt mechanism are enhanced by nitrogen atoms, and not by oxygen as for EZM:



Prompt NO is formed by the reaction of atmospheric nitrogen with hydrocarbon radicals in fuel-rich regions of flames, which is subsequently oxidized to form NO. Since the prompt NO mechanism requires a hydrocarbon or its radicals to initiate the reaction with nitrogen, this mechanism is much more prevalent in fuel-rich flames than in fuel-lean hydrocarbon flames. Prompt NO reactions are neglected in many NO models due to the increased complexity of the nitrogen chemistry and also due to intimate coupling of these reactions with the fuel oxidation steps. In addition, prompt NO is only significant in very fuel-rich systems and is a small portion of the total NO formed in most combustion systems. In practical combustion systems, which are usually operated fuel-lean or very close to stoichiometric, the contribution of prompt NO to the total NO formation is likely to be small.

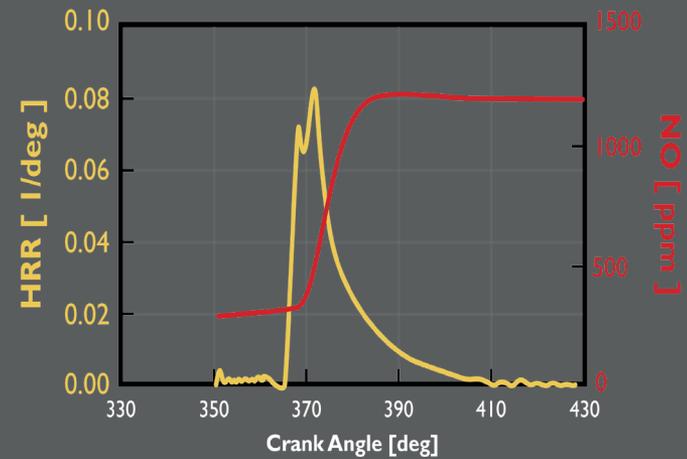


Figure 1.23 - NO typical trend in CI engine, 7 bar x 2000 rpm

Still referring to Figure 1.23, the NO trends quickly arises just after the increase of the HRR. Then, as the HRR drops down, the reactions producing NO actually freeze due to decreasing temperature given by the expansion stroke which mixes exhasuted gas with cooler air.

The formation mechanism of NO occurs mostly at the end of the combustion process, confirming that NO formation occurs during the diffusive combustion, as experimentally demonstrated through laser technique over optically accessible engines. As a matter of fact, the NO physically forms on the side of the fuel jet.

The main engine operating parameters affecting NO_x formation are:

- Engine load,
- Injection timing,
- Exhaust gas residuals - EGR,
- Injection pressure.

Engine Load

The decrease in equivalence ratio (increase of engine load), results in an overall decrease of NO_x emissions. This is because leaner mixtures produce combustion temperatures on average lower.

Consider a naturally aspirated Diesel engine and suppose to decrease the amount of fuel injected starting from a full load conditions (same injection timing and engine speeds). For the same amount of aspirated air, the combustion is obtained with a lower amount of injected fuel. Under this kind of experiment, a lower average temperature inside the cylinder is obtained, and, as a consequence, less favorable thermodynamic conditions for NO_x oxidation.

Figure 1.25 represents on the same plot the NO_x formation for both SI and CI engines: a decrease of NO_x emissions is experienced with an increase of A/F ratio (CI engines), whereas SI engines have a different sensitivity since they exploit homogeneous mixture. On the contrary, in CI engines operate with heterogeneous conditions for both temperature and concentration. The average temperature is, then, only representative, for a certain extent.

Injection Timing

Consider a simple injection system able to produce a single injection event per engine cycle. The evolution of this simple injection is represented in Figure 1.26. The target is to understand how a different injection timing can impact on nitrogen oxide emissions. In general, by increasing the injection advance, NO_x emissions tend to significantly increase.

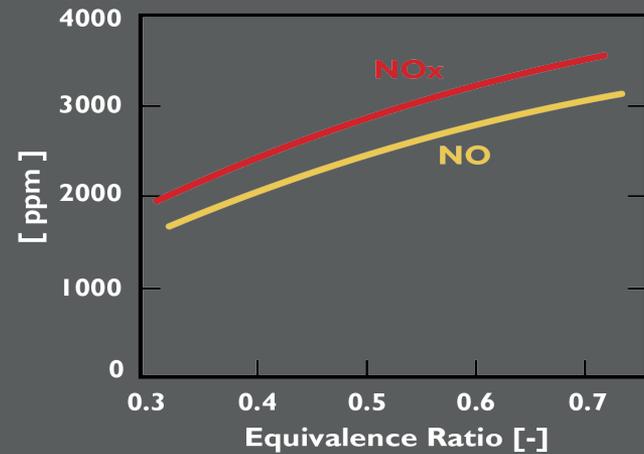


Figure 1.24 - Exhaust NO and NO_x concentration vs engine load

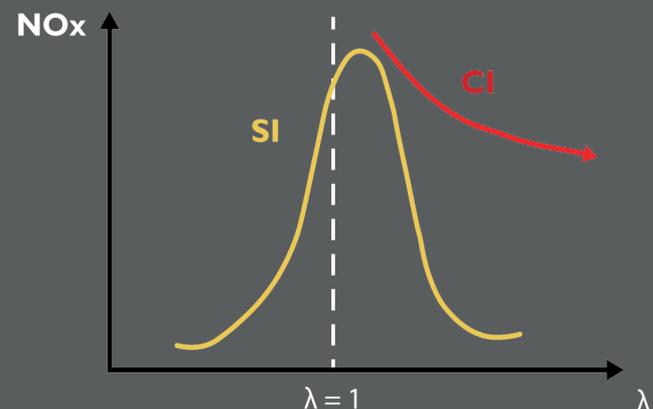


Figure 1.25 - NO_x formation in SI and CI engines vs A/F ratio

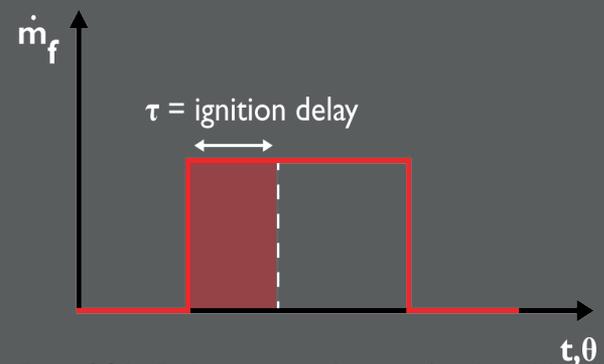


Figure 1.26 - Fuel mass injected vs time (crank angle)

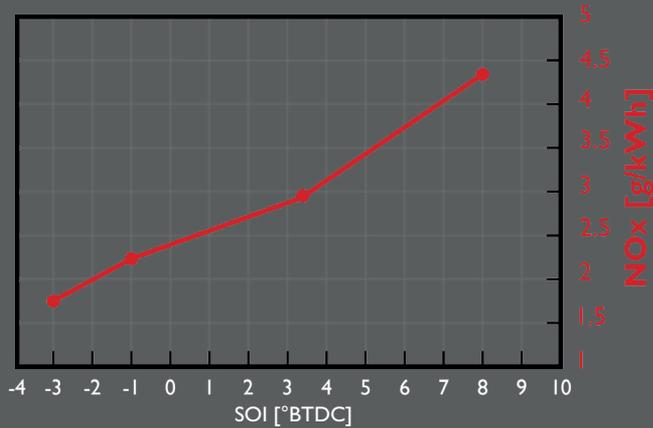


Figure 1.27 - NOx trend vs Injection advance

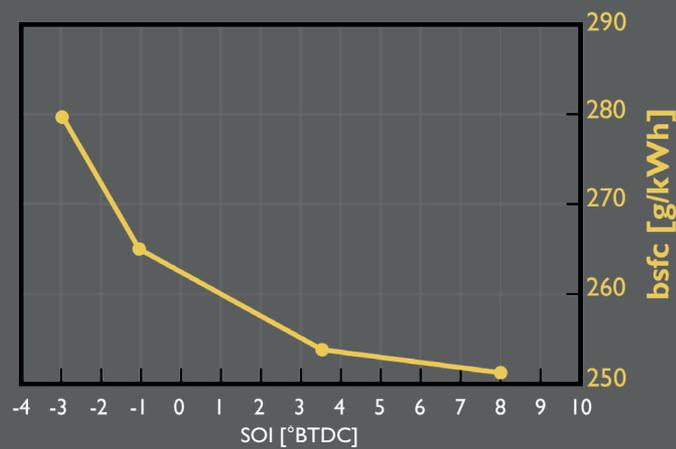


Figure 1.28 - bsfc trend vs Injection advance

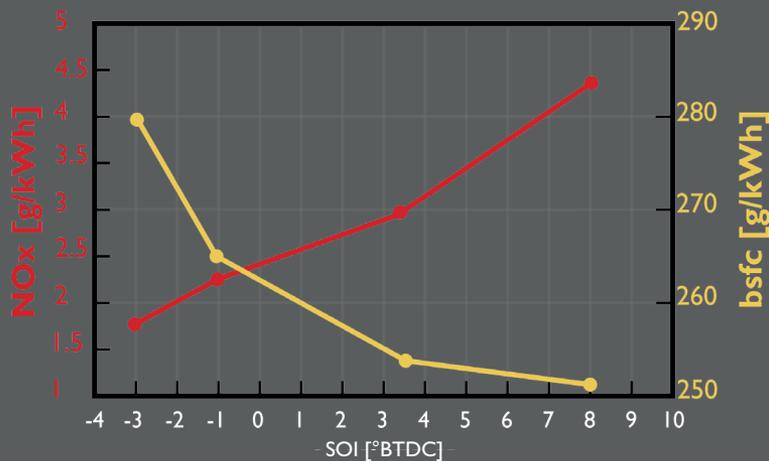


Figure 1.29 - NOx/bsfc trade-off vs Injection advance

The plot in Figure 1.28 [4] shows that, the more the injection advance is increased (higher degrees Before Top Dead Center - BTDC), the lower the temperature and pressure parameters in the combustion chamber at the Start of Injection - SOI.

This results in a longer injection delay: the injected fuel mass is higher, and it requires more time to break up from liquid spray to liquid droplets, and even more time to reach evaporation. Indeed, a lower temperature slows down all the chemical reactions from liquid to vapor:

The accumulated fuel mass, well-mixed with air, burns almost simultaneously, generating a peak in HRR. Despite the combustion approaches an isochoric combustion, it reaches higher temperatures that promote NOx formation.

It is worth mentioning that even though the NOx form during the diffusive burning, the premixed combustion trend determines the temperature of the mixed-controlled phase.

On the other side, because of such combustion, the efficiency of the system increases, leading to minor fuel consumption. The expression of the specific fuel consumption states that it is inversely proportional to the efficiency:

As pointed out by Figure 1.28, the higher the injection advance, the lower the brake specific fuel consumption.

$$q_b = \frac{\dot{m}_b}{P_u} = \frac{\dot{m}_b}{\dot{m}_b H_i \eta_u} = \frac{1}{H_i \eta_u}$$

Because of these advantages and disadvantages, it is required to calibrate an injection angle which is a suitable *trade-off* to minimize NO_x formation and specific fuel consumption.

In Figure 1.29 [4], it is possible to see the opposite trend of NO_x formation and Brake Specific Fuel Consumption (BSFC) with respect to the degree in advance of the SOI.

Exhaust Gas Residuals - EGR

The use of Exhaust Gas Residuals is the most spread solution in CI engines to drive down the temperature in the combustion chamber, and therefore NO_x emissions.

To measure the amount of exhaust gases which are recirculating, it is possible to exploit the carbon dioxide CO₂ in the intake as a smart tracer. In this way, the EGR may be expressed as:

$$\text{EGR [\% vol]} = \frac{\text{CO}_2 \text{ int \%}}{\text{CO}_2 \text{ exh \%}}$$

For example, consider a stoichiometric combustion of a Diesel fuel, which H/C ratio is 1.85. The products of the burning are represented, also, by water vapor, CO₂ and nitrogen.

The concentration of the carbon dioxide is around 13.4% by volume, while in the intake air is in the range of 400 ppm.

If the burned gases are recirculated and mixed with fresh air, targeting a 50-50% concentration, then the EGR concentration by volume is given by the average of the two.

Even though the primary effect of EGR is to reduce the peak combustion temperature to inhibit the NO_x formation, the combustion undergoes to other effects:

- Dilution effect

This is definitely the main effect since it has a greater impact over nitrogen oxides reduction. According to this effect, the mixture is diluted through inert gases that do not participate to the combustion but still they absorb heat. Therefore, the *same* thermal energy is distributed to higher number of particles so that the resultant temperature value of the combustion products is lower. Again, lower temperature means slower combustion reactions.

More in detail, to get the same amount of oxygen by increasing the EGR rate, for sure it is required to increase the amount of fresh intake charge. A higher air mass tends to absorb more heat, thus decreasing the temperature. By increasing the CO₂ through the EGR, more overall mass participate to the combustion, still with the same amount of oxygen.

Actually, this is equivalent to spread energy over a higher number of particles. In conclusion, the temperature is lower since these additional CO₂ molecules absorb part of the heat released by the combustion process

Note that the dilution effect can be misunderstood as the reduced oxygen concentration in the charge, but the main goal is to reduce the temperature. The lower temperature supplies lower kinetic energy to damaging the NO_x formation.

- Chemical effect

A higher EGR percentage at high temperatures enhance the dissociation phenomenon. As a result, CO_2 get dissociated into CO. This process absorbs heat causing an additional decrease of the temperature.

- Thermal effect

The overall specific heat of the mixture increases accordingly to the EGR rate. CO_2 and HC, which are present in the exhausted gas, retain more heat provided the same mass, compared to other chemical elements. In this way, the nitrogen is penalized in obtaining heat to form NO_x .

The scheme in Figure 1.30 [4] shows the investigation of the EGR effects over the NO_x formation. The most influential effect is the dilution with a small contribution of the chemical effect.

However, the thermal effect was found to be not relevant as dilution level goes up to 7 % even though CO_2 has a positively higher heat capacity than air (1.24 kJ/kg and 1.16 kJ/kg at 1000 K, respectively). This is hardly surprising since, adding 6% of CO_2 to air (the correspondent quantity for a 50% EGR), the specific heat capacity increases by less than 0.5%.

An important point is that the ERG % can have different effects on NO_x depending on the exhaust compositions, which, in a Diesel engine, is changing with engine load.

Typically, the higher the engine load, the higher the air utilization (close to stoichiometry). On the opposite, decreasing the load, a leaner condition is achieved.

If it is supposed to operate the engine without EGR, initially, at low load (e.g. 2 bar bmep), the mixture composition is, overall, extremely lean. This means that the exhaust gases contain a lot of fresh air, since the engine has been feed by air in excess.

Since diesel combustion is always lean, exhaust gases still contain significant amount of oxygen. The higher the engine load the lower the oxygen availability. Same percentages of recirculated exhaust gases may, therefore, produce different effects on nitrogen oxides (NO_x) emissions, according to the composition of the recirculated exhaust.

Test results indicate that high ratios of EGR need to be applied at low load but low ratios of EGR are sufficient for high load. When operating at lower loads, diesel engines generally tolerate higher EGR rate, because the exhaust contains a high O_2 concentration but low concentrations of combustion products, such as CO_2 and H_2O .

At high load, however, the exhaust oxygen becomes scarce and inert constituents become dominants.

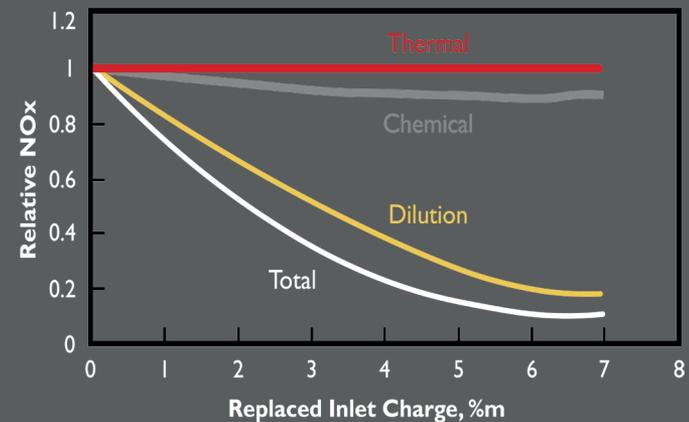


Figure 1.30 - Contributions of the EGR effects over NO_x formation

In general, if the EGR rate can benefit the reduction of NO_x, on the other side, it penalizes both the specific fuel consumption and PC emissions. In particular, Figure 1.31 [4] shows that the higher the EGR rate, the lower the NO_x emission, but the fuel consumption is worsened. Indeed, the dilution effect of the EGR decreases the temperature, meaning that the combustion efficiency is lower as well since the generated heat is lower. However, such consumption degradation is consider acceptable in respect of the greater benefit from NO_x reduction.

Figure 1.32 (2) points out that, concerning fuel consumption, the EGR strategy is the most suitable, rather than the injection advance. In contrast, the injection delay is the best technique to cut off PM emissions.

In conclusion, about the EGR effects, it should be ensured that:

- differently from SI engines, the EGR effectiveness reduces with respect to the engine load. The lower the fuel injected, the higher the oxygen availability, thus minor CO₂.
- The maximum rate of recirculating gases does not depend on combustion instability, but rather on pollutant emissions such as PM, HC and CO, provided the lower oxygen availability.
- High EGR rate may imply a heat exchanger (cooler) able to cool down the recirculating exhaust gases. In this way, it is possible to monitor the temperature of the intake charge and consequently regulate the maximum achievable temperature obtained from the combustion process.

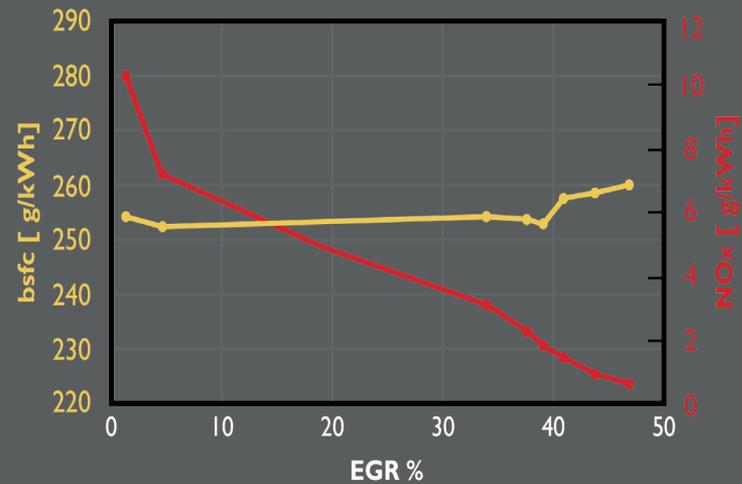


Figure 1.31 - NO_x/bsfc trade-off with respect to EGR rate

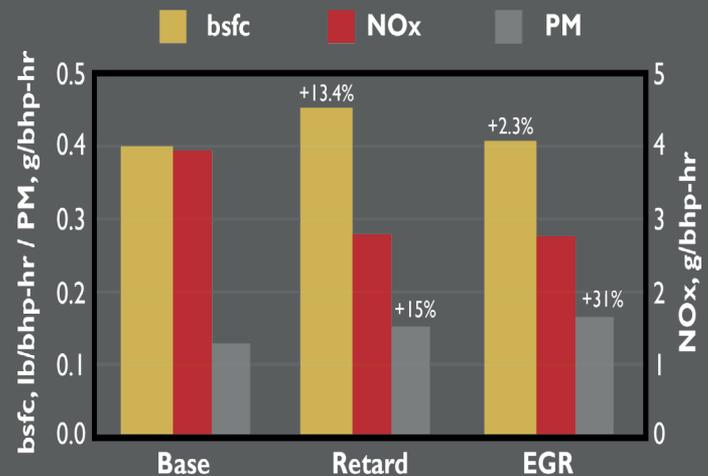


Figure 1.32 - EGR effect over later injection and bsfc

Injection Pressure

As seen, by increasing the injection advance, the NO_x emissions increase as well, together with combustion noise. The downside of such strategy is to worsen the brake specific fuel consumption. Trade-off between fuel consumption and NO_x emissions and between NO_x emissions and soot formation are necessary to obtain the desired target.

As far as injection pressure is concerned, it is necessary to focus on the possible modification of the mass fuel flow rate and the effect on the injection pressure on the Sauter Mean Diameter (SMD). Indeed, the droplets diameter is not constant, and the statistical distribution of the droplets diameter refers to the Rosin Rammler Distribution.

The SMD of a droplets population is the mean diameter that has the same volume-to-surface ratio. The higher the injection pressure, the lower the SMD, according to the trend presented in Figure 1.33. As a consequence of lower SMD, the evaporation is faster, thus shortening the ignition delay.

On the other hand, the effect of the shorter SMD when increasing the injection pressure from 1000 to 2000 bar is quite negligible. This is because the break-up phenomena are depending on the aerodynamic interaction between liquid fuel, jet and surrounding air. There is a kind of critical velocity: if the velocity is too low, the surface tensions which keep the droplet spherical will prevail on the aerodynamic forces, which tend to destroy the droplets.

Once that this level is exceeded, further increasing velocity will no more produce a further decrease in droplets.

Even if the injection delay of the black line is shorter, the main effect is still present: if the injection pressure is increased, the injection rate increases and the fuel is accumulated inside the combustion chamber, burning all together. This extremely intense combustion burning produces an important noise.

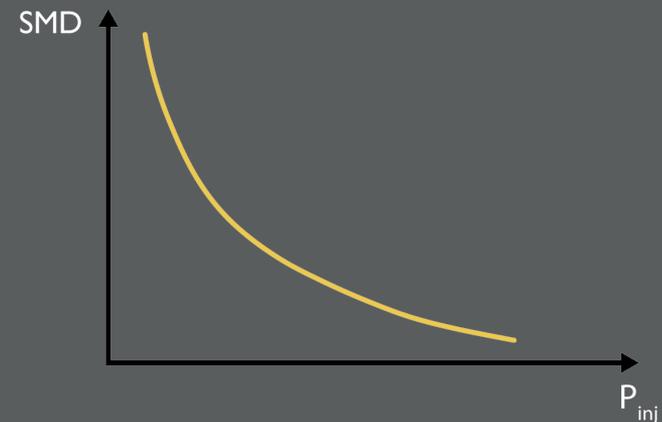


Figure 1.33 - Sauter Mean Diameter (SMD) trend as function of Injection Pressure

For a constant injection timing, the more the injection pressure is increased, the lower the bsfc. Usually an increase in injection pressure leads to a decrease of bsfc and soot, but to an increase in NO_x, combustion noise and peak in cylinder pressure.

Provided the same injection timing, the higher the injection pressure, the earlier the combustion process development, since both evaporation time and ignition delay are shorter.

Figure 1.34 ^[14] shows the benefit in terms of efficiency by increasing the injection pressure, however, such benefits are paid in terms of NO_x emission, as well as combustion noise.

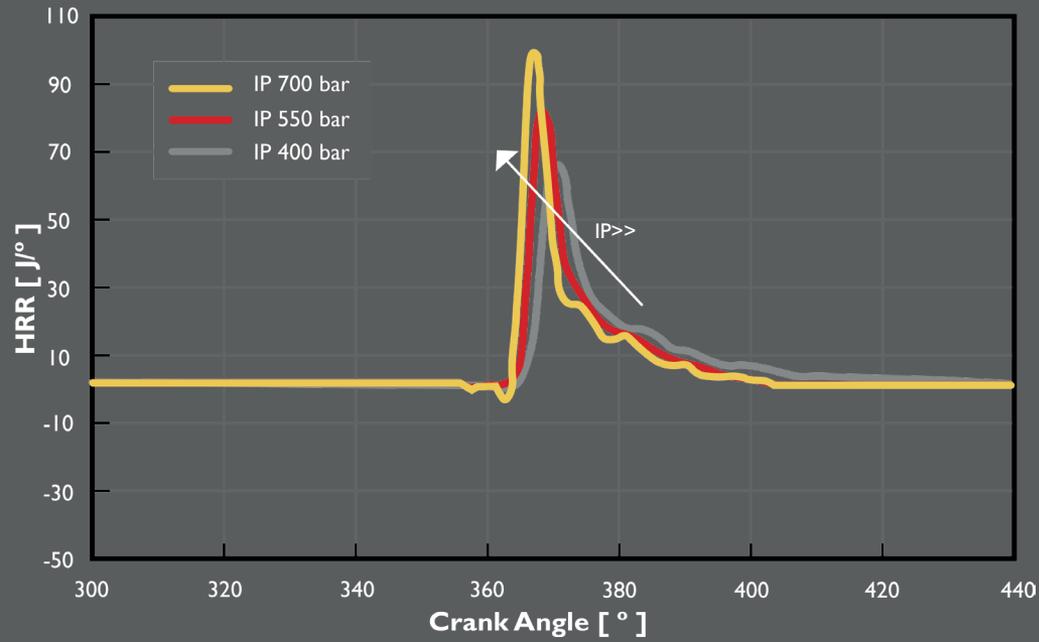


Figure 1.34 - HRR as function of Injection Pressure for the same Injection Timing (11° BTDC)

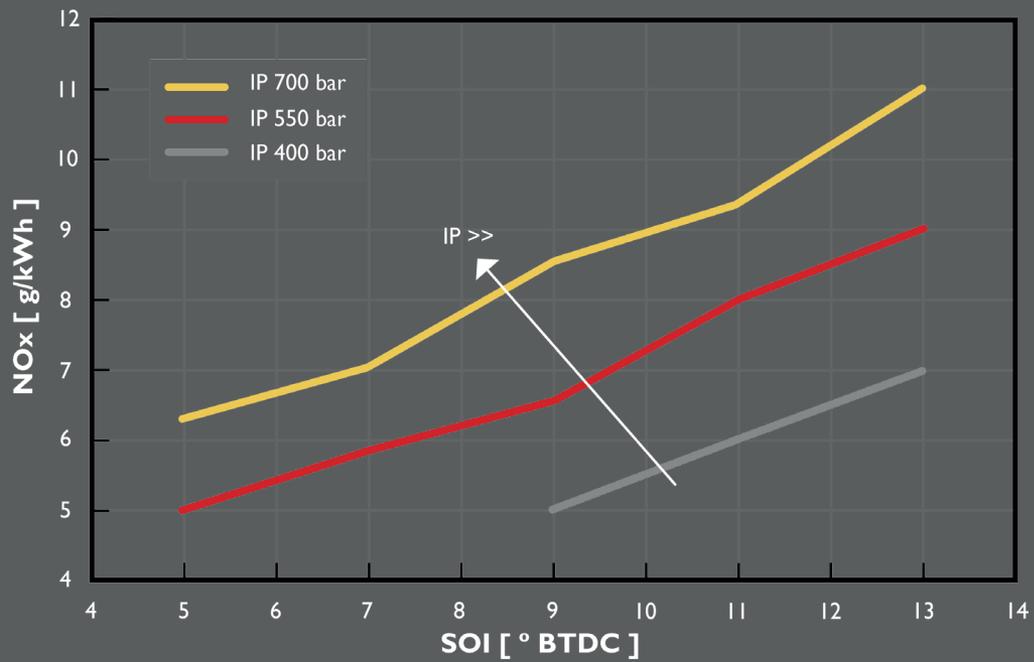


Figure 1.35 - NOx emissions as function of Injection Pressure for the same Injection Timing (11° BTDC)

1.5.2 Particulate Matter - PM

Particulate Matter (PM) does not have a specific chemical composition, it is an *aerosol* composed of solid, liquid and vapor substances.

The percentage of these contributions changes with the engine load, and that is why the measurement of PM is complex. In general, the chemical composition of PM refers to Figure 1.36 [15].

Because of this complexity of the chemical composition of PM, its definition is determined by the sampling method, whose detailed specification is an important part of all Diesel emission regulations.

PM sampling involves drawing an exhaust gas sample from the vehicles exhaust system, diluting it with air, and filtering

through sampling filters. The mass of particulate emissions is determined based on the weight of PM collected on the sampling filter.

It is quite obvious that any changes in the procedure, for example using a different type of sampling filter or different dilution parameters, may produce different result.

The main substances that form the PM are:

- SOL solid carbonaceous particles and ash
- SOF organic compounds from combustion and lubricant
- SO_4 sulphates generated by the presence of sulphur in the fuel

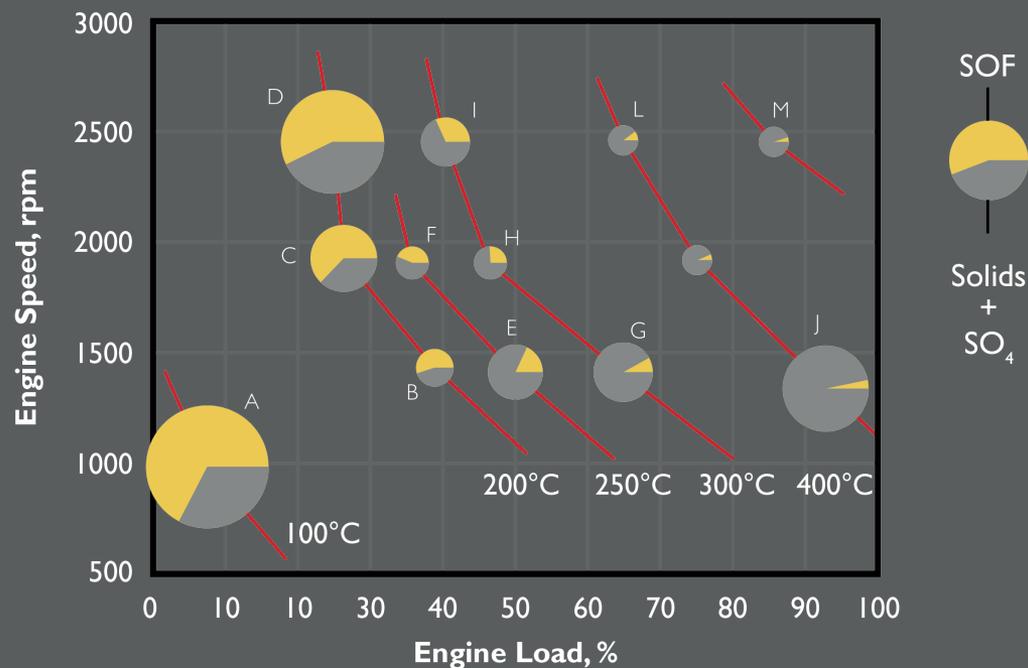


Figure 1.36 - Particulate Matter (PM) composition compared to Engine Load

Solid Fraction - SOL

The solid fraction of Particulate Matter can be further divided into carbonaceous particles, that is soot, and ash.

The carbonaceous particles is responsible for characteristic *black smoke* in Diesel engines. Soot initially forms in the premixed phase, in which the main combustion products are polyaromatic hydrocarbons (PAH). These HC are soot precursors both in low temperature combustion through a direct mechanism, but also in high-temperature through an indirect mechanism, provided stoichiometric mixtures. The direct mechanism is faster and relies on the thermal cracking of the fuel, whereas the indirect mechanism is slower and provide the composition of crystals once the hexagonal carbon structures gather into plates which are actually called *platelets*. These carbon sheets join together up to form the abovementioned crystals.

On the other side, the ash represents the non-combustible solid residuals. It forms both within the combustion chamber but also in the exhaust manifold of the engine. The first derives from the presence of additives in the lubricant oil that contains substances such as detergents, dispersants, and anti-oxidizing agents which are non-combustible. In addition, metallic oxides from engine wear are conveyed through oil directly into the combustion chamber; participating to the formation of ash which is then dragged into the exhaust system. Other ash is present in the exhaust manifold due to corrosion or as a consequence of additives in the fuel. As a matter of fact, fuel is also exploited to regenerate the anti-particulate filter by burning the stuck residuals. Thus, the combustion of the fuel into the catalyst generates ash in the exhaust system.

Soluble Organic Compound - SOF

Solid particles travels into the exhaust gases that are also formed by organic compounds *survived* to the combustion process.

In certain conditions, it is possible that such organic compounds undergo to condensation processes, or, they are absorbed by carbonaceous particles since these soot particles have a certain porosity degree.

In both cases, these organic compounds represents the soluble organic compound fraction (SOF). Since this fraction relies on the condensation of organic compounds, then it is generated mainly at low temperatures, i.e. in the exhaust system, and it is mainly composed of organic substances of the oil and fuel, such as ketones, aldehydes and paraffins which are mutagenic and harmful to human health.

Sulphates - SO_4

Automotive fuel may contain traces of sulphur that, during the combustion process may transform into sulphur oxide. It may further bind to water generating sulphur acids.

These acids produce sulphates through a hetero-molecular nucleation process. Such mechanism is promoted at low temperature and so, sulphates are mainly generated in the exhaust manifold.

Furthermore, sulphur acid may also react with oil additives and form sulphur salts.

As pointed out by Figure 1.36, the composition of the PM is highly dependent on the engine load. The soluble fraction becomes relevant in lot-to-medium load, whereas sulphur and solid fraction are much relevant in high loads. Both the composition and the measure of PM is highly complex.

When measuring the PM in laboratories, it is required to dilute PM into fresh air because it is necessary to simulate the release of PM to the external environment from the muffler of vehicles. Temperature variation causes a different PM composition since several substances may condense or evaporate. For this reason, regulations state that the PM must be measured at a standard temperature and standard dilution. Furthermore, one has to consider that the particulate must be oxidized in the anti-particulate filter; which implies an ash accumulation in the filter itself. All these considerations give rise to a common *operative definition* of PM, that is:

PM is the substance collected from the exhaust of an Internal Combustion Engine whose exhaust has been provided with a sampling filter kept at 52°C.

Concerning the particles formation, it is possible to define three different events:

- Nuclei mode

This give birth to nano-particles with a diameter lower than 50 nm. These are generated by condensation of organic compound in the exhaust system, right after the PM filter. Such nano-particles have high relevance among about the total number of particles in PM, but in contrast, they are weak in terms of mass contribution.

- Accumulation mode

This mode produces particles with a diameter lower than 1 μm but higher than 100 nm. These are formed during combustion and have high contribution in terms of mass.

- Coarse mode

This mode produces bigger particles with a diameter lower than 10 μm . These are formed during combustion due to engine and manifold wear. Despite these are the biggest particles, they have lesser contribution both in terms of mass, as well in number.

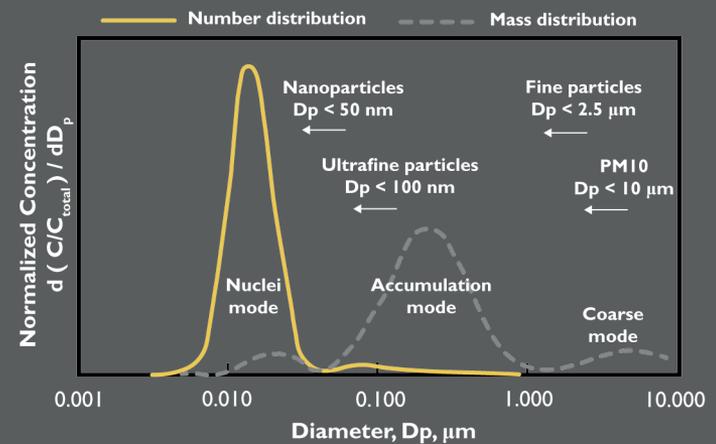


Figure 1.37 - Particulate particles formation mechanism ^[4]

As for NOX emission, the main engine operating parameters affecting the PM formation are:

- Engine Load,
- Injection timing,
- Exhaust gas residuals - EGR,
- Injection pressure.

Engine Load

Particulate Matter emissions usually show an abrupt increase for A/F ratio (engine load) below a certain threshold, which is usually referred to as *smoke limit*.

In Figure 1.38 [4], a naturally aspirated Diesel engine is reported, so that the effect of boost level variation effect are not present, and the same amount of air is always injected. Changing the A/F ratio actually means to vary the fuel injected depending on the load that has to be achieved.

This will correspond to a minimum A/F ratio for full load operation condition and a maximum A/F ratio for idle operation. The PM trend shows that it has almost no sensitivity to A/F ratio for quite lean mixture.

Then, a quasi-vertical asymptotic change is representing the so-called *smoke limit*. If the engine is charged with an amount of fuel exceeding this limit, the soot or PM emission dramatically increases.

This limit is a kind of characteristic data for the specific engine, depending on the combustion chamber geometry and on the air/fuel injection system characteristics. The higher the air utilization, the closer the soot limit will be to the stoichiometric value.

In a Common Rail automotive Diesel engine with good air utilization, it is possible to push the smoke limit very close to the stoichiometric value. This is also the reason why high amount of smoke can be produced by old vehicles on mountain roads: at higher altitudes, the density of the intake air is decreasing and, if the system continues to inject the same amount of fuel the operating point moves towards the left of the diagram of Figure 1.38.

Injection Timing

PM emissions are also extremely sensitive to injection timing, usually showing a significant increase with decreasing injection timing. A PM/NO_x trade-off, in Figure 1.39 [4], is therefore necessary. Even though, during premixed burning, the fuel is forming soot, this has good chances to be later on oxidized, since the fuel has been burned in advance in the combustion process. Substantial improvements in the trade-off can be found with multiple injections.

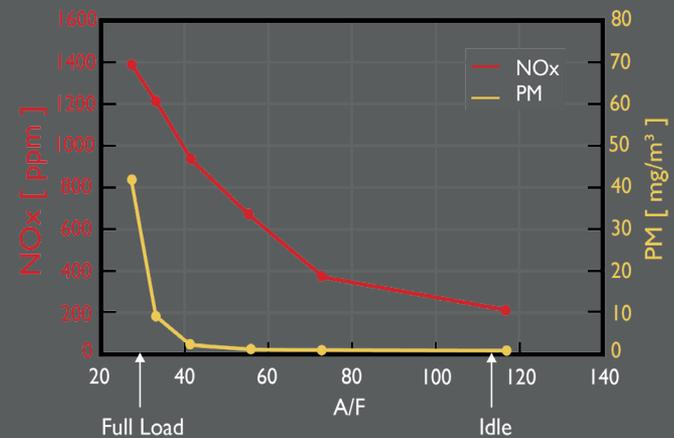


Figure 1.38 - PM/NO_x formation as function of Engine Load

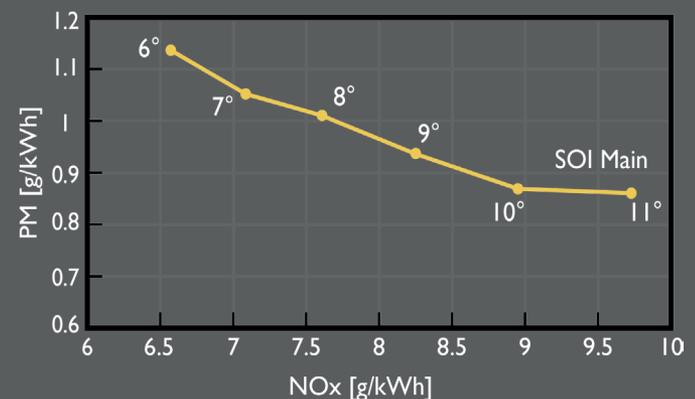


Figure 1.39 - PM/NO_x trade-off as function of Injection Timing

Exhaust Gas Residuals - EGR

The EGR rate mainly affects the temperature of the combustion through the dilution effects, but it also affects the oxygen availability and consequently it also regulates the soot oxidation. Furthermore, the EGR slows down the combustion process and improves the air-fuel mixture. As a result, the combustion noise decreases while efficiency increases.

Another positive effect of EGR is to recirculate part of the exhaust gases which have not been already oxidized. In this way, these gases have another chance to complete the oxidation, thus reducing soot emission at the engine outlet.

However, one must consider that the higher the EGR rate, the lower the oxygen sucked-in by the engine, since the volume within the chamber is fixed, and lower oxygen implies higher particulate emission. The trade-off presented in Figure 1.40^[4], targets acceptable results in terms of soot formation and NO_x emission to comply the limits imposed by the existing regulations.

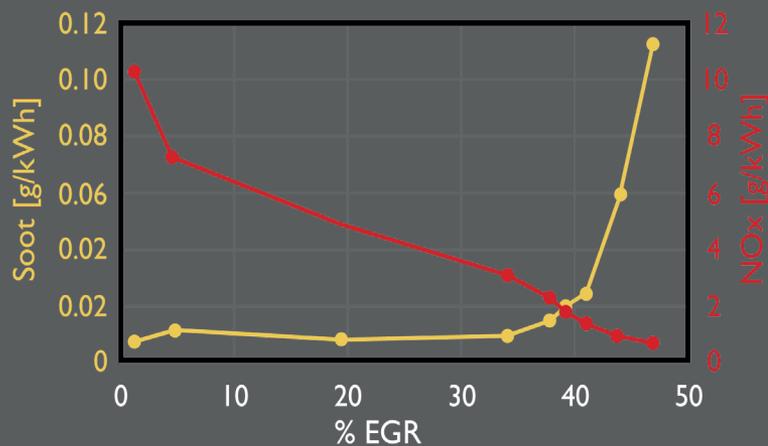


Figure 1.40 - PM/NO_x formation as function of Injection Timing

Injection Pressure

The injected pressure has direct effect on fuel atomization, i.e. on the mixture equivalence ratio ϕ . Higher pressure ratio benefits the particulate formation due to the better nebulization of the jet.

However, the end-stroke of the needle causes fuel bleeding due to the flux speed decrease. This bleeding negatively impacts on HC formation as well as particulate emission.

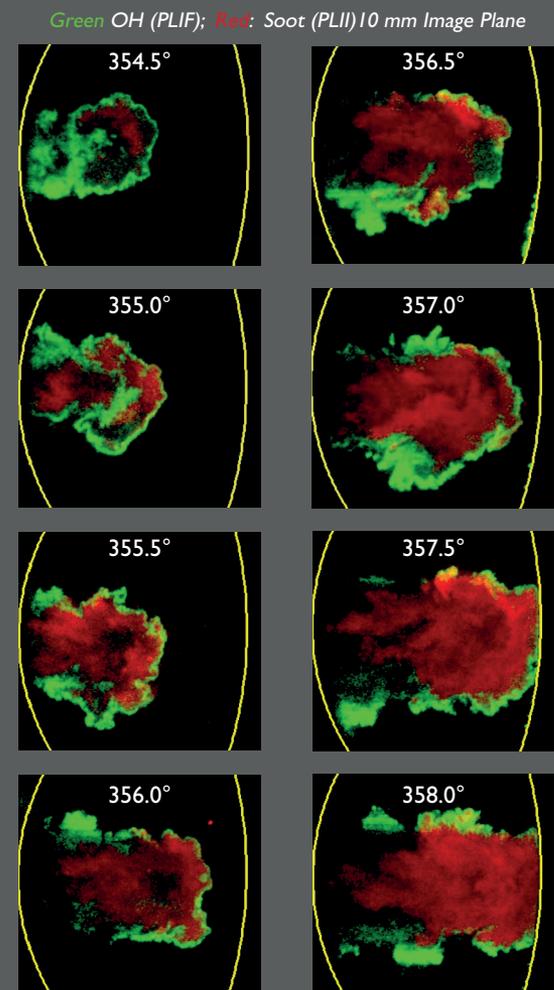


Figure 1.41 - Soot analysis (red) in the diffusive flame (green)^[9]

Higher...	Engine Load	Injection Timing	EGR Rate	Injection Pressure
Nitrogen Oxides - NOx	<p>↑</p> <p>Higher engine load means higher equivalence ratio ϕ, With rich mixture the combustion temperature increases promoting the thermodynamic condition for NOx formation</p>	<p>↑</p> <p>The higher the advance (longer ignition delay), the lower the the temperature and pressure in the chamber. Premixed combustion increases promoting NOx formation</p>	<p>↓</p> <p>The higher the EGR rate, the lower the combustion temperature. Inert exhaust gases absorbs heat obtaining a dilution effect of the charge. EGR effectiveness depends on exhaust composition</p>	<p>↑</p> <p>The higher the injection pressure, the higher the fuel rate, the atomization resulting in higher HRR, combustion temperature and pressure. Faster premixed burning promotes NOx in the diffusive burning</p>
Particulate Matter - PM	<p>↑</p> <p>Severe lack of oxygen leads to soot production in premixed burning and soot formation later on. Emission rapidly increases at engine smoke limit. The higher the air utilization, the closer th soot limit to stoichiometric value</p>	<p>↓</p> <p>Premixed combustion gives rise to soot production, howeverm the earlier the fuel burns, the higher the probability to oxidize the soot during the rest of the combustion due to the entrainment of fresh air</p>	<p>Negligible</p> <p>EGR mainly does no affect soot production. However, EGR is representative of the intake oxygen availability which is crucial for soot production</p>	<p>↓</p> <p>The higher the injection pressure, the better the atomization resulting in a more homogeneous equivalence ratio.</p>

Table 1.2 - NOx and PM qualitative trend vs engine parameters

1.5.3 Unburned Hydrocarbons - HC

Unburned Hydrocarbons are compound that do not participate to the combustion process. Even though they are called *unburned*, their chemical composition at the exhaust is not the same as in the intake. Actually these pollutants are generated by the partial combustion of the hydrocarbons making up the fuel.

Also, fuel burning in abnormal conditions significantly increases the HC emissions (about 10%). These unexpected burning refers to:

- fuel trapped into cylinder crevices,
- fuel trapped in wall thermal boundary layers,
- bulk quench due to the high quantity of residuals.

That is why the chemical composition of HC at the engine exhaust is very complex and involves a very extended range of molecular masses (more than 400 different organic compound have been detected).

During the Diesel fuel combustion, part of the hydrocarbons undergoes to pyrolysis reactions giving rise to solid particulate particles because a substantial liquid fraction of HC is absorbed by carbonaceous nuclei. This directly affects the technique to collect HC at engine exhaust for analysis purpose

As for PM, the hardware needed for the measurement is represented by a particulate filter and a sampling line, both maintained at 190°C in order to keep in vapor phase all the HC compound, since they have lower condensing temperature. The remaining fraction is considered as particulate emission (SOF - soluble fraction).

Concerning HC, the main formation mechanism are:

- Overmixing

The local equivalence ratio ϕ within the combustion chamber is strongly affected by the turbulent motion. However, excessive turbulences, especially in the expansion stroke., may generate too lean regions in the chamber that do not promote combustion. Such areas are located in the periphery of the fuel jet and, despite they do not burn completely, they are nevertheless subjected to a slow oxidation process. This behavior of the jet is one of the most crucial cause about HC formation.

- Undermixing

A poor local mixing also leads to an abnormal HC formation. When the mixture is highly rich, the oxygen cannot easily reach all the regions of the combustion chamber. This phenomenon usually takes place the last injection event, when the outflowing velocity of the fuel is lower and atomization is poor. Furthermore, fuel bleeding may have non-negligible contribution to undermixing. The high content of liquid fuel cause the increase of HC as well as carbonaceous deposits that may interfere with the injection event (*cocking* phenomenon).

- Wall impingement

The excessive fuel penetration may cause the jet to clash with the wall of the piston bowl. The relatively cold wall temperature delays the fuel evaporation, and so the overall liquid content may increase, thus enhancing the HC formation.

The engine parameters affecting the HC formation are:

- Injection timing,
- Injection pressure,
- Exhaust gas residuals - EGR.

Injection Timing

The injection timing has high relevance on HC formation because it directly affects the ignition delay, as already mentioned in the NO_x analysis. In turn, the ignition delay impacts over the overmixing and undermixing phenomena. In particular, the early injection, i.e. when the piston is relatively far from the TDC, may enhance overmixing, whereas undermixing is promoted for late injections because the fuel does not have the time required to create a sufficiently-mixed charge. Figure 1.42^[3] shows that despite the earlier injection benefits the mixture of the charge, meaning that the overall equivalence ratio is more homogeneous, the dominant effect is the increase of HC.

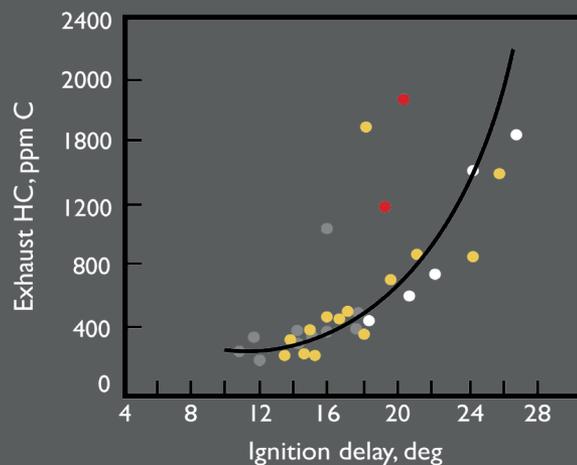


Figure 1.42 - HC trend as function of Ignition delay

In case of big bore engines, this phenomenon does not occur, and so higher injection pressure only brings benefit.

On the other hand, smaller engine must exploit recent technologies which permitted to obtain nozzles with more holes, with lower diameters so to avoid wall impingement.

That is the reason why from the last decades, the trend of increasing the injection pressure was not penalized by wall impingement.

Exhaust Gas Residuals - EGR

The EGR degree affects the intake temperature and the mixture of the charge.

In particular, higher EGR rate decrease the intake temperature, thus reducing the HC formation because both liner wall and sucked charge are hotter.

On the other hand, higher EGR rate modify the equivalence ratio, being locally more poor. As a result, too high EGR degrees enhance overmixing phenomena and HC formation.

1.5.4 Carbon Monoxide - CO

Carbon Monoxide is an extremely dangerous pollutant for human health, since it tends to combine with the hemoglobin present in blood, preventing the oxygen to reach the proper tissues.

In CI engines, its formation occurs during the oxidation process of hydrocarbons, that is performed through different steps. The molecular is modified until the combustion products CO_2 and water vapor are obtained. The oxidation reaction is:



At high temperatures which are reached during the combustion process ($T > 2800 \text{ K}$), reaction rates are high enough to reach a chemical equilibrium. Therefore, even if a sufficient amount of oxygen is present to fully oxidize CO to CO_2 , significant CO concentrations are to be expected, due to the reverse reaction that leads to the dissociation of CO_2 into CO and OH.

However, the sudden temperature decrease that occurs during expansion phase freezes the chemical reactions, preventing further CO_2 oxidation to CO.

Since the reaction with CO_2 as a product is enhanced by the excess of oxygen, the CO emission for a CI engine are quite negligible whether they operate with highly-poor mixture, that is, with a high overall equivalence ratio.

The engine load represents the most crucial parameter since higher fuel quantity requires to be mixed with air. In conclusion, all the rich operating condition, such as cold start, transients and full load, cause significant CO emission.

1.6 Injection Technologies

The most widely used technology about injection for Diesel combustion is the *Common Rail* (CR). This system, represented in Figure 1.43 [3], is made possible by separating the functions of pressure generation and fuel injection. The Common Rail offers a significantly higher level of adaptability to engine design.

The low-pressure system comprises the components of the fuel-supply system; the high-pressure system is composed of high-pressure fuel line.

About common rail injectors, the first thing to be considered is the dual typology:

- Indirect Acting Injectors,
- Direct Acting Injectors.

The difference stands in the command given to lift the main needle. In Direct Acting the command is electrical and directly opens the main needle without using a pilot valve. A further classification is provided depending on how the electrical command is delivered, through:

- a solenoid; the electromagnetic field moves the valve,
- a piezo-actuator; a stack moves the valve/needle.

Practically, an indirect injector may use both technologies, whereas a direct injector never uses solenoidal layout, otherwise the solenoid required to move the main needle would be too big. Considering the cost, the cheapest solution is the solenoid injector while the most expensive is the direct acting, piezo injector. However, performances follow the cost direction.

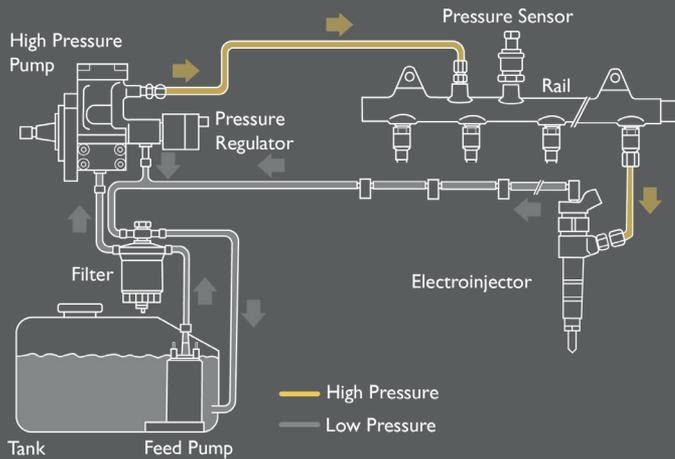


Figure 1.43 - Scheme of a Common Rail (CR) system

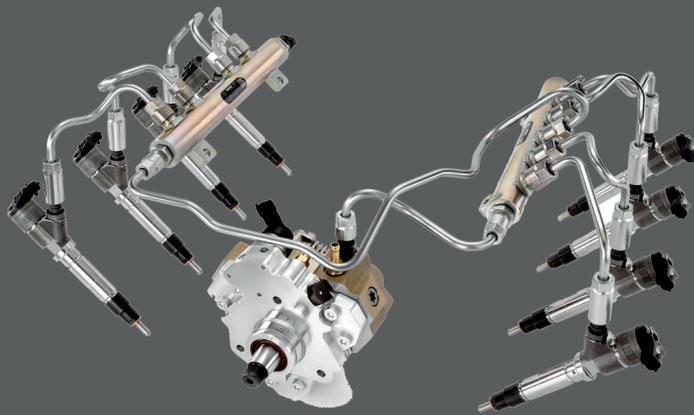


Figure 1.44 - High pressure circuit of a Common Rail system

1.6.1 Solenoid CR Injectors

In general, the layout of a solenoid injector for CR systems is represented in Figure 1.46 [1]. The fuel is fed from the high-pressure connection (4), to the nozzle through the passage (10), and to the control chamber (8) through the feed orifice (7). The control chamber is connected to the fuel return (1) via bleed orifice (6) which is opened by the solenoid valve. With the bleed orifice closed, the

hydraulic force applied to the valve control plunger (9) exceeds that at the nozzle-needle pressure shoulder (11). As a result, the needle is forced into its seat and seals off the high-pressure passage from the combustion chamber.

When the fuel enters the injector, there are two possible directions. The first is along the feed pipe to the sac before being injected across the injector holes. The pressure difference between the sac and the environment is the injector pressure drop Δp .

When, on the contrary, the needle is closed, the pressure upstream the sac, that is the feed pipe, is equal to the rail pressure. Then, if the needle is closed, the fuel traces another path reaching the control chamber. The Z-hole is the gate in which the fuel pass to reach the control chamber.

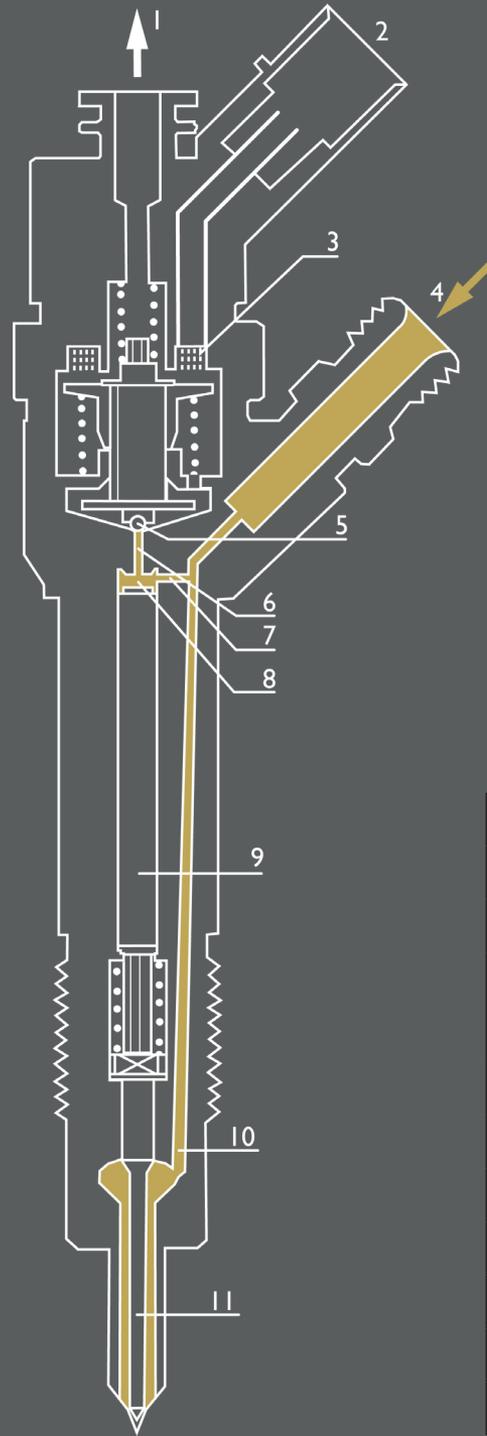
About the opening of the injector, the bleed orifice must be open to lift the needle. In this way, the pilot valve is energized and the fuel in the control chamber returns back into the tank. By opening the bleed orifice (A-hole), the pressure level in the control chamber drops down to a lower value.

This is made possible by the Z-hole that decouples the environments so that only part of the fuel goes actually into the pilot valve. In this way, the control chamber experiences a lower pressure.

Consequently, the pressure acting on the stem, is no more the rail pressure but instead is the control chamber pressure which is lower. This pressure drop determines the lift of the needle. Still, the opening is not immediate because the fuel requires time to enter the control chamber and set the desired pressure level.



Figure I.45 - CR Solenoid Injector



Injector (schematic)	
1	Fuel return
2	Electrical connection
3	Triggering element
4	Fuel inlet
5	Valve ball
6	Bleed orifice (A)
7	Feed orifice (Z)
8	Control chamber
9	Valve control plunger
10	Feed passage to nozzle
11	Nozzle needle

Figure I.46 - Schematic section of the CR Injector

1.6.2 NOD, NCD, Hydraulic Duration of Injection

Once the solenoid system for Diesel engines has been reviewed, it is worth mentioning how an injection system works in terms of signals.

Figure 1.47 [3] shows the duration of the electric and hydraulic signals in an experimental injection. Note that the hydraulic duration of the injection is also named Injection Temporary Length (ITL).

First evidence is that, in this specific case, the electric signal is shorter than the hydraulic duration. Because of this difference, one must consider two delays:

- Nozzle Opening Delay - NOD

the time interval between the start of the electric command input and actual instant at which fuel injection begins,

- Nozzle Closure Delay - NCD

time interval between the electric current shut-off and the instant at which nozzle actually closes.

Since usually $NCD > NOD$, this determines that the hydraulic duration and the electric command are different. The duration of the electric command is named energizing time ET and so:

$$ITL = ET + NCD - NOD$$

More in detail, consider the two plots in Figure 1.48 [3]. The top graph evaluates the trends of the pressures in both control and delivery chambers, as well as the lifts of pilot valve and needle, as function of time.

Starting from the lifts, the pilot valve opens once received the electric command and closes when such signal is stopped.

The *lift of the pilot valve* l_{pv} is relatively small. The main *needle lift* l_n indeed, is triggered by the previously describe process.

From the electric command, which marks the start of the pilot valve motion, to the time in which the main needle starts moving there is a time delay composed of two contributions:

- Δt_1 , control chamber pressure p_{cc} discharge due to pilot valve opening,
- Δt_2 , time needed to recover the nozzle axial deformation.

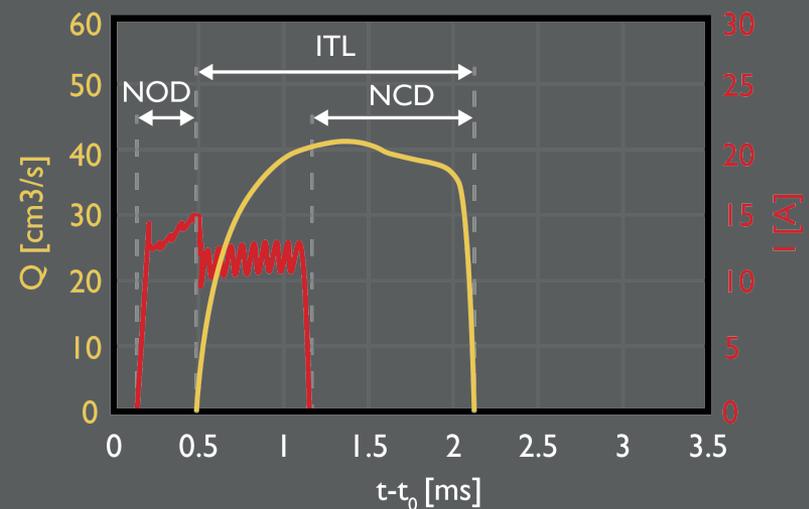


Figure 1.47 - Hydraulic and electrical signals in an injection event

Note that during the p_{cc} discharge process, the velocity of the needle remains nil. However, before the needle actually starts moving out of the seat, it requires a Δt_2 delay to counterbalance the elastic deformation. That is why the relative displacement of the needle takes places after a while even if the velocity is different from zero. For this reason the total Nozzle Opening Delay is defined as:

$$\text{NOD} = \Delta t_1 + \Delta t_2$$

Still considering Figure 1.48, the additional represented pressure is the *pressure of the delivery chamber* p_{dc} . This pressure is actually causing the pressure drop along the nozzle holes and is produced by the upstream rail pressure. As the needle opens, the delivery pressure tends to decrease.

Later on during the injection, the pressure in the control chamber p_{cc} increases even though the pilot valve is open, this because of the upwards travel (upstroke) of the needle causes the scavenge of the control chamber. The counterpressure to control the needle velocity is performed by the bleed orifice (A-hole), that therefore must be sufficiently large to make the control chamber empty quickly, but at the same time, it must not be too large otherwise the velocity of the needle would be too high.

Also about closing, the delay contributions are two:

- Δt_3 , time required to slow down and invert the needle motion,
- Δt_4 , time needed by the needle to reach the final closure position.

When you want to conclude the injection, the electric command is stopped, so that the pilot valve starts clo-

sing through a return spring. The pressure inside control chamber is increasing to the higher level and such pressure is actually stopping the needle motion.

Note as the needle slows down, it remains open, or at least not sealed into the seat. Thus, the pressure of the control chamber is also tasked to make the needle travel back at constant velocity until it reaches the closure position. The longer the energizing time ET, the longer the delay Δt_4 to close the needle. The Needle Closure Delay is then defined as:

$$\text{NCD} = \Delta t_3 + \Delta t_4$$

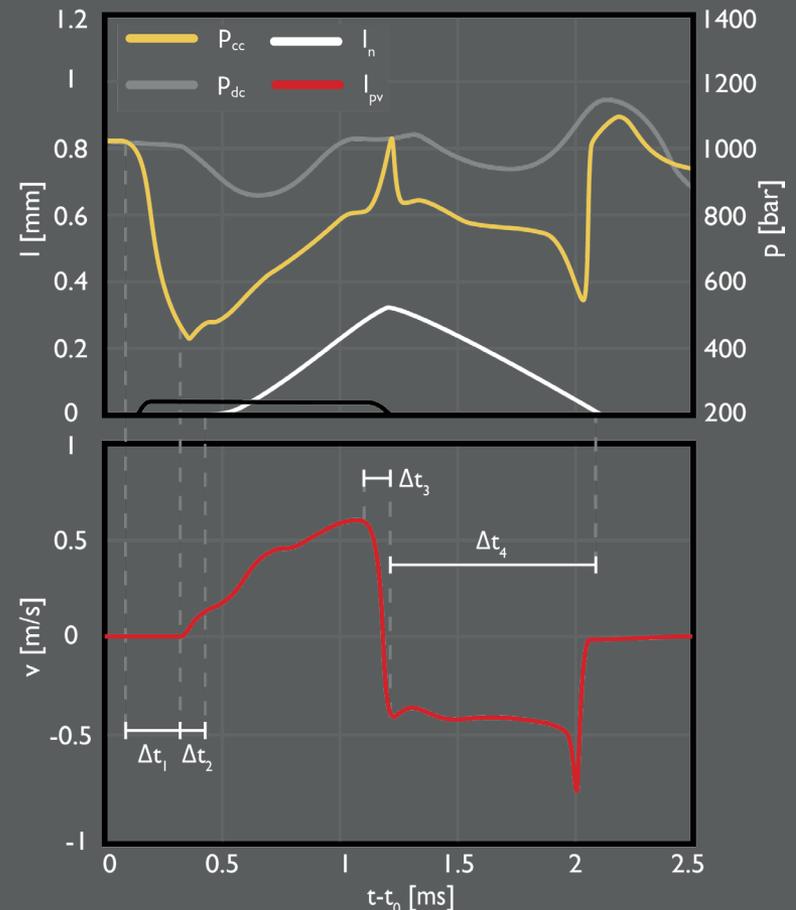


Figure 1.48 - Delay contributions in an injection event

In indirect injectors, the nozzle closure delay is always bigger than the opening delay. On the contrary, in case of direct injector the two delays may be very similar, so that the hydraulic and electric durations almost coincide.

In the following, Figures 1.49 [3] and 1.50 [3] highlights the differences between NOD and NCD for indirect injectors.

First, note that the range of the vertical axis is very different since the NOD ranges between 350 to 430 μs whereas the NCD goes from 300 up to 1500 μs . The dominant value is the closure delay because of the higher scale of y-axis.

In particular, the NCD relies on the operating condition such as the energizing time ET and the rail pressure p_{rail} while, as second approximation, the NOD is less affected by these quantities.

As stated, the trend of NCD is determined by the duration of ET which drive the needle motion, and the higher the lift, the longer the time required for closure.

However, beyond a certain energizing time, the closure delay is not affected anymore owing to the mechanical end-stroke of the lift. Fixing the energizing time ET, the higher the rail pressure p_{rail} , the higher the nozzle closure delay NCD.

By the way, the rail pressure also influences the opening force of the needle so higher rail pressure leads higher momentum transferred to the needle meaning higher lift.

Indeed, in the end part of the graph, the trend of NCD is reversed since with higher rail pressure, and with the needle at end-stroke, the pressure in the control chamber is restored quicker thus reducing NCD.

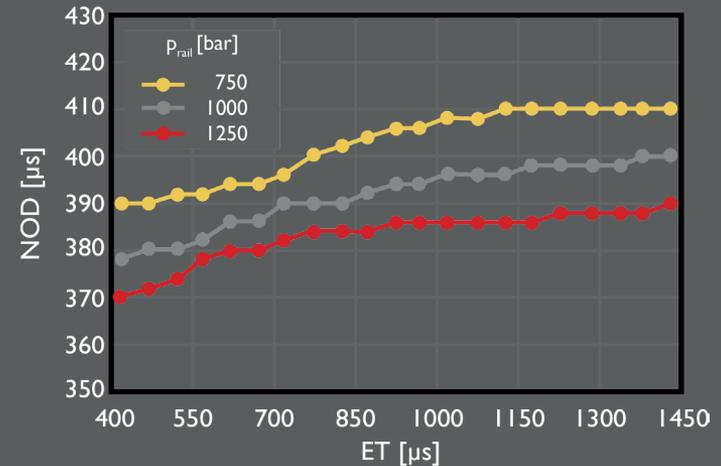


Figure 1.49 - NOD as function of ET for different rail pressures

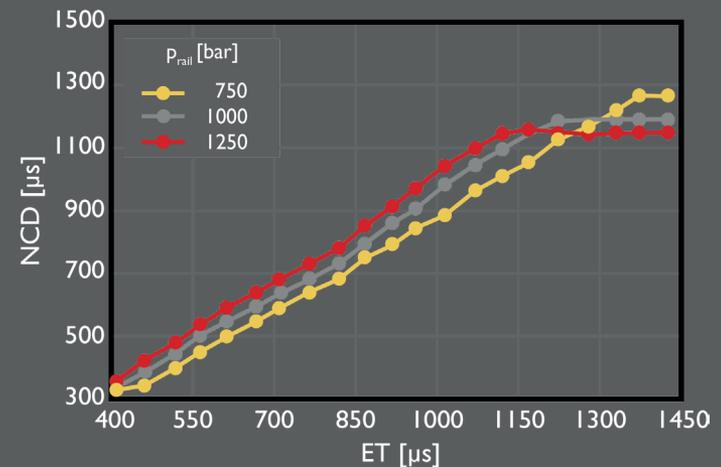


Figure 1.50 - NCD as function of ET for different rail pressures

As seen, the hydraulic duration ITL is defined as:

$$ITL = ET + NCD - NOD$$

Plotting ITL as function of ET, one obtains the diagram of Figure 1.51 [3]. The change of slope is owing to the end-stroke of the needle in the upward travel. However, nowadays, the needle never reaches the end-stroke (ECU strategy does not permit this to happen).

The ECU, that supplies the current to match the desired energizing time, is provided of a torque-based control. This is a linear torque-to-fuel map that directly determine the hydraulic duration of injection and the suitable energizing time, thus avoiding the really end part of the ITL vs. ET diagram.

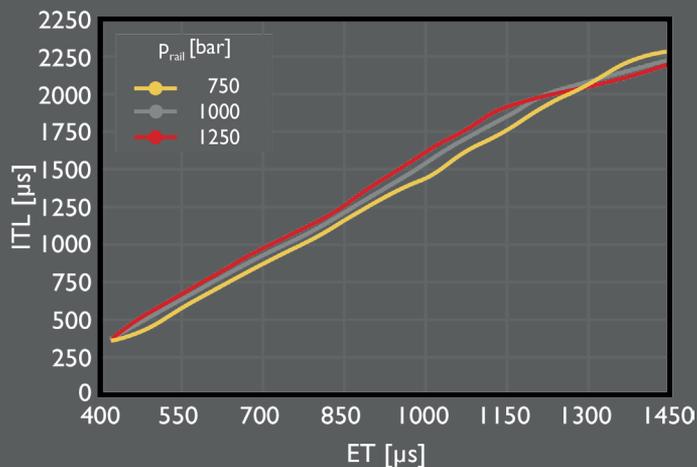


Figure 1.51 - ITL vs ET trend for different rail pressures

Furthermore, it is possible to state that also the Nozzle Opening Delay is slightly dependent on the energizing time and rail pressure. As far as the rail pressure is concerned, by increasing the pressure, the NOD decreases due to the time needed to refill the pressure within the

control chamber. The higher the pressure, the faster the velocity through the bleed orifice. About the energizing time, such energy does not produce work and the correspondent effect is heat dissipation which increases the fluid temperature resulting in different velocities of sound. Since all the phenomena previously mentioned depend on wave propagation, then also rate of pressure is embraced. This is reason why also the NOD tends to slightly increase accordingly with ET. To sum up, Figure 1.52 and the correspondent table, summarize the effect of the rail pressure and the energizing time over the NOD and the NCD.

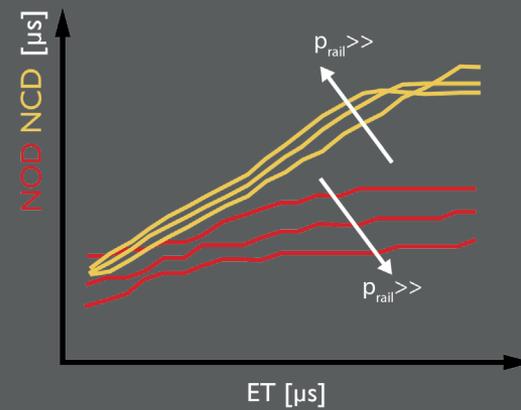


Figure 1.52 - NOD/NCD trend vs ET as function of the rail pressure

	NOD	NCD
P_{rail}	$P_{rail} \uparrow$ NOD \downarrow Higher P_{rail} means higher velocity in the bleed orifice speeding up the opening	$P_{rail} \uparrow$ NCD \uparrow Higher P_{rail} leads higher momentum to the needle, thus higher lift
ET	$ET \uparrow$ NOD \uparrow Higher energy brings more heat to the fluid, changing wave propagation laws	$ET \uparrow$ NCD \uparrow The longer the electrical command, the longer the lift. The limit is the end-stroke

Table 1.3 - NOD/NCD qualitative trend vs ET and P_{rail}

1.6.3 Multiple injection events

All the previous considerations about injection were related to a single-event injection, that is, a single electric command which causes a single injection. However, in a Common Rail system, usually the strategy provides the so called *multiple-event injection*. The fuel delivery does not occur in a single event, but it can be delivered according to different pulses in the same cycle. A schematic example of multiple injection is portrayed in Figure 1.53. Today, current Common Rail injector for passenger car application might have up to 9 pulses per each cycle.

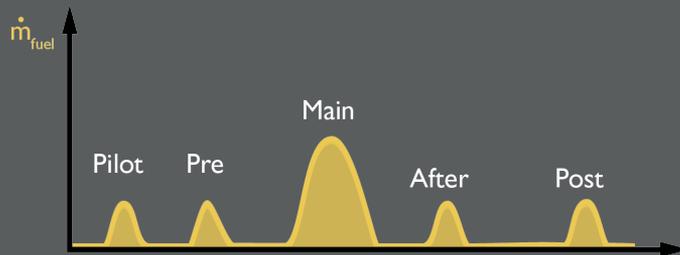


Figure 1.53 - Multiple Injection Strategy scheme

The reason why this last solution is widely adopted is because it may be adapted to a broad engine requirements such as:

- torque, since it is proportional to the fuel flow rate,
- drivability, concerning combustion noise control,
- emission control, to manage NO_x and soot emissions,
- ATS, to oxidize soot and regenerate DPF.

The Common Rail injection system should operate correctly both for single-event and also for multiple-event injections.

In multiple injection events the pilot valve opening and closure determine depressure waves travelling back and forward. Other fluctuation are generated by wave reflections which keep on travelling through the pipe.

Since the electric command does not unchanged but rather the SOI does, actually the variable parameter is the so-called *Dwell Time* (DT). Depending on the ECU, the DT may be defined as the time between the start of the first pulse and the start of the second pulse, otherwise as the time between the end of the first pulse and the start of the second pulse:

$$DT = \frac{SOI_{pil} - SOI_{main} - ET_{pil}}{n}$$

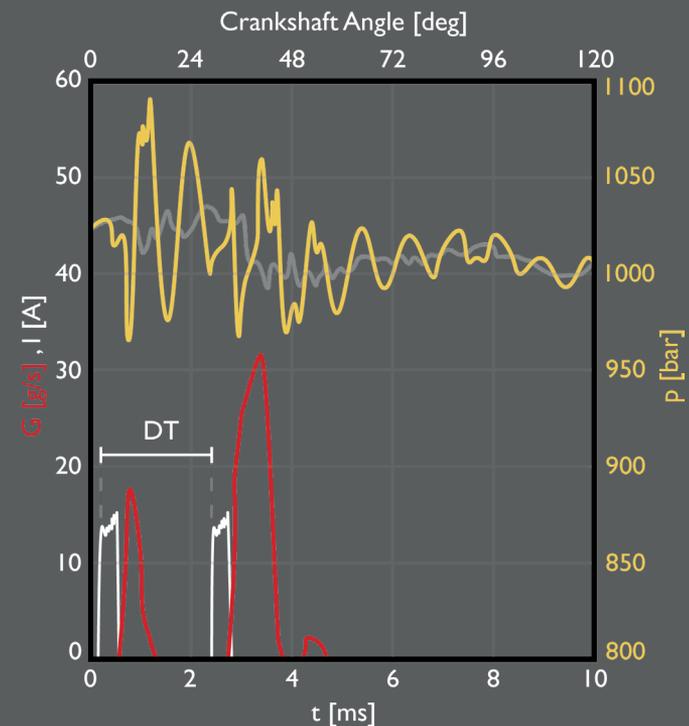


Figure 1.54 - Dwell Time for a Pre and Main injection strategy

Consider an example in which only the DT is changed (same SOI and ET). Despite the expectations would suggest that the two systems share the same mass of fuel injected, actually this is not verified. Figure 1.55 [3] indeed plots the percentage deviation of the injected mass volume from its mean value as function of the Dwell time.

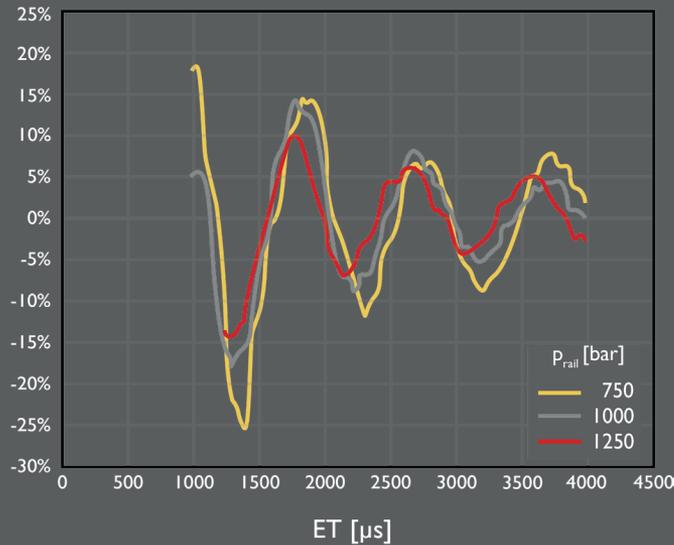


Figure 1.55 - Fuel mass percentage deviation for a DT variation

As seen, by only changing the Dwell time, the quantity of fuel injected can be increased up to 20%, which is definitely not negligible.

Also the fluctuation is dependent on the Dwell time. This may be an issue whether a specific operating point needs to raise the Dwell time, but the higher the injected fuel, the higher the produced torque, which may be not desired.

However, the main injection may start or in a peak pressure, or in the contrary during a minimum in the pressure.

The pressure wave generated by the pilot affects the opening force of the needle in the main injection.

If the pressure wave is high, it transfers high momentum to the needle causing higher lift. Consequently, the nozzle closure delay NCD is longer as well as the hydraulic duration of injection ITL, still considering fixed energizing time ET.

In case that the injection event occurs in a minimum pressure wave, the injected fuel mass would be lower because lower pressure means lower lift, shorter closure delay and thus shorter hydraulic injection. This adverse reaction is of course unfavorable.

By exceeding with the Dwell time reduction, then the two pulses become a sort of a unique pulse. This effect is named *Injection Fusion*. To understand the condition for this to happen, consider Figure 1.56 which reports the limit case of the phenomenon.

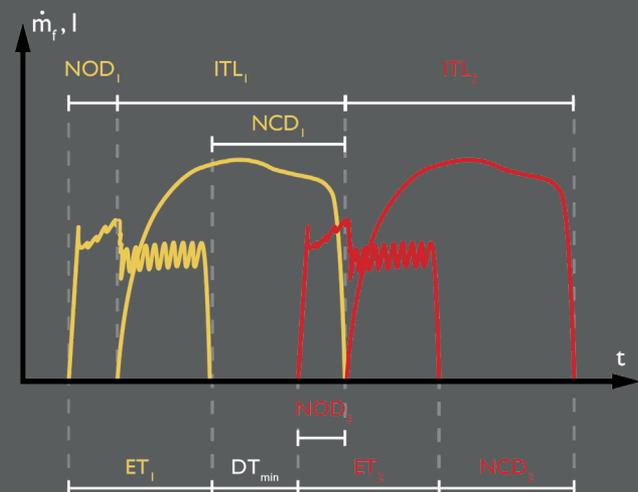


Figure 1.56 - Scheme of an Injection Fusion event

Still considering Figure 1.56, it is clear that the minimum Dwell time DT_{\min} is the minimum time for which the nozzle sprays two different pulses. Basing on the definition of hydraulic duration of injection:

$$ITL_1 = ET_1 + NCD_1 - NOD_1$$

then, the geometric criteria to set the minimum Dwell time is:

$$ET_1 + DT_{\min} + NOD_2 = NOD_1 + ITL_1$$

$$ET_1 + DT_{\min} + NOD_2 = NOD_1 + ET_1 + NCD_1 - NOD_1$$

$$DT_{\min} = NCD_1 - NOD_2$$

To obtain the values of the delays of the first pulse, one has to investigate the diagram concerning the delay compared to the energizing time (Figures 1.49 and 1.50). However, for the second pulse, the diagram for the single injection is no more valid. The difference stands in the pressure level of the control chamber that may be under transient conditions and consequently the needle is not fully closed. Therefore, in injection fusion the needle does not close before starting the second pulse, as shown in Figure 1.57 [3]. Note that the second peak of the pressure in the control chamber does not reach the pressure level as it was stand-alone, but at higher level.

For this reason, the nozzle opening delay of the second pulse is shorter and for sake of clarity, this new value of the nozzle opening delay is named Nozzle Re-opening Delay (NRD_2). From experience, NRD_2 may also be half compared to NOD_2 .

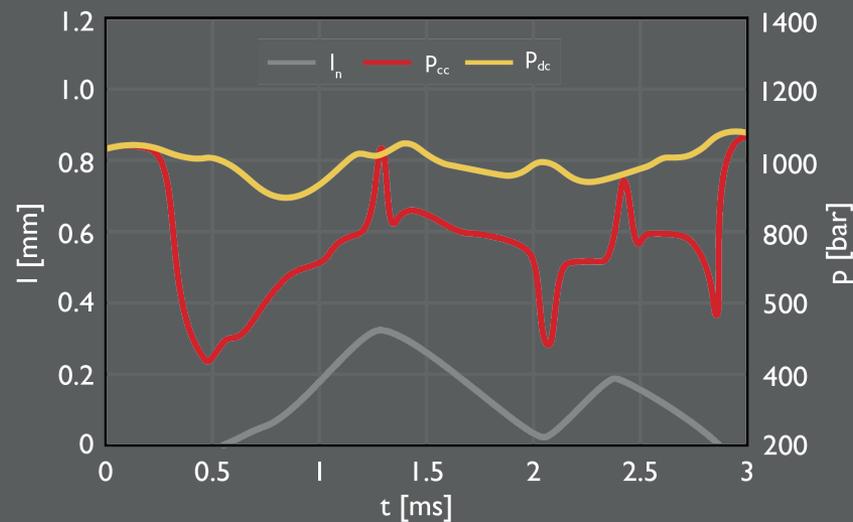


Figure 1.57 - Needle lift for a second pulse in multiple injection event

2

FPT

F1A

CI ENGINE

Engine Characteristics

2.1 Engine Characteristics

The engine of studies is the FPT FIA and it has been provided by FCA. It is a four cylinder in-line, turbocharged, PCCI engine with dual loop EGR circuit. This unit also complies with Emission Regulation EU6b.

The following table shows the datasheet of the engine.

Manufacturer	FPT Industrial
Engine Family	FIA
Application	Light-Duty
Displacement	2.3 L
Cylinders	4 In-Line
Bore	88 mm
Stroke	94 mm
Con-rod Length	146 mm
Compression Ratio	16.2:1
Dead Volume Height	0.7 mm
Injection System	Common Rail
Turbocharging	VGT
EGR System	LP + HP
Max Power	102 kW @ 3250 rpm
Max Torque	350 Nm @ 1750 rpm
ATS	DOC + DPF

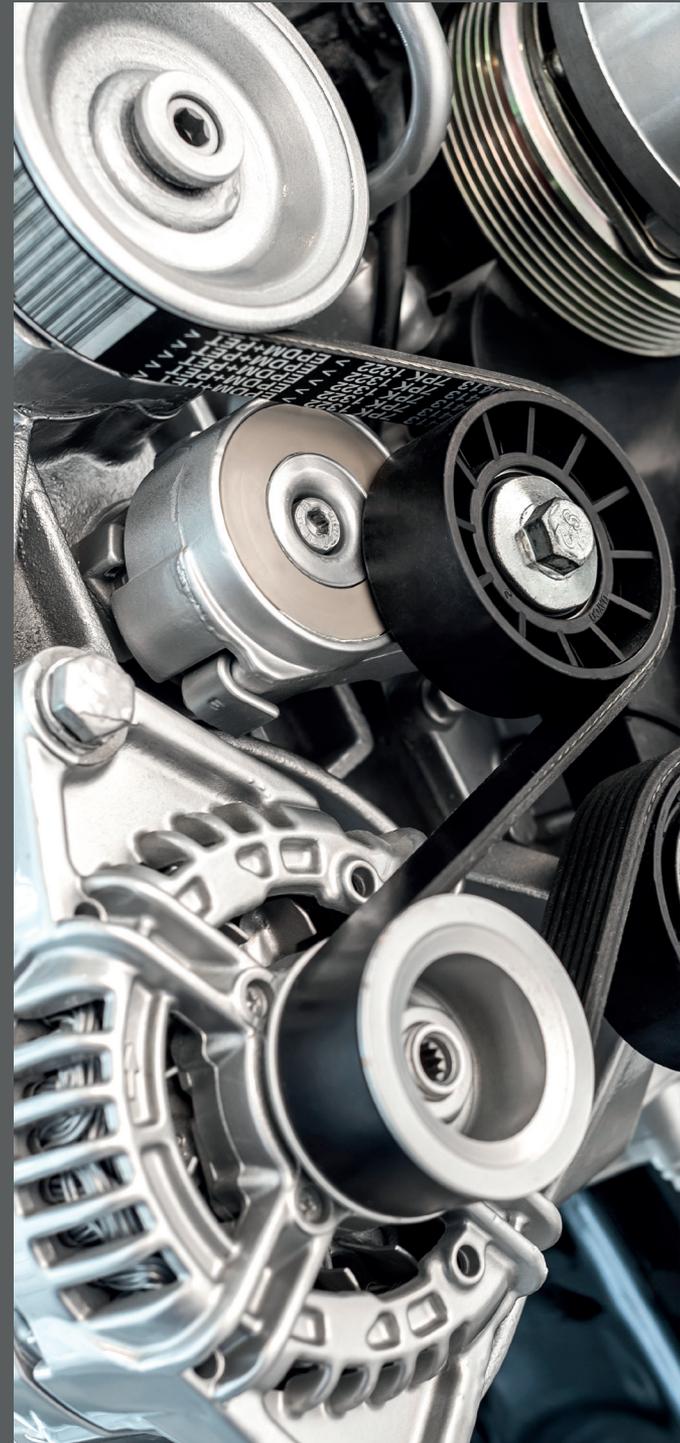


Table 2.1 - FPT FIA Engine data sheet

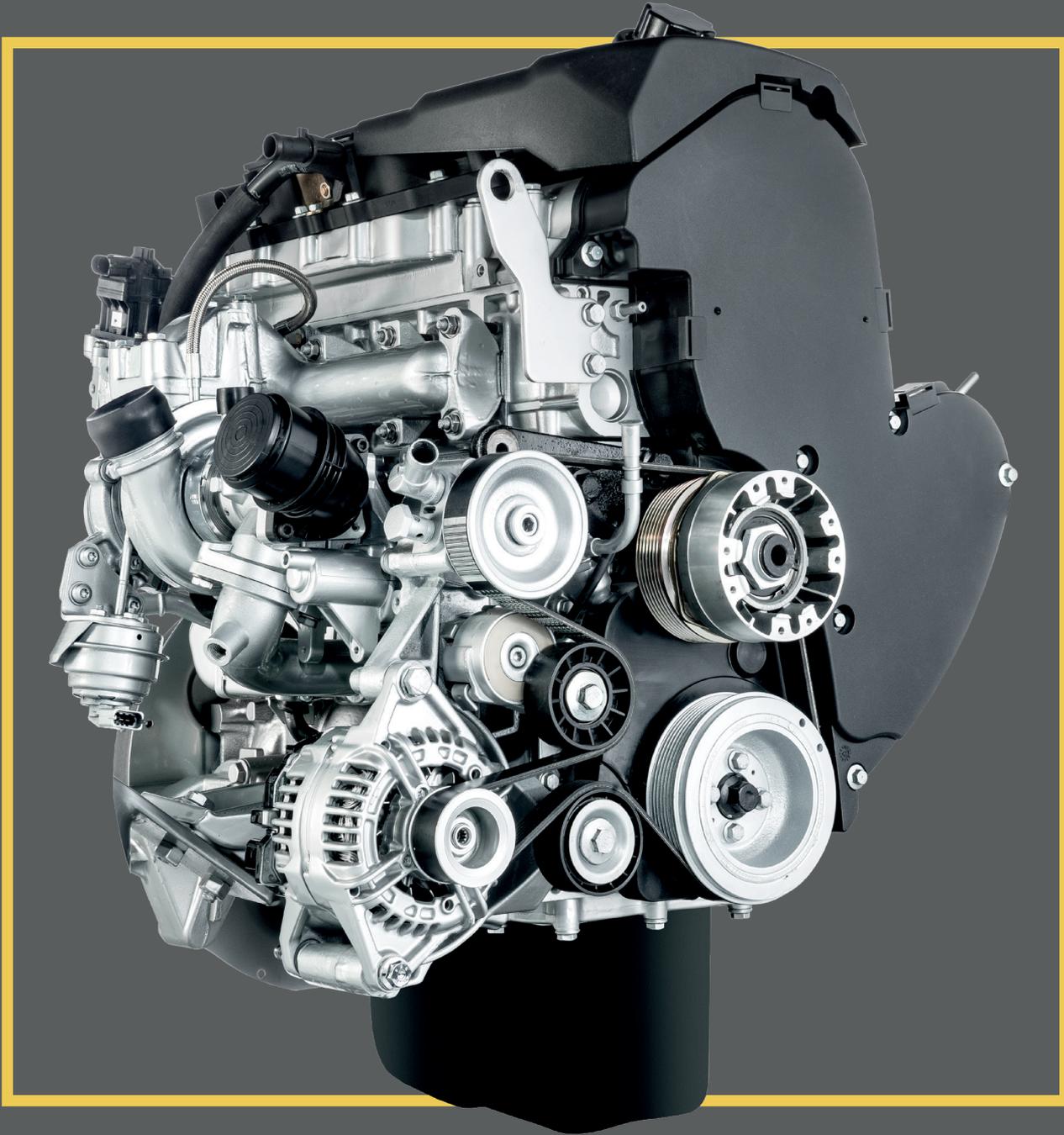


Figure 2.1 - FPT Industrial FIA Diesel Engine ⁽³⁾

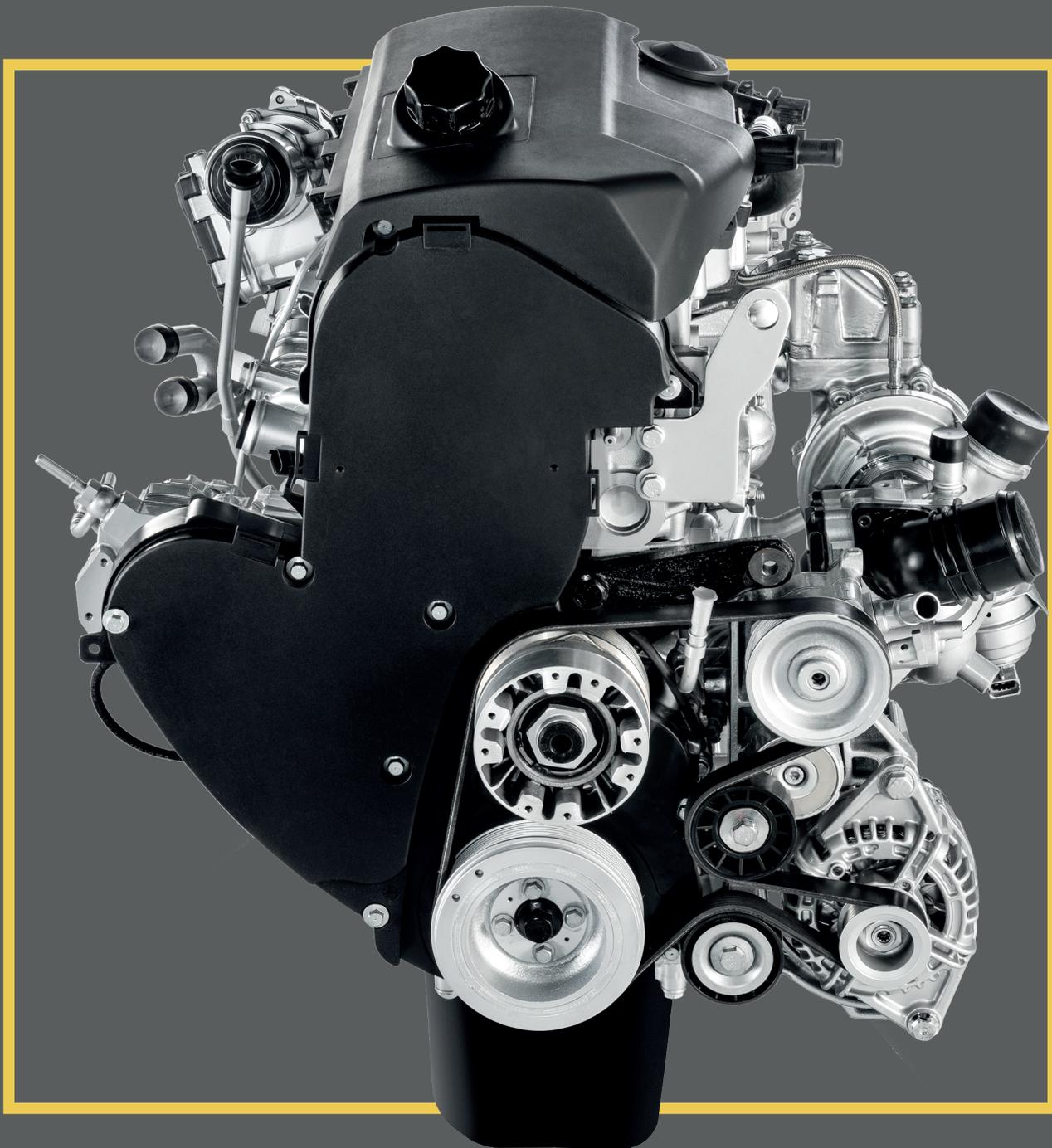


Figure 2.2 - Front view of the FPT Industrial FIA Diesel Engine ⁽³⁾

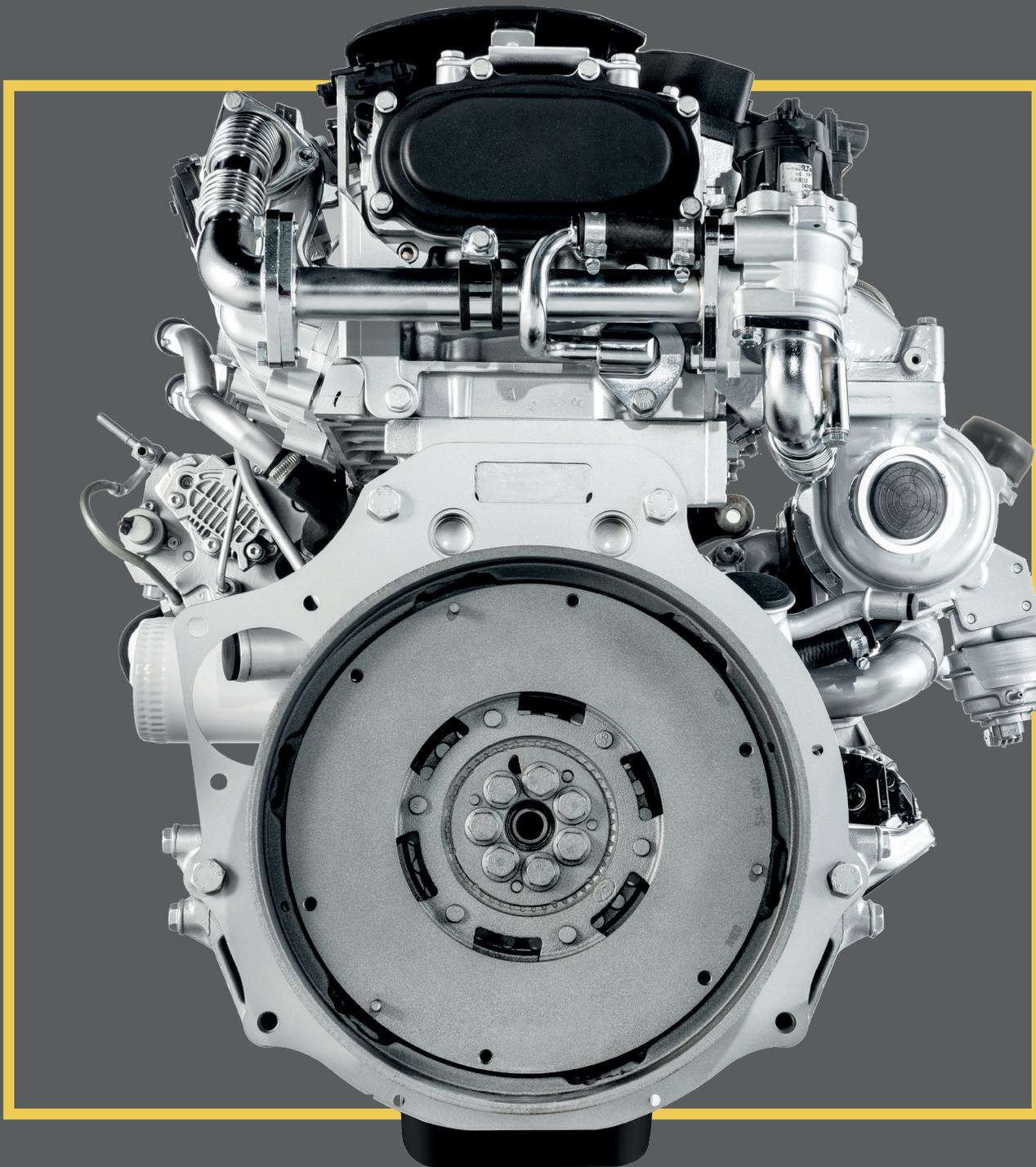


Figure 2.3 - Back view of the FPT Industrial FIA Diesel Engine ⁽³⁾



Figure 2.4 - Left side of the FPT Industrial FIA Diesel Engine ⁽³⁾

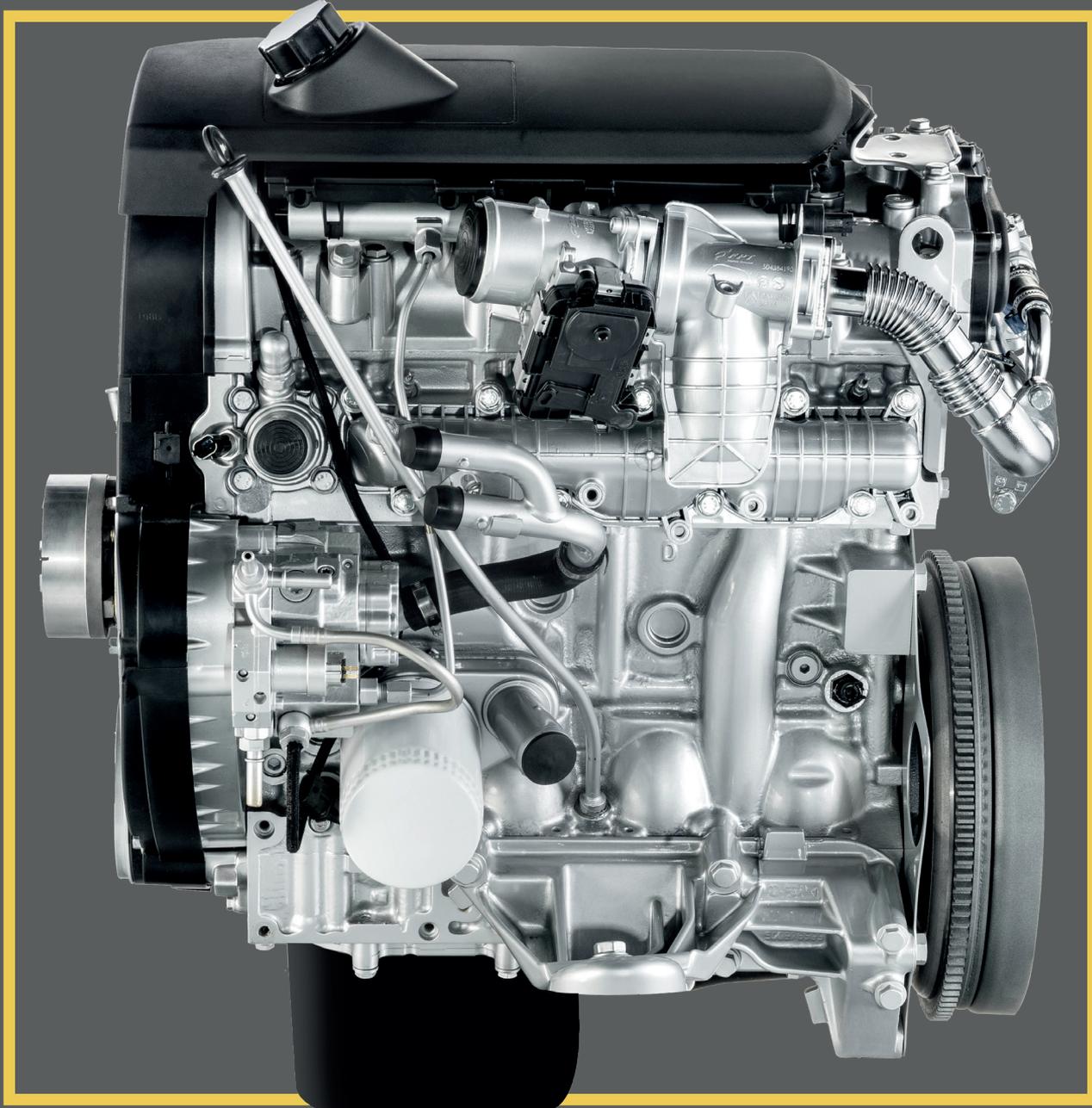


Figure 2.5 - Right side of the FPT Industrial FIA Diesel Engine ⁽³⁾



Figure 2.6 - Front Left view of the FPT Industrial FIA Diesel Engine ⁽³⁾



Figure 2.7 - Front Right view of the FPT Industrial FIA Diesel Engine ^[3]

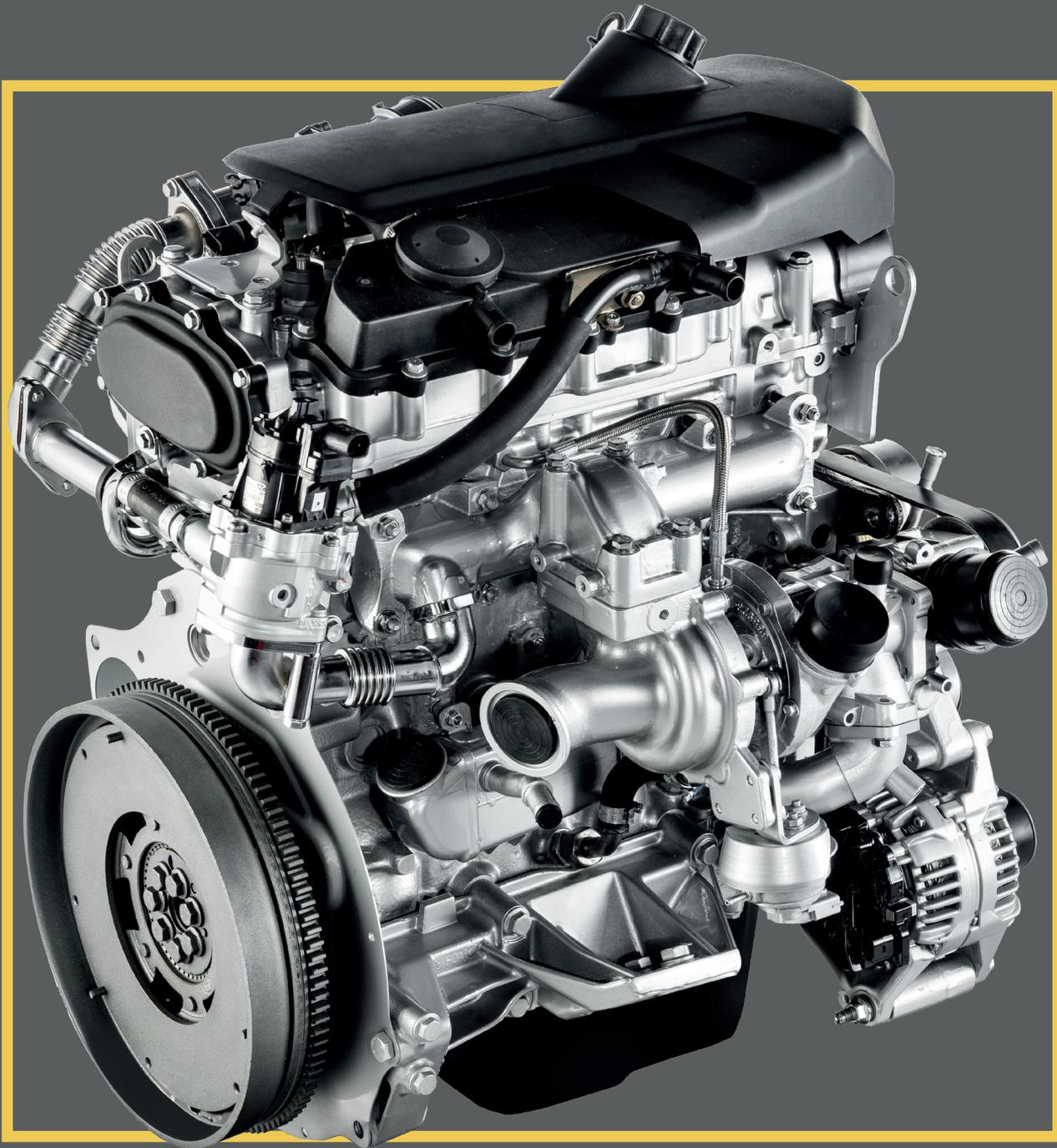


Figure 2.8 - Left Back view of the FPT Industrial FIA Diesel Engine ^{3}



Figure 2.9 - Right Back view of the FPT Industrial FIA Diesel Engine ⁽³⁾



***GT POWER
SIMULATION
SOFTWARE***

GT-Power Software

Combustion Models

DI-Pulse Predictive Combustion Model

Burn Rate Analysis

The simulation models represent a very useful tool to emulate the engine working conditions. Their use is of paramount importance whether the engine is not physically accessible or because performing real simulations is costly. Once the engine model has been built up, it is possible to study the behavior and the influence of the parameter of interest. On the other hand, the time and effort to obtain a model, as well as the simulation time, may be too long. Also consider that the results coming from a virtual simulation are different from that obtained through a real simulation in a test bed; still, the calibration and validation of such models return comparable results.

The two main advantages in using virtual simulations are:

- to save time and money, since it is possible to quickly realize cycles and reduce the number of prototypes,
- to overcome the limits of a traditional models, since virtual environments allow to simulate a high range of different operating conditions.

The simulating models are essentially divided between: mono-dimensional (1-D) and three-dimensional (3-D) models, also known as CFD models (Computational Fluid Dynamics). Both categories are able to solve partial differential equations that rule the fluid motion. CFD models are for sure more reliable and complete, but they require high computational time. For this reason, they are usually involved to study single components.

On the other hand, the computational time of 1-D and 0-D models is much shorter, and the results are reliable as well. Because of these advantages, 1-D models are used to study the overall engine system.

3.1 GT-Power Software

GT-SUITE is a multi-physics Computer-Aided Engineering (CAE) simulation software developed by Gamma Technologies, used for full vehicle design and analysis. The library developed for engine simulation is GT-Power.



Figure 3.1 - Gamma Technologies logo ⁽⁴⁾

The main software is GT-SUITE which is composed by several operative components:

- GT-ISE, the graphical interface to get the database of the engine parts and connectors needed by the user to actually build up the engine model,
- GT-Solver, the tool that performs calculations to solve the equations,
- GT-Post, the software that analyze the results from any GT-Power simulation, both numerical and graphical, including plots and maps.

Thus, the working area is represented by the so-called *project map* in which all the objects (pipes, cylinders, turbocharges, connectors...) obtained in the library are placed so as to build the engine model.

After the model has been placed onto the map, it is required to define the properties of each object (thermal, chemical, geometrical) and set the desired outputs. In this way, one can reproduce very accurate engine models very close to reality.

As stated, GT-Power is the library for engine application. It is used to provide both stationary and transient simulations suitable for engine control and powertrain applications because it can run any engine typology. Through GT-Power one can:

- analyze and predict engine performances (fuel metering, mean effective pressure - MEP, indicated mean effective pressure - IMEP, fuel consumption...),
- estimate the influence of the engine geometrical parameters over its performance,
- model the combustion process and evaluate the heat transfer, heat release, turbulence, fuel evaporation and pollutants emission,
- measure quantities which are tough to be analyzed in real applications, such as instantaneous mass flow rate through intake valves,
- test pressure, temperature and flow rate in different part of the engine,
- get the best trade-off between complexity and accuracy of the model considering the available input data and the desired output data,

GT-Power solver is based on one-dimensional solution of the fully unsteady, nonlinear Navier-Stokes equations. It employs thermodynamic, chemical and phenomenological model solvers to simulate the behavior of internal combustion engines .

The working principle of GT-Power is to divide the overall system into smaller sections, each of them with specific boundary conditions.

The solver computes the following equation in each time-instant:

- Mass Conservation Equation,

$$\frac{dm}{dt} = \sum_{\text{edges}} \dot{m} = \sum_i \dot{m}_i - \sum_o \dot{m}_o$$

- Energy Conservation Equation, in which e is the total internal energy per unit mass, p and V are pressure and volume, H is the total enthalpy, h is the heat transfer coefficient, T is the temperature, A_s is the heat transfer surface

$$\frac{dt(\dot{m}e)}{dt} = -p \frac{dV}{dt} + \sum_{\text{edges}} (\dot{m}H) - hA_s (T_{\text{fluid}} - T_{\text{wall}})$$

- Momentum Conservation Equation, in which A is the area of the transversal flux. C_f is the friction coefficient, K_p is the coefficient of the pressure drop, u is the edge velocity, dx is the length of the mass element in the fluid section, dp is the differential pressure acting along dx , while ρ is the fluid density.

$$\frac{d\dot{m}}{dt} = \frac{dpA + \sum_{\text{edges}} (\dot{m}u) - 4C_f \frac{\rho_u |u| dx A}{2D} - K_p (\frac{1}{2} \rho_u |u|) A}{dx}$$

The mass and energy conservation equations are solved in sub-volumes through scalar variables, such as the temperature, the pressure and the species concentration, that are computed in the center of each sub-volume and considered constant throughout its volume. On the contrary, the momentum conservation energy is solved along the edges through variable vectors, such as speed and mass flux.

The scheme of Figure 3.2 [16] represents the working principle of GT-Power. The obtained solution does not rely upon an iterative numerical process, as for CFD simulations, but rather it is based on the initial state of the system (time t_0) and computed for a new time instant, which must be as close as possible to instant t_0 to ensure the solution feasibility. Such maximum time step is evaluated per each step, always.

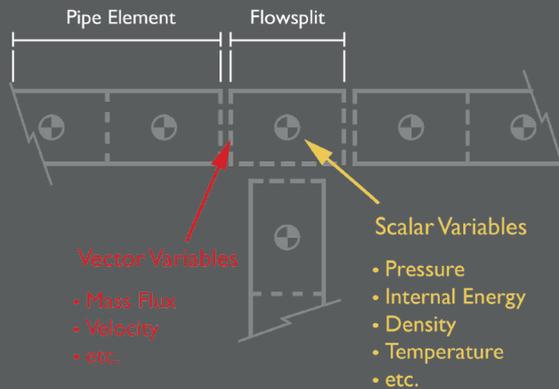


Figure 3.2 - GT-Power working principle

The calculus procedure is then based upon the idea that the solution in a specific instant is computed on the values of the centroid and quantified in the previous time instant. The time difference between these two instant is named *timestep*.

However, all numerical computations about fluid-dynamics simulations in GT-Power must comply to the Courant condition, that forces the discretization length and the timestep.

The Courant-Friedrichs-Lewy condition (CFL) states that, in order to converge a numerical explicit problem, the Courant number C must be lower to the maximum value C_{max} in each point.

Whether the numerical scheme is explicit, then such maximum value is at most unitary. The formula for the Courant number C is as follows:

$$C = \frac{u \cdot \Delta t}{\Delta x} < C_{max}$$

in which u is the local speed of the fluid particle, and Δt is the timestep.

At operational level, this rule affects the discretization length and the timestep in the following way: during the development of model, it is required to set the discretization length per each channel, and while the simulation is running, GT-Power defines its timestep.

The channel which discretization length is the shortest are those that actually set the timestep to satisfy the CFL condition.

For this reason, it is crucial that the model is well optimized so as to avoid bottlenecks along the circuit that may slow down the calculations.

To get an idea of the discretization effect over the calculation time, consider that, through an explicit method, if the discretization length is halved, then the computation time is extended up to four times longer; indeed the timestep is halved as well to comply the Courant condition, causing the doubling of the number of sub-volumes.

3.2 Combustion Models

GT-Power provides several combustion model depending on the engine (SI or CI) and on the simulation typology.

Before discussing the models, it is preferable to define the quantities that recur most frequently:

- Combustion

It represents the fuel transfer from the unburned to the burn region. The combustion provides the release of the chemical energy and the species calculation that form the exhaust gases.

- Burn Rate

It is the instantaneous velocity of the air and gas molecules at which they become combustion products. When the combustion provides a diffusive flame, the burn rate is the velocity of an infinitesimal volume dV at which it moves from the unburned to the burned gas. In such transition, the composition of the sample volume changes from a rich mixture to combustion products and air. Since the combustion requires a certain amount of time to complete, the burn rate never matches the heat release rate trend. Indeed, this last actually develops only when the pre-reactions before the combustion are completed. The burn rate is an input that may be set or requested by the user, depending on the used model.

- Heat Release Rate

It represents the instantaneous thermal energy that is released by the combustion. The HRR differs from the burn rate owing to the partial combustion. The air-fuel particles do not burn all together instantaneously, and so, part of the combustion energy is released later on. That is the reason why the HRR is delayed compared to the Burn Rate.

The HRR takes into account the energy useful for the conversion into mechanical energy, i.e. the Net Heat Release Rate (NHRR), but also the energy which is dissipated through the liner wall. The evaluation of the real HRR is complex because it is very hard to determine the instantaneous chemical compositions within the cylinder; therefore it is preferable to estimate the HRR starting from the pressure inside the cylinder. This HRR is actually the already mentioned Apparent Heat Release Rate (AHRR).

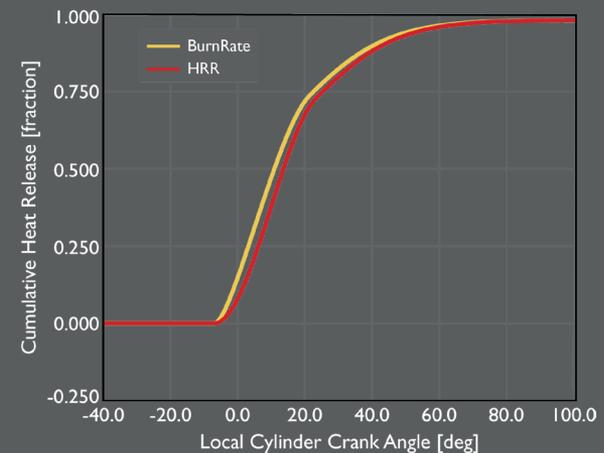


Figure 3.3 - Difference between Burn Rate and Heat Release Rate ^[16]

- Forward Run Combustion Calculation

This is the main strategy to operate the simulations in GT-Power. The Burn Rate is the input of the simulation either it is imposed or estimated, and the output is represented by the in-cylinder pressure. The air-fuel particles are transferred from the unburned zone to the burned region as specified by the Burn Rate. The in-cylinder pressure represents the energy released by the combustion.

- Reverse Run Combustion Calculation

In the Reverse Run, the input of the simulation is the in-cylinder pressure (measured experimentally in the real engine), whereas the Apparent Heat Release Rate is the output. This last, should be then used as an input to return the values of the pressure so as to exploit the Forward Run Calculation. As a matter of fact, the reverse run is used in GT-Power using the same methods as for the forward run. In each time step, the fuel quantity from the unburned to the burned region is looped multiple times until the value of the pressure matches that measured.

- Mass Fraction Burned 50% - MBF50

It is the angular position of the crankshaft at which the 50% of the fuel mass has been burned away. This value is used as an index of the combustion trend.

- Peak Firing Pressure - PFP

It is the maximum pressure achieved in the combustion process occurring the in mixing controlled phase. Together with the MFB, the PFP is also considered an index of the combustion. With the same start of injection (SOI), a lower PFP reveals a slower combustion or a lower injected fuel quantity.

The *real* Burn Rate depends upon a wide range of variables (injection timing, EGR rate, flux through the cylinder...). Consequently, a predictive physical model would be more suitable most of the times, still it is a more complex building and computational times are longer, compared to non-predictive models. Accordingly, predictive models are used whenever the variables directly affects the burn rate-

For example, predictive models are implemented to study the influence of injection strategies, injection timing, EGR effect and swirl motion. On the contrary, when the variables do not impact over the burn rate, then non-predictive models are to be preferred since they are simpler and require less effort to perform calculations. These are used mostly to estimate the influence of the cylinder geometry, to analyze the supercharging, and also in acoustic applications.

3.2.1 Non-predictive combustion models

A non-predictive combustion model requires the burn rate as function of the crank angle to be imposed as input. Thus, the simulation takes the imposed burn rate and considers that there is always a certain amount of fuel that guarantees the burn rate, regardless of the in-cylinder conditions. For this to happen, the burn rate is not affected by any factors, such as the residual fraction or the injection timing.

The air and fuel mixture simply burn at the design speed. Through this combustion model, one can study the variables that do not affect the burn rate. The most significant benefit of this approach is the very fast calculation time.

3.2.2 Semi-predictive combustion models

The semi-predictive combustion models may be feasible alternatives to predictive models. This because they are positively sensible to the variable that have a direct effect over the burn rate and supply fast response to the chan-

ges of these variables. In addition, semi-predictive models exploit the strategies of a non-predictive model, that is the burn rate is imposed to shorten the calculation time.

3.2.3 Predictive combustion models

Predictive combustion models are suitable for any kind of application; however, they are more complex and calculation time may be too long. Here, the burn rate is computed every cycle, depending on the in-cylinder conditions and because of this, usually they are substituted by other models.

The burn rate, indeed, changes accordingly to the variation of the variables of interest. These models provide more accurate results because the burn rate considers all the input data (pressure, temperature, equivalence ratio...) which are measured on the engine test bed.

The drawback lies in the calculation time due to more complex computations but also owing to the calibration with experimental data which is needed to obtain accurate results (this procedure requires a long time). The results of a predictive model concern all the performance levels of an engine, such as power, fuel consumption and pollutants emission.

3.3 DI-Pulse Predictive Combustion Model

The DI-Pulse (Direct-Injection Diesel Multi-Phase) is one of the most recent combustion model, realized by Gamma Technologies, to estimate the combustion rate and pollutants emission for direct injection Diesel engines, both with

single and multiple injection events. Definitely, this model is more convenient rather the old predictive model, named DI-Jet. Indeed, the DI-Pulse provides a faster computational time still supplying accurate results, or even better.

Actual Diesel engines are much more complex because the inputs (injection advance, EGR, pulses quantity...) affect the engine performances. Phenomenological models, such as DI-Pulse, provided an accurate calibration, are able to supply a good prediction about combustion parameters and pollutants emission, and at the same time keeping reasonable computational time.

Through DI-Pulse combustion model, it is possible to predict the fluid-dynamic evolution of the charge within the cylinder in any condition, with specific reference to the injection events, mixing, evaporation and burning of the fuel.

Furthermore, DI-Pulse is a multizone model, meaning that the injected fuel *packets* are subdivided into different zones, each of them treated as an open thermodynamic system. The evolution of any single packet is developed on its own and the predictions are performed considering the trajectory after the injection, the evaporation, the mixing with surrounded gas and the burning time.

The combustion equations are solved separately per each zone and they rely on the temperature, pressure and equivalence ratio of each zone. In this way, one obtains the best predictions.

It is worth noting that the very accurate results are possible only if the injection profiles (input of the model) are very precise too.

The model approaches the combustion by dividing the cylinder volume into three distinct thermodynamic regions, each of them with its specific temperature and concentration, as shown in Figure 3.4 [17]. The charge within the combustion chamber is divided as follows:

- Main Unburned Zone - MUZ

It contains all the air mass trapped into the cylinder once the intake valve closes,

- Spray Unburned Zone - SUZ

It includes the injected fuel and the entrained air;

- Spray Burned Zone - SBZ

It contains all the combustion products.

The main advantage of this technique is to determine the effect of the single injection pulse over the combustion.

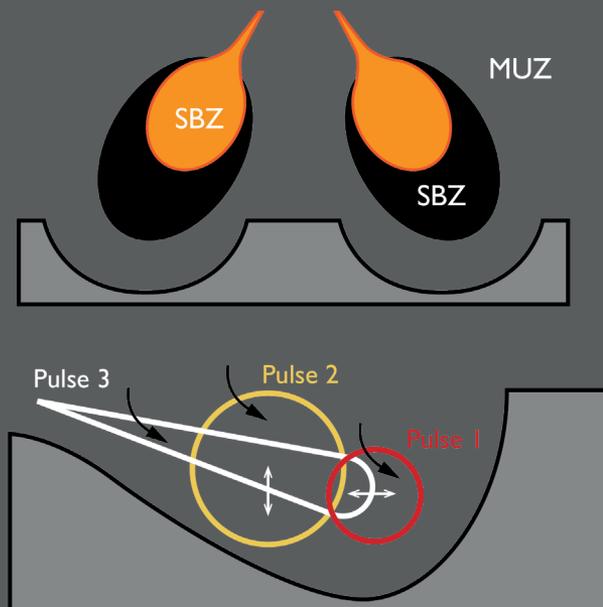


Figure 3.4 - Thermodynamic zones of DI-Pulse model

The calibration of the DI-Pulse model is achieved acting on four multipliers:

- Entrainment Rate Multiplier C_{ent}
- Ignition Delay Multiplier C_{ign}
- Premixed Combustion Rate Multiplier C_{pre}
- Diffusion Combustion Rate Multiplier C_{dif}

These parameters control any phase of the combustion process, starting from fuel penetration up to diffusive combustion, i.e. from the start of combustion up to MBF90 (90% of the fuel has burned away).

GT-Power exploits several models for the various combustion phases:

- Fuel injection

Each injection event is considered as an injection pulse and each pulse is treated separately from the others. This actually means that there is no difference between a pilot, main or post injection.

The DI-Pulse model has no limits about the number of pulses. The injected fuel straight joins the SUZ, hence the injector model requires a high degree of accuracy.

- Dragging model

The injected fuel mixes with air, residuals and fuel from other injections through the *dragging* phenomenon.

To model this behavior, it is required to apply the momentum conservation that may be modified through

the multiplier coefficient C_{ent} . The equation describing the dragging is:

$$\frac{dm}{dt} = -C_{ent} \frac{m_{inj} u_{inj}}{u^2} \frac{du}{dt}$$

in which u is the dragging velocity; t is the time, m_{inj} is the injected fuel mass; u_{inj} is the injection velocity at the nozzle tip.

- Evaporation

This model of the evaporation of fuel droplets occurs through a transfer of both heat and mass that properly represent the evaporation at limited diffusion as well as limited boiling.

- Injection delay

The mixture undergoes to a certain delay at every pulse. These delays are modeled accordingly to Arrhenius equation that may be tuned through the multiplier C_{ign} .

The injection delay is modeled per each pulse separately depending on the condition within the pulse itself:

$$\tau_{ign} = -C_{ign} \rho^{-2} \exp\left(\frac{300}{T}\right) [\text{O}_2]^{-0.5}$$

where T and ρ are the temperature and density of the pulse; $[\text{O}_2]$ is the oxygen concentration

- Premixed combustion

The premixed combustion takes place as soon as the autoignition conditions of the fuel are satisfied. The air-fuel mixture, developed after the injection delay, is involved in the premixed combustion process

The speed of this combustion phase is kinematically limited by the multiplier C_{pre} which is used to tune the model:

$$\frac{dm_{pre}}{dt} = C_{pre} m_{pre} k (t - t_{ign})^2 f([\text{O}_2])$$

in which k is the turbulent kinematic energy; t_{ign} is the ignition time, m_{pre} is the premixed mass.

- Diffusive combustion

After the premixed combustion, the fuel that has not burned yet and the gases dragged by the pulse, keep on burning in the diffusive phase.

The speed at which the mixture burns in the diffusive combustion mainly depends on the EGR rate, on the oxygen concentration $[\text{O}_2]$, on the cylinder volume V_{cyl} and on the mixture mass m . The multiplier C_{dif} is exploited to modify the following model:

$$\frac{dm}{dt} = C_{dif} m \frac{\sqrt{k}}{\sqrt[3]{V_{cyl}}} f(\text{EGR}, [\text{O}_2])$$

3.4 Burn Rate Analysis

Since it is difficult to measure the Burn Rate, then GT-Power becomes a very useful tool even if it provides only the measures of the in-cylinder pressure, obtained from the engine test bed. To compute the burn rate, GT-Power exploits a two-zone thermodynamic model: the unburned and burned gas regions.

In GT-Power, the combustion process takes place in the following way:

1. At the beginning, only the unburned gas region is present within the cylinder, also including residual gases and EGR from the previous cycle.
2. Next, part of the air-fuel mixture is moved from the unburned to the burned gas region, at each time instant. The speed at which this happens is defined *Burn Rate* and it is computed, or imposed by a combustion model.
3. As the mixture keeps on moving towards the burned gas region, GT-Power performs the chemical equilibrium for the entire burned region. In this operation, the software considers all the chemical species in the burned region (C, H, O, N, S, Ar) so as to estimate the equilibrium concentration at the end of the combustion (N_2 , O_2 , H_2O , CO_2 , H_2 , N, O, H, NO, OH, SO_2 , Ar).

The equilibrium concentration of the several species highly rely on the temperature of the burned region, but also on its pressure.

4. Next, GT-Power computes the internal energy of each species. The energy of the overall burned region

is obtained by the sum of all its contributions. Through the energy conservation equation, one obtains the new values of the temperature in the unburned and burned gas region and the equilibrium pressure.

In GT-Power the simulations are compiled by forward run, i.e. the in-cylinder pressure trend is obtained from the burn rate. However, the burn rate is actually estimated from the in-cylinder pressures measured from the real engine.

Whenever it is not possible to define the burn rate, but still the pressure values are available, then GT-Power can perform a reverse run, i.e. the burn rate is computed from the in-cylinder values of the pressure.

To this purpose, the software provides two different analyses, depending on the available input data:

- Cylinder Pressure Only Analysis CPOA
- Three Pressure Analysis TPA

The execution of the CPOA requires:

- the geometry of the engine,
- the values of the wall temperature of the liner,
- a heat transfer model,
- the fraction of the residual gases,
- the filling coefficient at the Intake Valve Closing (IVC).

The initial conditions are fundamental since they represent the starting variables affecting the outputs. Indeed, the CPOA studies the fluid-dynamic evolution of the trapped mass from the intake valve closure to the

exhaust valve opening (EVO). The computational time of CPOA is very short since only two cycles are performed per each operation point of the engine.

On the other hand, the TPA requires:

- the values of the in-cylinder pressure,
- the intake pressure (air and EGR),
- exhaust pressure (residuals, air and EGR),
- geometry of the runners (from the manifold to the cylinder),
- valves characteristics (lift plots).

The TPA is made possible only if the trend of the pressure in the intake and exhaust manifolds as function of the crank angle and the trend of the in-cylinder pressure, are known. This analysis does not need the fraction of residuals, or the filling coefficient.

Indeed, these values are estimated from the simulation. However, the TPA needs to be run at least for 20 cycles because it works for stationary conditions of the engine.

3.4.1 Cylinder Pressure Only Analysis - CPOA

In the CPOA, the in-cylinder pressure is the main parameter to be supplied in order to get the burn rate.

The pressure values must be measured, or for a single cycle or for an average cycle, as well as some results from the base average cycle. The model runs two cycles to converge the results. For this reason, all the valves, ports and connections are not necessary.

As already mentioned, the input parameters are the geometry of the cylinder, the wall temperature of the liner, a heat transfer model and the initial conditions.

In GT-Power, all of these data input are set in the model *EngCylinder*. The required initial conditions are represented by the state of the trapped air during the IVC, and so the trapping ratio, and the residual gas fraction.

Actually, these quantities are tough to be estimated or calculated during the test on the real engine so, this represents the main drawback of the CPOA. However, the benefit is the fast response owing to the approach of the analysis which only requires the instantaneous pressure trace developed inside the cylinder to compute the burn rate.

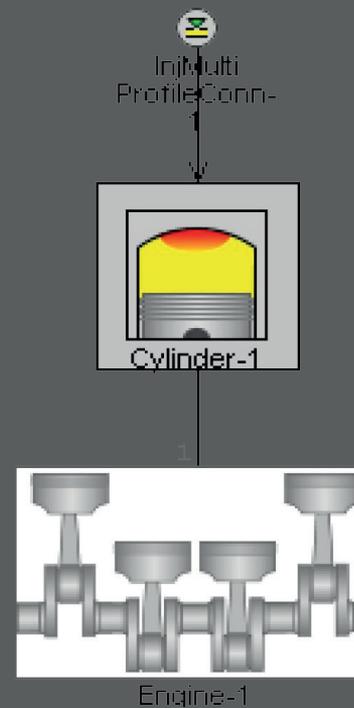


Figure 3.5 - Scheme of a CPOA in GT-Power

3.4.2 Three Pressure Analysis - TPA

The TPA considers three pressures to define the burn rate: intake, in-cylinder and exhaust pressure. The engine model must be provided with valves and connections. This analysis performs the reverse run simulation, which consists in looping for every time instant the fuel quantity passing from the unburned to the burned gas region until the estimated pressure matches the measured in-cylinder pressure. One of the advantages of the TPA is that the trapping ratio and the residual gas fraction are not necessary input because these last are estimated by the model. This is very useful whenever they are tough to be measured or, difficult to calculate, as for partial load operations.

The drawback is the development of a detailed model with more experimental data and consequently higher computational time since the simulation runs over multiple cycles up to reach the convergence. The TPA may be performed in two different ways:

- TPA-Steady

it includes in GT-Power a unique pressure trend per each operation point. This single trend is the average value of the pressure among all the consecutive measured pressure cycles. As a result, TPA-Steady supplies a single burn rate per each engine point.

- TPA-Multicycle

it includes in GT-Power the values of the pressures measured in all consecutive cycles per each operation point of the engine, without any preliminary tuning. The results of TPA-Multicycle focus on the cyclic variation of the pressure trend, cycle by cycle.

The simulation algorithm that computes the burn rate is divided into three calculus phases:

1. Cycle I

a fictitious Burn Rate is imposed without measuring the pressure,

2. Cycle II

the simulation pauses as soon as the Forward Burn is achieved. Starting from the point named Start of Cycle, the software estimates the Apparent Burn Rate by exploiting the thermodynamic conditions of the fluid trapped into the cylinder, and the correspondent pressures. In the same time, the injection profiles and the Heat Transfer Rate are derived from the previous cycle,

3. Cycle III

the Forward Run continues with its development of the mathematical model, this time by imposing the Burn Rate of Cycle II until the results are stationary.

Thus, the analysis determines a burn rate per each engine operational point, at the same time performing the consistency check to verify the convergence between the experimental data with the trend of the computed burn rates.

TPA Consistency Checks

The calculus of the burn rate from the in-cylinder pressure signal always implies several errors in experimental input data, or in the assumptions needed to balance the difficulties to precisely measure certain phenomena.

For example, the heat exchange between the charge and the wall must be predicted based on assumptions. Consequently, all these assumptions play a key-role concerning the error that marks the burn rate computation. In order to meet this problem, GT-Power tunes the lower specific heat power of the fuel through a multiplier coefficient named *LHV multiplier* which therefore represents the overall error rate. The LHV multiplier is determined per each case of the TPA, and it must range in between 0.95 and 1.05,

In addition, GT-Power carries out tests over different resultant quantities coming from the simulation, and communicate whether these quantities fail. These tests are named *Consistency Checks* and are supplementary to the checks needed to verify the convergence of the experimental data with the computed burn rate trends.

GT-Power performs the following consistency checks:

- In-Cylinder Pressure at IVC

in TPA, the measured pressure profile is shifted automatically to match the start of the cycle and the 40 degrees preceding the TDC. If the quantity of the shift is higher than 0.5 bar, GT-Power returns an error;
- Cumulative Burn Rate during Compression

in Diesel engines, the fuel is injected at the end of the compression stroke. The combustion is expected to occur only after the fuel is actually injected in the cylinder. However, sometimes the burn rate is registered before the SOI. If the integral of such burn rate reaches the 2% of the total injected fuel mass, GT-Power returns an error.

- Compression Slope

on the $\log P$ - $\log V$ diagram of an engine cycle, it is required that the compression slope is constant. In direct injection engines, the value of the polytropic compression transformation is approximable to the ratio between the air specific heats. This value is about 1.4 at 300 K and reduces up to 1.3 at 1000 K. This consistency check determines the value of the polytropic coefficient downstream of the simulation and tests if this value overcome the admissible limit, otherwise, GT-Power returns an error;

- Fraction of fuel Injected Late

this check only apply to direct injection engines and points out that there is no sufficient fuel to sustain the produced burn rate. Thus, the missing fuel quantity is defined instant per instant and returned at the end of the simulation. If the quantity of the missing fuel break the 2% of the total quantity, GT-Power returns an error;

- Large LHV change required

this check returns that the LHV multiplier coefficient is out of the permitted range.



4

F1A ENGINE MODELS IN GT POWER

Detailed Engine Model

The Controllers

Fast-Running Engine Model

Comparison between DETM and FRM

At this point, the analysis has been focused on the working principles of CI engines and on the theory behind the GT-Power simulation software. Now, it is required to merge the information about the FPT Industrial FIA engine and the virtual environment. In the following, the study carries out the development of the FIA engine model in GT-Power:

4.1 Detailed Engine Model

Figure 4.2 shows the scheme of the FIA model in GT-Power in which the main components have been highlighted. In the detail modelling, the target is to reproduce the engine as close as possible to reality, also considering the geometry, bendings and connections of the pipes.

The model was developed starting from a non-predictive version, i.e. with no HRR simulation, provided by FPT in September 2019.

This version includes improved intake/exhaust ports flow coefficients and swirl based on values measured on the head in Arbon, Switzerland (this model is an improved version compared to a previous GT-power model version provided to PoliTO by FPT in November 2018).

The geometry of the model has been revised on the basis of the measurements taken on the FIA engine installed at PoliTO test bench and it was calibrated and assessed using data provided by FPT in July 2019, which were acquired on the FIA engine provided to PoliTO. These data include a full engine map in which only HP-EGR was used.

The DI Pulse heat release predictive model has been calibrated on the bases of these data.

In the following, the main components are further investigated to fully appreciate the model characteristics.

Engine

The Engine accounts for the firing order; geometry of the cylinders, engine speed, inertia, and frictions. In this specific model, A variable compression ratio *epsilon* has been adopted (function of BMEP and speed), based on the measurements provided by FPT. The variable CR approach has been agreed with FPT, as it allows to obtain a better matching with in-cylinder related quantities.

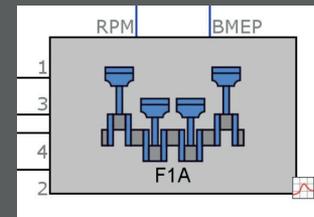


Figure 4.2 - Engine FIA part in the Detailed Model

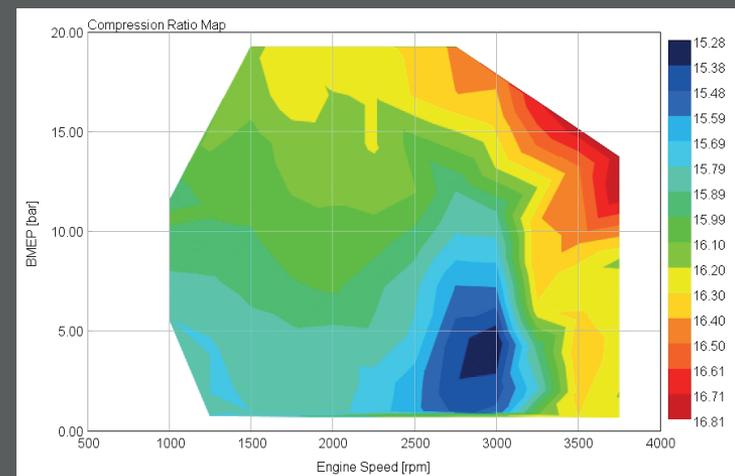


Figure 4.3 - Compression Ratio Map

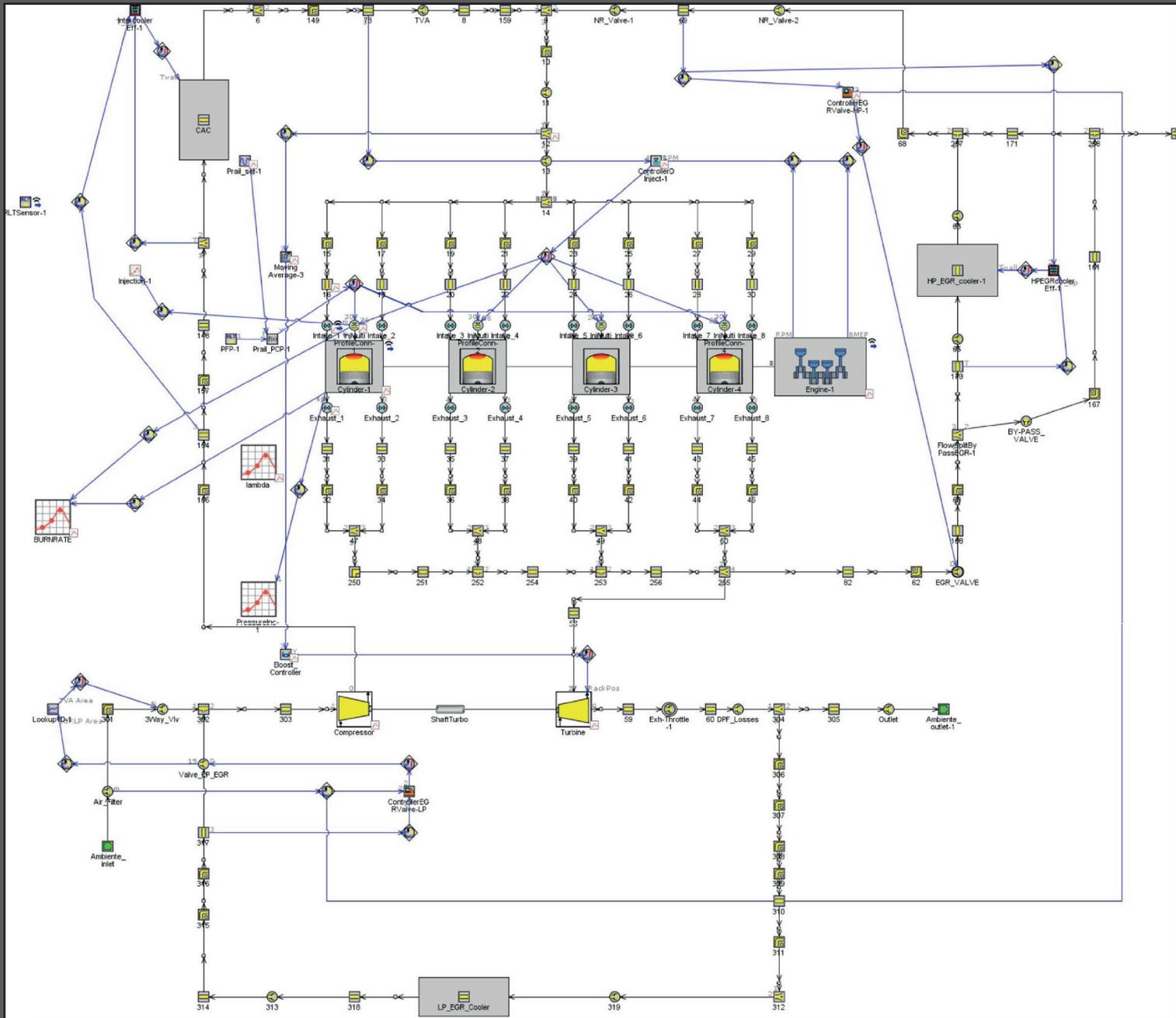


Figure 4.1 - FIA Detailed engine model in GTPower provided by FPT

Injectors

Currently, the installed nozzle is 7x300cc / 30s. The SOI values of pilot and main that were obtained from the curves of the injector currents provided by FPT. These values have been organized in look-up tables as function of BMEP and speed. Figure 4.5 shows the SOI trend for the Pilot, Pre and Main events. Referring to Table 4.1, remind that the nozzle discharge coefficient is the ratio between the instantaneous, actual fuel rate and the correspondent ideal (isentropic) fuel rate measured at the same upstream conditions.

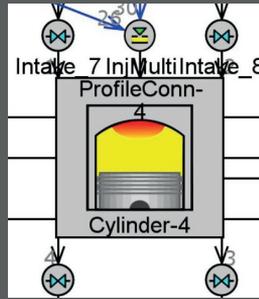


Figure 4.4 - Injector part in the Detailed Model

Atomizer Diameter	0.124 mm
Number of holes per injector	7
Nozzle discharge coefficient	0.86

Vapor Fluid Object	diesel-vap
Heat of Vaporization @ 298 K	250000 J/kg
Density	830 kg/m ³
Absolute Entropy @ 298 K	2913.25 J/kgK
(T-T _{ref}) Coefficient, a1	2050
(T-T _{ref}) ² Coefficient, a2	0
(T-T _{ref}) ³ Coefficient, a3	0

Table 4.1 - Nozzle and Fuel Specification in GT-Power

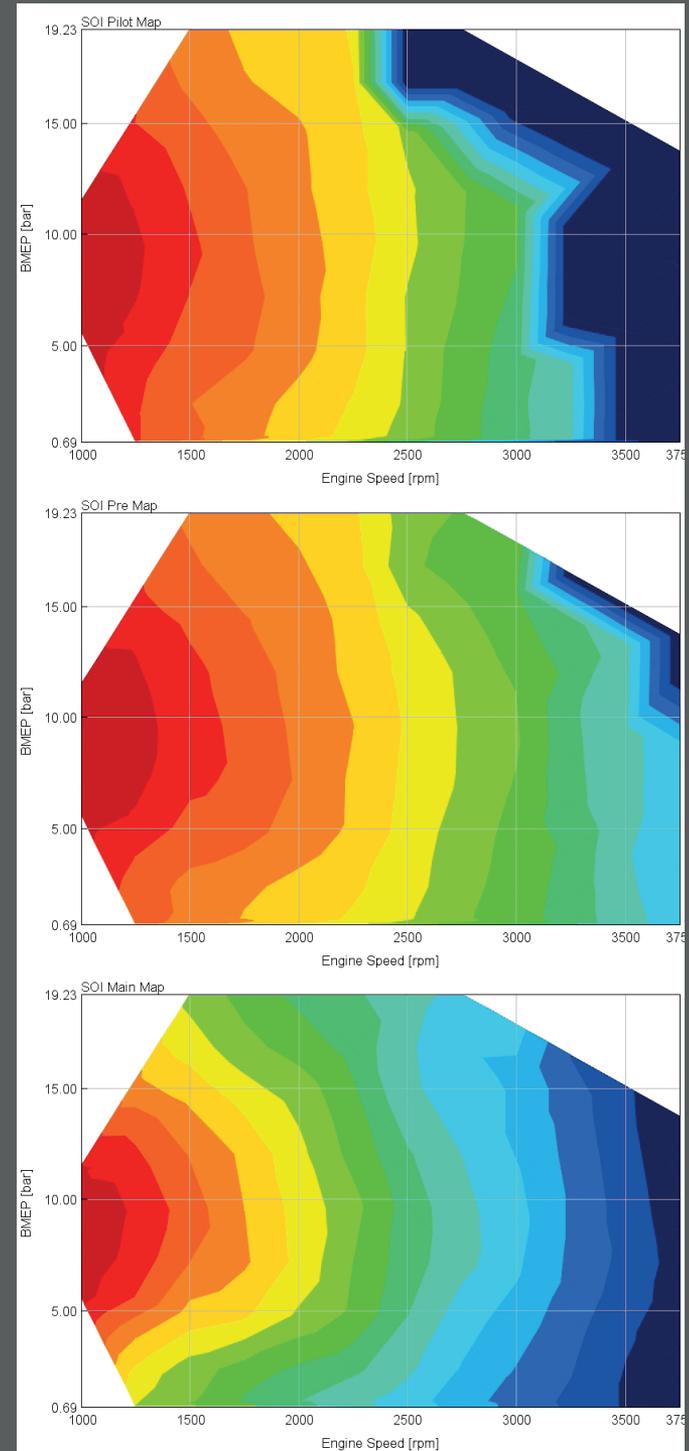


Figure 4.5 - SOI Map for Pilot, Pre and Main Injections

Reminding paragraph 1.6, the data provided by FPT are:

- electric commands for the injectors,
- NOD basing on the upstream rail pressure of the Common Rail system.

Starting from these supplied data, it has been possible to determine the SOI_{hydr} , that is the crank angle for which the injection actually takes place. Note that SOI_{hydr} differs from SOI_{el} which is the crank angle degree for which the electrical command is applied to the injector. The steps to determine SOI_{hydr} are:

- assessment of SOI_{el} per each injection event starting from the trend of the electrical command applied to the injectors,
- computation of SOI_{hydr} through the following relation:

$$SOI_{hydr}[\text{deg}] = SOI_{el}[\text{deg}] + \frac{6 \cdot n [\text{rpm}] \cdot \text{NOD} [\mu\text{s}]}{10^6}$$

- construction of the Injection Rate Map (fuel flow rate vs crank angle degree) per each combination of rail pressure and ET.

The tab *Profile Object* of the injector have been implemented with the table *Injection Rate Map*, which relates the fuel flow rate, defined as function of the crank angle degree, both with the rail pressure and the Energizing time (ET).

These injection profiles are determined by flushing the injector keeping constant the upstream pressure and measuring the fuel flow rate. GT-Power exploits the rail pressure and the map profiles to extract the ET and so building up the injection profile that is used for the correspondent engine working point.

Attribute	Unit	Object Value
Units for Time ("XData" below and X values of Injection ...		micro s ...
Units for Pressure ("Y Data" below)		bar ...
Units for Mass Flow Rate (Y values of Injection Profiles)		g/s ...

Inj Rates	Y Data #	1	2	3	4	5	6	7	8	9	10	11	12	13	14
XDat #		77.09 ...	300 ...	400 ...	500 ...	600 ...	700 ...	800 ...	900 ...	1000 ...	1100 ...	1200 ...	1300 ...	1400 ...	1500 ...
1	400 ...	400x300min ...	400x300 ...	400x400 ...	400x500 ...	400x600 ...	400x700 ...	400x800 ...	400x900 ...	400x1000 ...	400x1100 ...	400x1200 ...	400x1300 ...	400x1400 ...	400x1500 ...
2	500 ...	500x300min ...	500x300 ...	500x400 ...	500x500 ...	500x600 ...	500x700 ...	500x800 ...	500x900 ...	500x1000 ...	500x1100 ...	500x1200 ...	500x1300 ...	500x1400 ...	500x1500 ...
3	600 ...	600x300min ...	600x300 ...	600x400 ...	600x500 ...	600x600 ...	600x700 ...	600x800 ...	600x900 ...	600x1000 ...	600x1100 ...	600x1200 ...	600x1300 ...	600x1400 ...	600x1500 ...
4	800 ...	800x300min ...	800x300 ...	800x400 ...	800x500 ...	800x600 ...	800x700 ...	800x800 ...	800x900 ...	800x1000 ...	800x1100 ...	800x1200 ...	800x1300 ...	800x1400 ...	800x1500 ...
5	1000 ...	1000x300min ...	1000x300 ...	1000x400 ...	1000x500 ...	1000x600 ...	1000x700 ...	1000x800 ...	1000x900 ...	1000x1000 ...	1000x1100 ...	1000x1200 ...	1000x1300 ...	1000x1400 ...	1000x1500 ...
6	1200 ...	1200x300min ...	1200x300 ...	1200x400 ...	1200x500 ...	1200x600 ...	1200x700 ...	1200x800 ...	1200x900 ...	1200x1000 ...	1200x1100 ...	1200x1200 ...	1200x1300 ...	1200x1400 ...	1200x1500 ...
7	1400 ...	1400x300min ...	1400x300 ...	1400x400 ...	1400x500 ...	1400x600 ...	1400x700 ...	1400x800 ...	1400x900 ...	1400x1000 ...	1400x1100 ...	1400x1200 ...	1400x1300 ...	1400x1400 ...	1400x1500 ...
8	1600 ...	1600x300min ...	1600x300 ...	1600x400 ...	1600x500 ...	1600x600 ...	1600x700 ...	1600x800 ...	1600x900 ...	1600x1000 ...	1600x1100 ...	1600x1200 ...	1600x1300 ...	1600x1400 ...	1600x1500 ...
9	1800 ...	1800x300min ...	1800x300 ...	1800x400 ...	1800x500 ...	1800x600 ...	1800x700 ...	1800x800 ...	1800x900 ...	1800x1000 ...	1800x1100 ...	1800x1200 ...	1800x1300 ...	1800x1400 ...	1800x1500 ...
10	2000 ...	2000x300min ...	2000x300 ...	2000x400 ...	2000x500 ...	2000x600 ...	2000x700 ...	2000x800 ...	2000x900 ...	2000x1000 ...	2000x1100 ...	2000x1200 ...	2000x1300 ...	2000x1400 ...	2000x1500 ...
11	2200 ...	2200x300min ...	2200x300 ...	2200x400 ...	2200x500 ...	2200x600 ...	2200x700 ...	2200x800 ...	2200x900 ...	2200x1000 ...	2200x1100 ...	2200x1200 ...	2200x1300 ...	2200x1400 ...	2200x1500 ...

Figure 4.6 - Injection Rate Map supplied to GT-Power evaluated from experimental data

Turbine and Compressor

The Compressor/turbine maps, and valve lift profiles, have been maintained as in the original GT-power model provided by FPT.

In the turbine, an efficiency multiplier of 0.95 and a mass flow multiplier of 1.05 were set in order to better match exhaust p/T. Because of the VGT system, the turbine is provided of a map per each operating condition. More in detail, the VGT system is treated in-depth in the paragraph concerning the controllers.

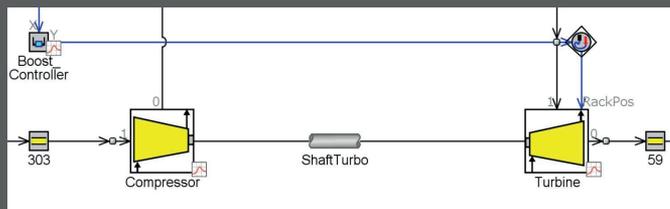


Figure 4.7 - Turbocharger part with VGT controller

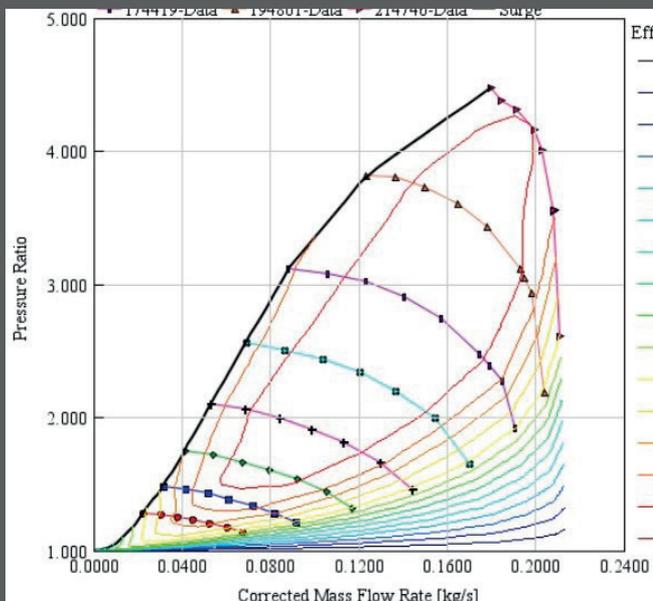


Figure 4.8 - Compressor efficiency map, max speed 226 krpm

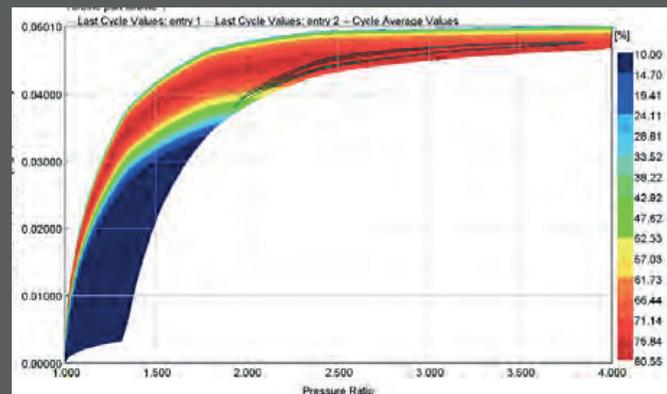


Figure 4.9 - Turbine efficiency map for Efficiency and Mass multipliers equal to 1

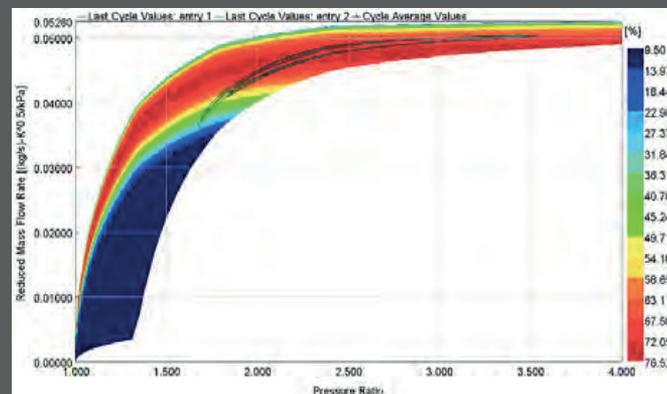


Figure 4.10 - Turbine efficiency map for Efficiency and Mass multipliers equal to 0.95 and 1.05

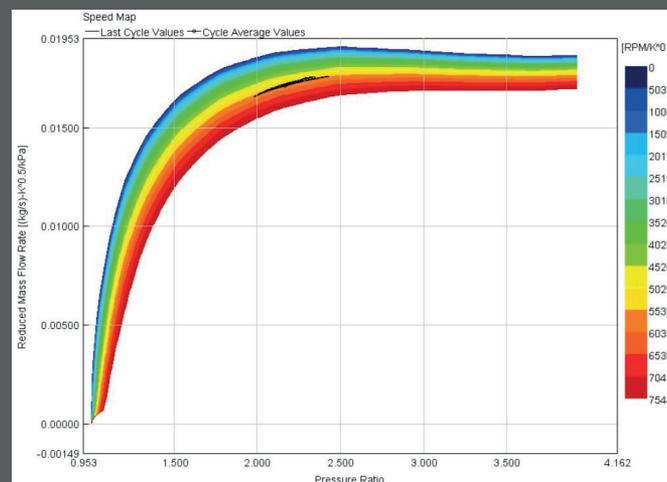


Figure 4.11 - Turbine Speed Map, 3750 rpm x 13.74 bar

Intercooler

A turbocharger group increases the intake air pressure as well its temperature.

This side effect causes some troubles because by increasing the air temperature, its density reduces, and so the increase in temperature tends to vanish the benefit of the compressor which raises the pressure up.

Since the air mass sucked by the engine is directly proportional to the intake pressure and inversely proportional to the root square of the intake temperature:

$$m_a \propto \frac{P_a}{\sqrt{T_a}}$$

This actually means that if the intake temperature increases, then the sucked air reduces, as well as the expected power and BMEP.

For this reason, an intercooler; also known as *CAC - Charge Air Cooler*; is placed in between the intake manifold and the compressor:

The intercooler is an heat exchanger in which two air flows clash each other to transfer heat as fast as possible. Indeed, the intercooler is always mounted in the front of the vehicle so as to be hit by the air before everthing else.

The air pressurized by the compressor travels along small pipes which are immersed by the fresh airflow. The temperature reduction allows:

- to raise the air density in the intake manifold and so increase the amount of air through the intake valves,

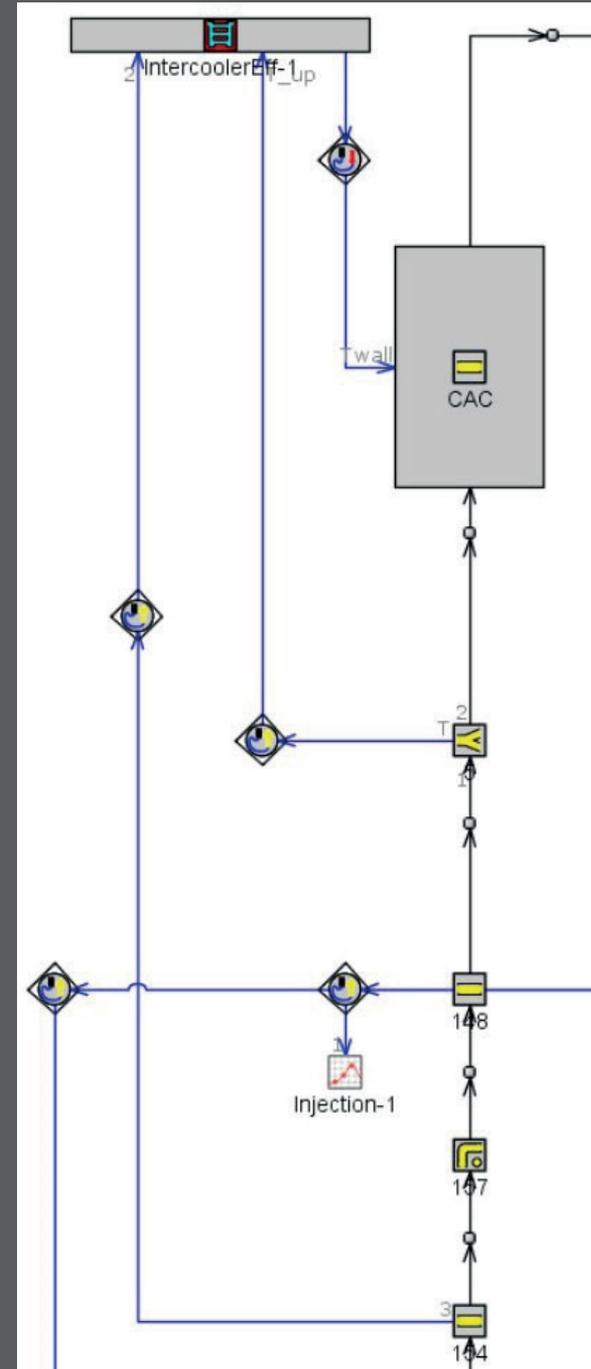


Figure 4.12 - Intercooler (CAC) part in the Detailed Model

- to reduce the temperature values during the cycle, leading lower thermal load on the engine and lower thermal stress over the mechanical components,
- to enhance the organic efficiency since the power is increased without changing friction forces,
- to reduce the pollutants emission, in particular that of NO_x since it is very sensible to the maximum temperature (thermal formation mechanism),
- to reduce knocking in SI only.

In Gt-Power, the intercooler performance has been modeled on the basis of the efficiency:

$$\text{Efficiency} = \frac{T_{in} - T_{out}}{T_{in} - T_{coolant}}, \quad T_{coolant} = 20^{\circ}\text{C}$$

which experimental data have been provided by FPT. The obtained results are visible on in the plot of Figure 4.13.

The gas temperature at the outlet of the cooler is estimated from the previous formula:

$$T_{out} = T_{in} - \text{Eff}(T_{in} - T_{coolant})$$

where T_{in} and Eff are known in the model.

The resultant T_{out} is then imposed as wall temperature of the intercooler, and the heat transfer multiplier inside the CAC is set at 1000, so that the outlet gas temperature will be coincident with the wall temperature.

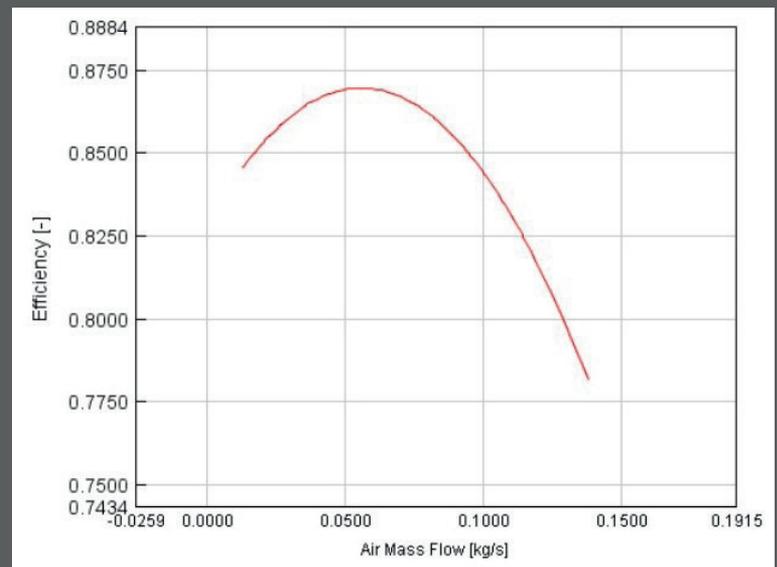
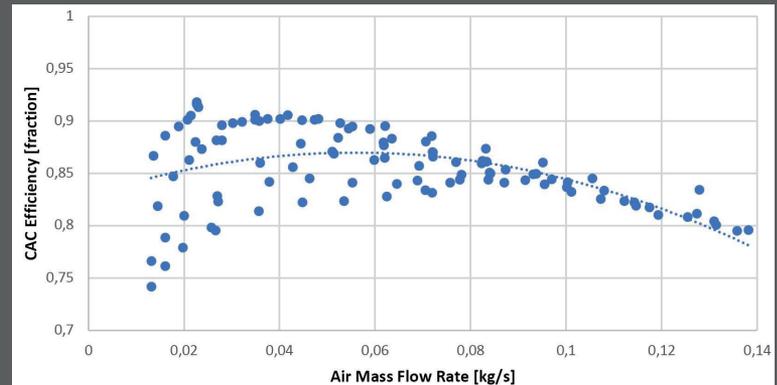


Figure 4.13 - Intercooler (CAC) efficiency, experimental (above) and supplied to GT-Power (below)

Exhaust Gas Recirculation System - EGR

The FIA engine is provided of a dual EGR system:

- high-pressure, short-route circuit HP EGR,
- low-pressure, long-route circuit LP EGR.

Figure 4.14 represents the model of the LP EGR of the FPT FIA engine. This EGR layout is still in development and it is not yet validated owing to the unavailability of tests. For this reason, in all the following studies, the diameter of the LP EGR valve is set to zero.

Differently from a HP EGR circuit, in the LP EGR the exhaust gases are taken downstream the Diesel Particulate Filter (DPF). Before being actually recirculated, the exhaust gases must be cooled down through a LP cooler (which efficiency has not yet been derived due to unavailability of experimental tests). After cooling, the gases are sent back in the intake system before the compressor, meaning that the gases of the LP EGR mix with the uncompressed, fresh air. The gases coming from the DPF contain much less soot and highly reduce fouling in the intake system.

The LP EGR system implies that both fresh air and exhaust gases pass through the compressor; consequently it must be larger to accommodate a higher mass flow rate. This is useful at low speed, both high and low loads, where usually EGR rate are higher and mass flow rate are lower.

Another difference between the two system is in the so-called *EGR lag*. In the LP system, the gases are much better distributed over the cylinders since these gases must pass through all the intake system, mixing with air.

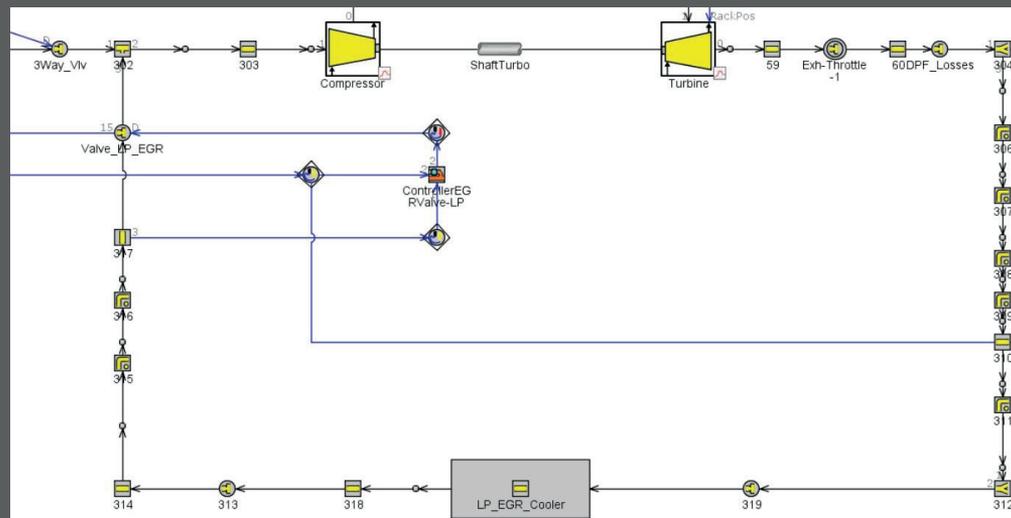


Figure 4.14 - LP EGR system. not yet validated

The HP EGR system is the most used in automotive application, both passenger car, light-duty or heavy-duty engines. The model of the high-pressure, short-route EGR of the FIA engine is represented in Figure 4.15. Here, the exhausted gases are recirculated from the exhaust manifold before passing through the expansion of the turbine. The non-recirculated gas, as usual, enter the turbine scroll and then through the ATS.

While instead, the recirculating gases pass through a HP water-cooling heat exchanger (the temperature of the coolant is about 90°C) to avoid the overheating of the intake manifold. In some critical operating conditions (cranking for example), the HP EGR is bypassed until the coolant does not reach about 60°C. The cooler efficiency has been extracted from the FPT data.

After the cooling, the gases are regulated through a electronically-controlled valve that adjusts the quantity of exhausted gases into the intake manifold. The recirculation from the exhaust to the intake side is made possible only if the pressure in the intake manifold is significantly lower compared to that of the exhaust manifold. Indeed, at low loads, this condition may occur and the gases do not recirculate even though the valve is completely open. To counter this problem, a butterfly valve is placed upstream the intake manifold to increase the pressure gradient. Despite this valve introduces dissipation, it is useful to control pollutants emission.

The advantages of HP EGR are represented by the easiness and the quick response of the system. However it must face the fouling of the cooler, the dissipation, and the lack of supercharging since part of the gases do not reach the turbine.

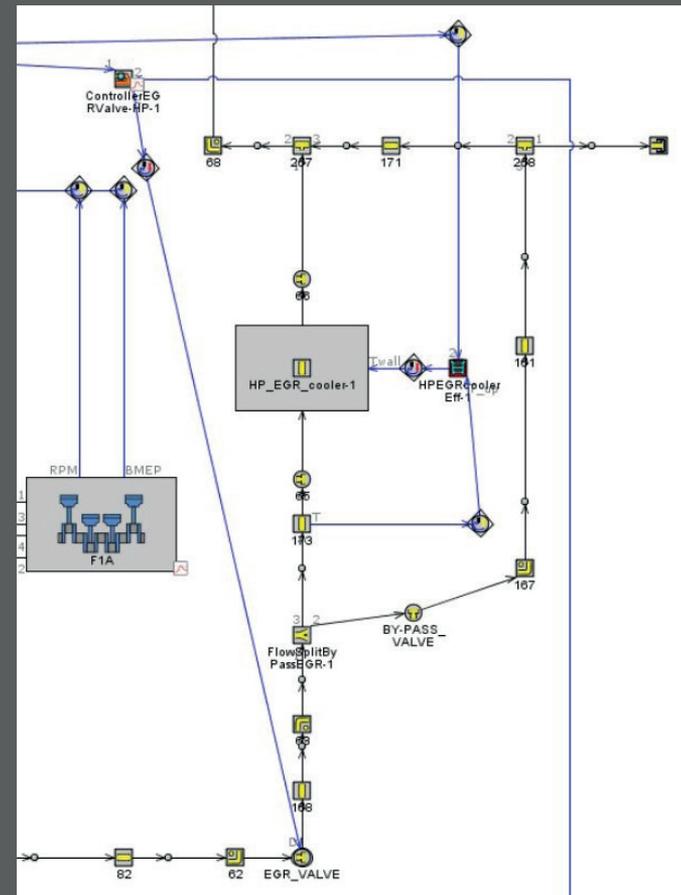


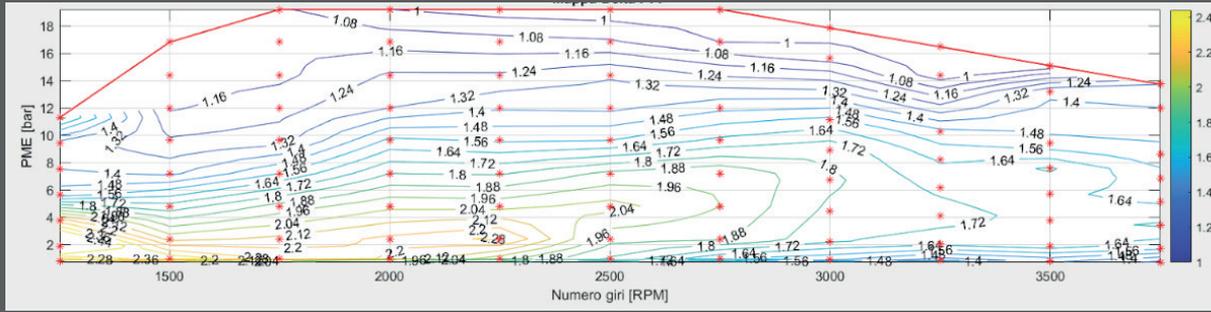
Figure 4.15 - HP EGR system, validated

DI-Pulse Heat Release Model

The model has been calibrated using the data and the related experimental pressure traces provided by FPT.

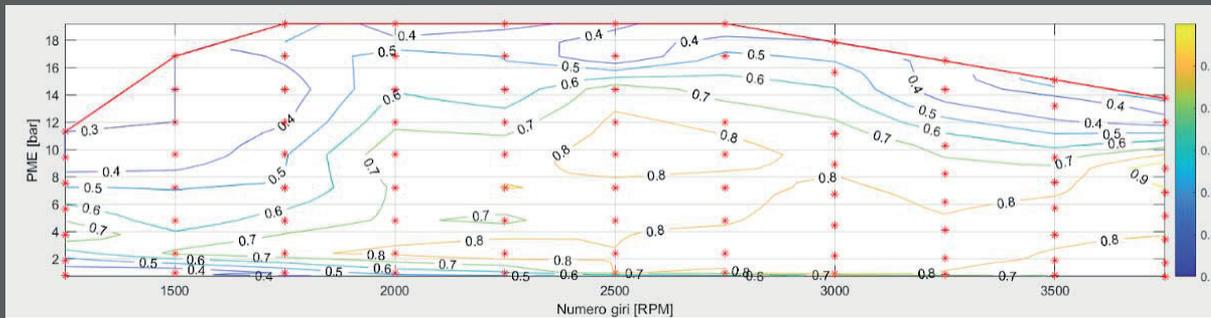
The points with poor combustion ($\lambda < 1.2$, $N = 1000$ rpm) were excluded from the calibration dataset.

In the end, a look-up table for the 4 tuning parameters of the DIPulse model was obtained. These maps are shown in Figure 4.16.



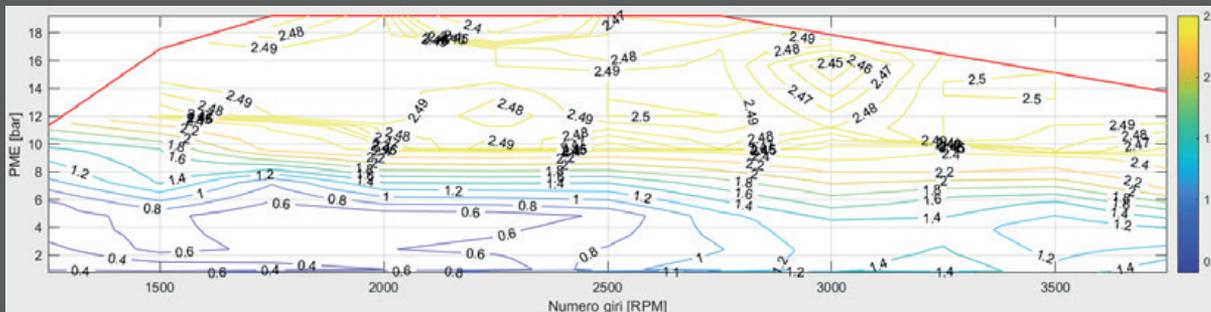
Entrainment
Multiplier

$$C_{ent}$$



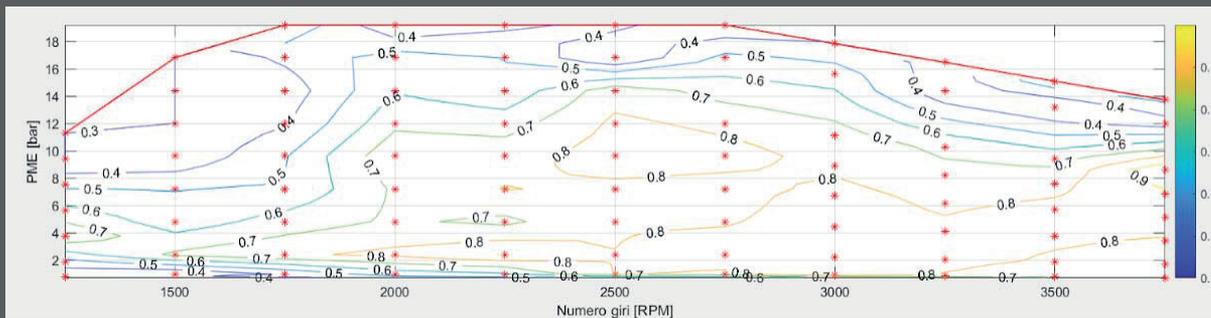
Ignition Delay
Multiplier

$$C_{ign}$$



Premixed Comb.
Multiplier

$$C_{pre}$$



Diffusive Comb.
Multiplier

$$C_{dif}$$

Figure 4.16 - DI-Pulse coefficients maps

4.2 The Controllers

The GT-Power model is provided of *controllers* that regulate the quantities during the simulation. These controllers acts on:

- BMEP,
- VGT,
- LP EGR,
- HP EGR.

4.2.1 Mean Effective Pressure (MEP) Controller

This controller is used to target various engine performance parameters at part-load operation by adjusting the injected fuel quantity. It contains a model-based feed-forward controller in parallel with a continuous feedback, model-based controller with anti-windup protection.

Its purpose is to eliminate the need for the user to create the 'PIDController' for the particular application.

At every timestep, the controller calculates the desired injection quantity based on several physical quantities of the engine including fuel energy, engine size, and engine speed.

The feed-forward portion estimates the injection quantity based on modeling fundamentals without taking into account the "error" between the specified Target Signal and the Input Signal. The feedback controller uses an adjustable-gain PI technique to minimize the remaining error.

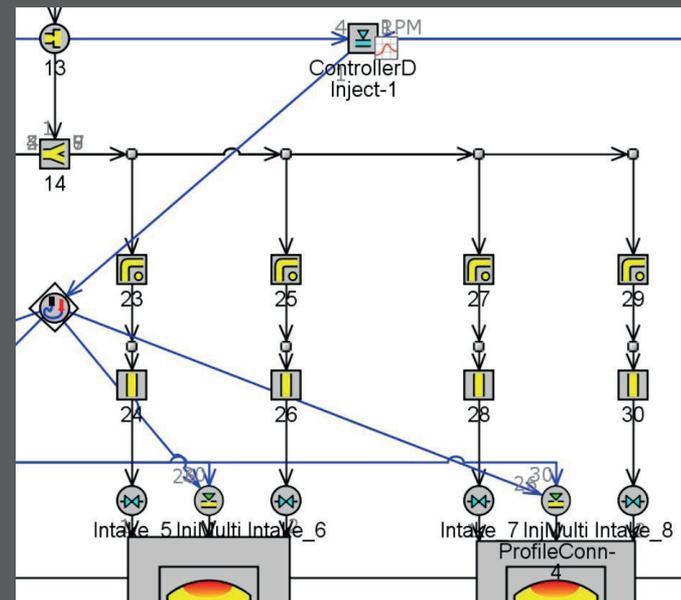


Figure 4.17 - BMEP controller

4.2.2 Variable Geometry Turbocharger (VGT) Controller

The Variable Geometry Turbocharger (VGT) is a widely used in CI engines because exhaust temperatures are lower, compared to SI engine, and so there are no limitation in the turbine construction.

In a naturally aspirated (NA) engine, the quantity of intake air is limited by the available space, i.e. the displacement. To have more air within the cylinders, it is required that the cylinder environment must be at higher pressure and density than external environment. The specific power of an engine can be expressed as:

$$\text{Specific Power: } \frac{P_u}{iV} = \text{MEP} \cdot \frac{n}{m}$$

To increase power, one can increase the rotational speed of the engine n , or the MEP.

Since higher rotational speed leads to higher inertial stresses (they are dependent on the square of the speed), then more power passes through higher MEP, even though engine components must be reinforced. The higher the intake air density, the higher the MEP.

$$\text{MEP} = \eta_u \cdot \lambda_u \cdot \rho_a \cdot \frac{H_i}{\alpha}$$

In a turbocharger, the compressor is moved by the shaft connected to the turbine. The exhaust gas residual energy is exploited in the turbine through the expansion in the scroll. The engine is not directly providing work to the shaft, and so the turbocharging system does not subtract mechanical power from the engine, still this system affects the engine with higher backpressure at the exhaust due to the higher pressure drop (expansion) in the turbine. This fluid-dynamic connection between the turbogroup and the engine suffers a controllability issue. To this purpose two main strategies are applied:

- the *wastegate* valve upstream the turbine opens to bypass the gas expansion, in this way the maximum tolerable boost is avoided. Note that the wastegate valve also opens to maintain the boost pressure. However, in this way part of the energy is dissipated,
- the VGT has no wastegate valve because the boost pressure is regulated by the channels of the stator blade: different sections of the pipes means different velocity triangles. The controller acts on the rack position that defines the channel profiles. Since the stator blade may change direction, the correspondent metal angle (angles related to the profile of the stator and rotor blades) changes too. When the engine is at low rpm, the blades tend to close because the mass flow

rate through the turbine is significantly lower. The reduction of the channel section causes the flow to hit the blades with almost all the tangential component of the speed. On the contrary, at high rpm, the channel section is higher because the mass flow rate is high as well. In this case, the tangential component of the gas is much lower. The target is to diminish the work developed by the turbine to control the boost pressure. Thus, the VGT controller regulates the rack position, i.e. stator blade inclination, so as to reach the designed boost pressure.

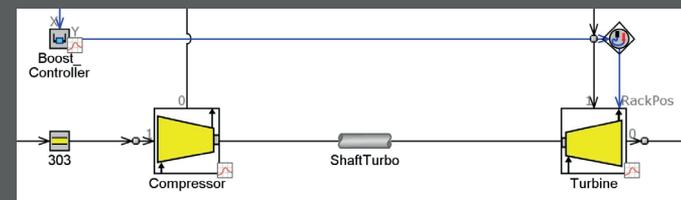


Figure 4.18 - VGT controller

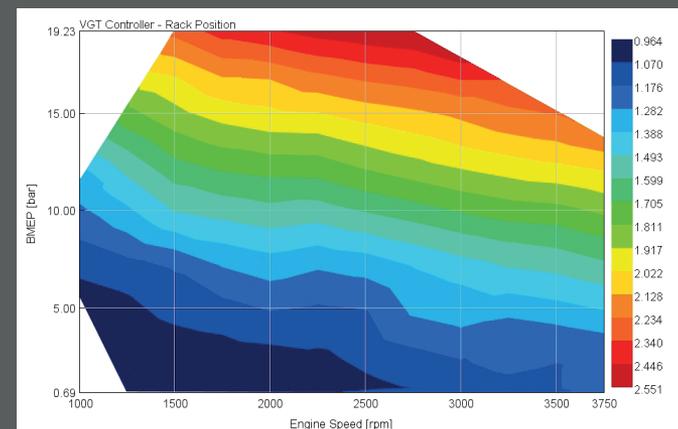


Figure 4.19 - VGT controller map

4.2.3 HP EGR Controller

As already mentioned, LP EGR system, and its controller, is not of interest since the lack of experimental tests does not allow the validation of the system yet.

On the contrary, the HP EGR controller is fully operational. This controller contains a model-based strategy used to target EGR rate by controlling throttle angle or an orifice diameter:

Again, as for BMEP controllers, the purpose of this component is to eliminate the need for using the 'PIDController' template, which requires significant effort to determine gains, for this particular application.

There may be some instances where it is desirable to have a restriction valve, but where the mass flow is imposed by some means other than by an EGR fraction (maybe a turbine bypass or a restrictor to impose a certain flow rate in the exhaust).

The input map for this controller points out the quantity of EGR desired specified as a function of EGR mass flow rate (as determined by the input signal connected to the controller part) divided by the sum of EGR mass flow rate and air mass flow (as also determined by input signal into the controller part). This calculus leads to the so-called *Target EGR Fraction (EGR + air)*.

If multiple EGR circuits are modeled then it also required to determine the *Total Target EGR Fraction* which needs to be specified in each controller:

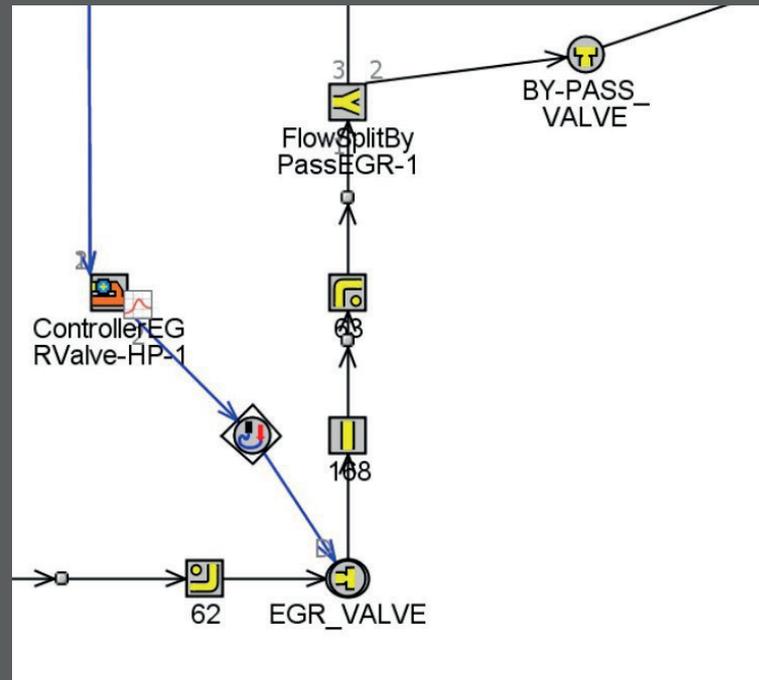


Figure 4.20 - HP EGR controller

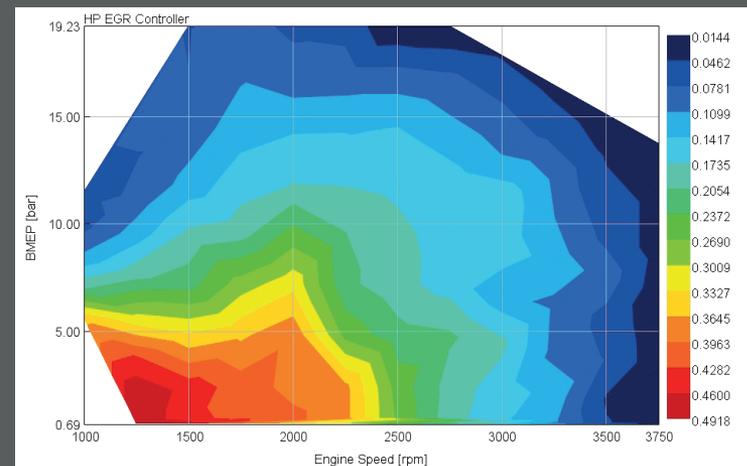


Figure 4.21 - HP EGR controller map

4.3 Fast-Running Engine Model

The Fast-Running engine model represents a step ahead of the Detailed model.

GT-Power is a CFD ID simulation software through which it is possible to model and then simulate engine performance with different degrees of detail. Depending on the test, the response of a Detailed model may be too slow to be implemented in virtual tests. For this reason, the Detailed model is converted in a more responsive Fast-Running model. The kind of experiments which require a Fast-Running model are discussed in the following chapter:

In general, a Fast-Running model is developed by simplifying the detailed GT-Power engine model. These simplifications consist in reducing the number of pipes, in increasing the discretization length and in simplifying the EGR cooler and intercooler systems, while maintaining the same detailed simulation as the in-cylinder combustion process.

The ultimate target of this operation is to reduce the Factor of Real Time (FRM), while maintaining the degree of accuracy of the original Detailed model. The real-time factor is the ratio between the time required for the simulation and the simulated time interval length. The lower the FRM, the more responsive the model, i.e. the lower the time to run the simulation.

Considering the FIA engine model, the differences from the Detailed model are:

- the exhaust ports and runners have been simplified into flowsplits. It was verified that the 4ports cannot be further simplified into a singleflowsplit, otherwise oscillations take place in the solution.
- the exhaust manifold has been simplified into a flow-split.
- The intake ports been simplified into a single flowsplit, which also includes the original flowsplit of the intake manifold. It was verified that this simplified approach does not lead to any significant detriment of the accuracy in the predicted air mass flow, and no oscillations take place in the solution.
- The geometry of the intake side downstream from compressor and of the HP EGR circuit has been kept as in the original model. It was in fact verified that if a simplification is carried out, a significant detriment in the trends of the predicted intake temperature and pressure occur especially in transient operation. Moreover, the computational time benefit that would be obtained by simplifying this part is negligible.
- The geometry of the LP EGR circuit has been simplified. However, this part has not yet been validated.
- Intercooler and HP EGR cooler objects have been replaced with "HeatExchangerConn" elements, which allow to reduce the computational time.
- The pipe "59" downstream from the turbine was one of the bottlenecks in terms of computational step constraint. Its original length of 10 mm has been extended to 100mm. It was verified that the impact on the results is negligible.
- For the same reason as above, the pipe "60" has been extended from 30 to 100 mm, the pipe "58" from 44 to 100 mm, and pipe "303" from 34 to 100 mm. It was verified that the impact on the results is negligible.

Figure 4.22 shows the Fast-Running model obtained from the evolution of the Detailed model of the FPT FIA engine. The Factor of Real Time has been reduced to 1/10, indeed the FRT of the Detailed model is on average 60, whereas in contrast the FRM of the Fast-Running model is about 6-8. The comparison of the FRT between the two models, as well as any other comparison, has been performed running the models in the steady state conditions represented by the engine map of Figure 4.24, and its correspondent table.

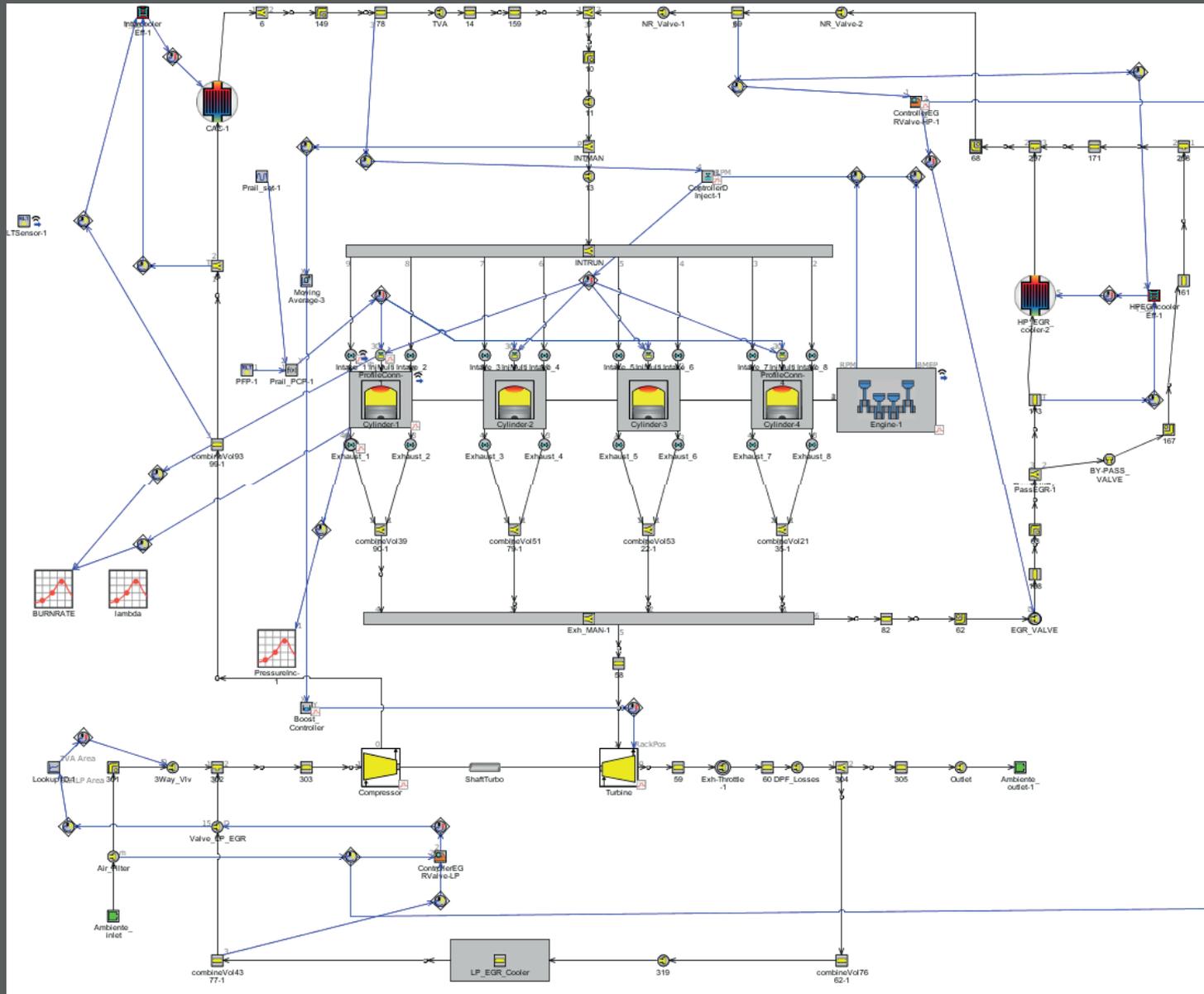


Figure 4.22 - Fast-Running model, Factor of Real Time about 6-8

In the simulations, the BMEP controller has been removed because the experimental quantity of the main has been set as input data. In addition, the opening /closing rate of the HP EGR valve controller was limited to 40, otherwise frequent oscillations take place. To further reduce the computational time, the discretization length of all the pipes was set at 1000 mm (i.e. each part is not discretized in sub-volumes), and the maximum time step, in the 'Time Step and Solution Control Object' in the Run Setup, was set at 720.

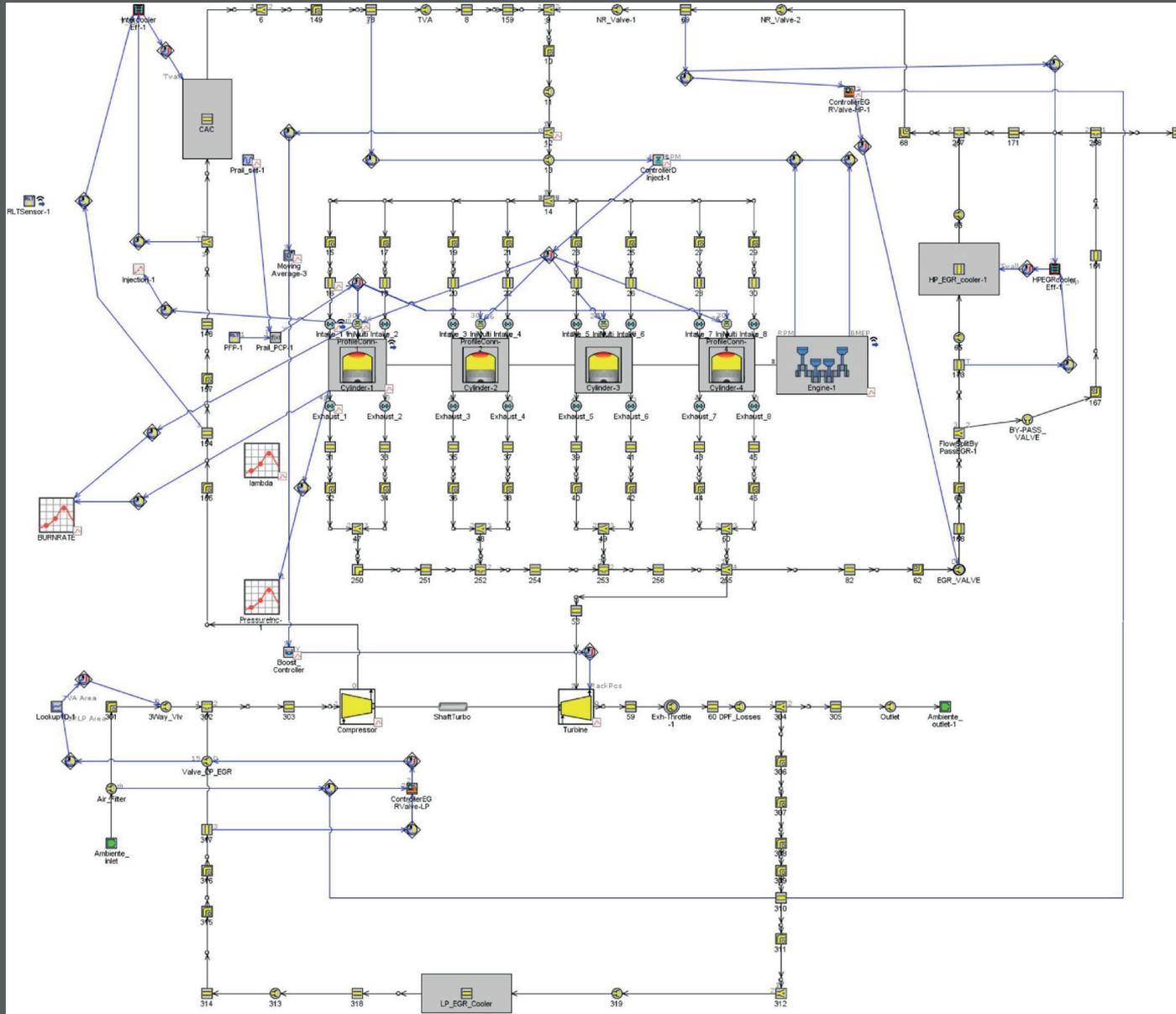


Figure 4.23 - Detailed model, Factor of Real Time is on average 50

Case	N [RPM]	BMEP [bar]
1	3750	13,7
2	3500	15,1
3	3250	16,5
4	3000	17,9
5	2750	19,2
6	2500	19,2
7	2250	19,2
8	2000	19,2
9	1750	19,2
13	3750	12
14	3500	13,2
15	3250	14,4
16	3000	15,6
17	2750	16,8
18	2500	16,8
19	2250	16,8
20	2000	16,8
21	1750	16,8
22	1500	16,8
25	2500	14,4
26	2250	14,4
27	2000	14,4
28	1750	14,4
29	1500	14,4
30	1250	11,3
32	3750	8,6
33	3500	9,4
34	3250	10,3
35	3000	11,2
36	2750	12

Case	N [RPM]	BMEP [bar]
37	2500	12
38	2250	12
39	2000	12
40	1750	12
41	1500	12
42	1250	9,4
44	3750	6,9
45	3500	7,6
46	3250	8,2
47	3000	8,9
48	2750	9,6
49	2500	9,6
50	2250	9,6
51	2000	9,6
52	1750	9,6
53	1500	9,6
54	1250	7,5
56	3750	5,2
57	3500	5,7
58	3250	6,2
59	3000	6,7
60	2750	7,2
61	2500	7,2
62	2250	7,2
63	2000	7,2
64	1750	7,2
65	1500	7,2
66	1250	5,6
68	3750	3,4
69	3500	3,8

Case	N [RPM]	BMEP [bar]
70	3250	4,1
71	3000	4,5
72	2750	4,8
73	2500	4,8
74	2250	4,8
75	2000	4,8
76	1750	4,8
77	1500	4,8
78	1250	3,8
80	3750	1,7
81	3500	1,9
82	3250	2,1
83	3000	2,2
84	2750	2,4
85	2500	2,4
86	2250	2,4
87	2000	2,4
88	1750	2,4
89	1500	2,4
90	1250	1,9
92	3750	0,7
93	3500	0,8
94	3250	0,8
95	3000	0,9
96	2750	1
97	2500	1
98	2250	1
99	2000	1
100	1750	1
101	1500	1
102	1250	0,8

Table 4.2 - Engine Steady-State working points

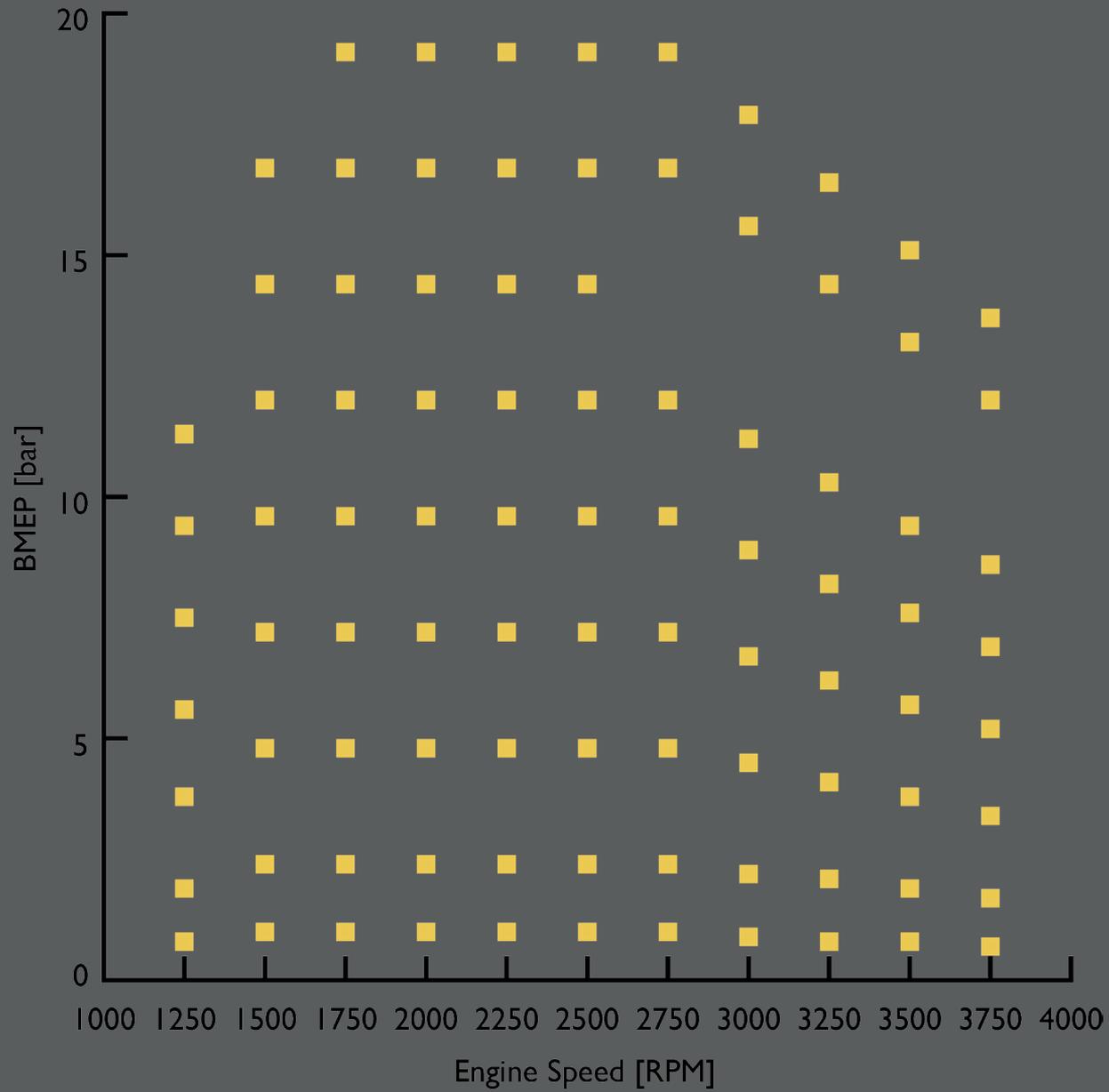


Figure 4.24 - Engine Map

4.4 Comparison between DETM and FRM

The comparison between the Detailed and the Fast-Running model is made possible because both the models perform the same simulation accordingly to the Engine Map described in Table 4.2 and the corresponding Figure 4.20.

The models have been tested to the following quantities:

1. Factor of Real Time	Turbine Average Speed .11
2. Brake Torque	Turbine Outlet Temperature .12
3. Air Flow Rate	VGT Controller .13
4. BMEP	Compressor Outlet Pressure .14
5. Volumetric Efficiency	Intercooler Outlet Pressure .15
6. Injected Average Mass Flow Rate	Intercooler Outlet Temperature .16
7. Intake Manifold Pressure	EGR Rate .17
8. Intake Manifold Temperature	EGR Valve Mass Flow Rate .18
9. Exhaust Manifold Pressure	EGR Controller .19
10. Exhaust Manifold Temperature	EGR Cooler Outlet Temperature .20

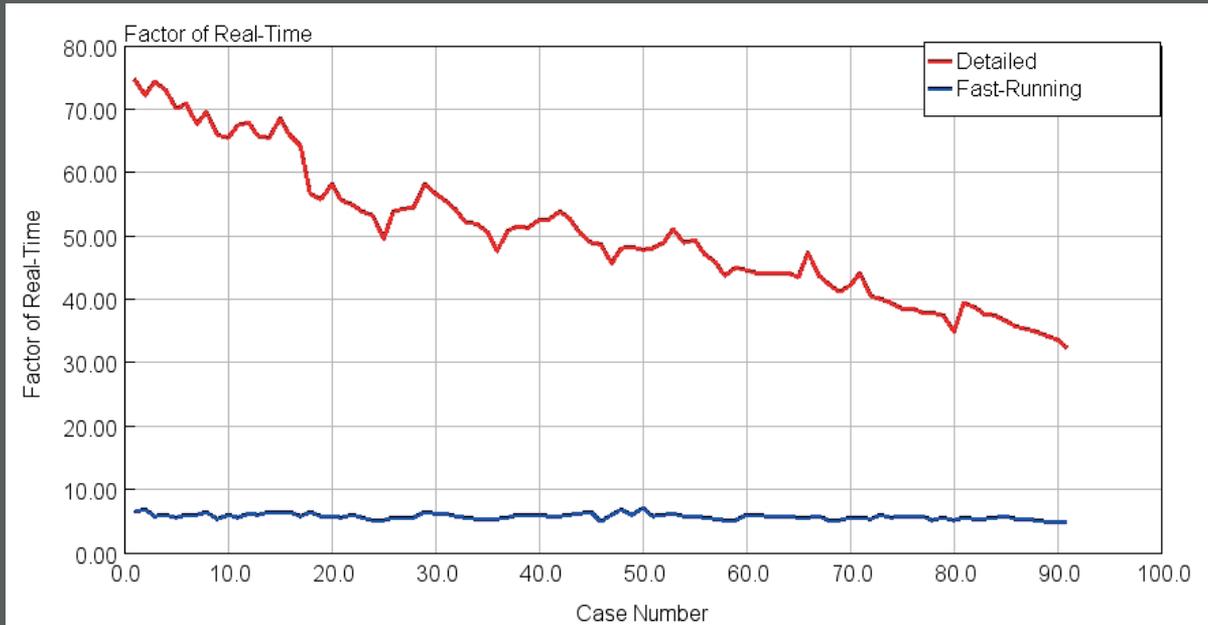


Figure 4.25 - FRT trend, DETM vs FRM

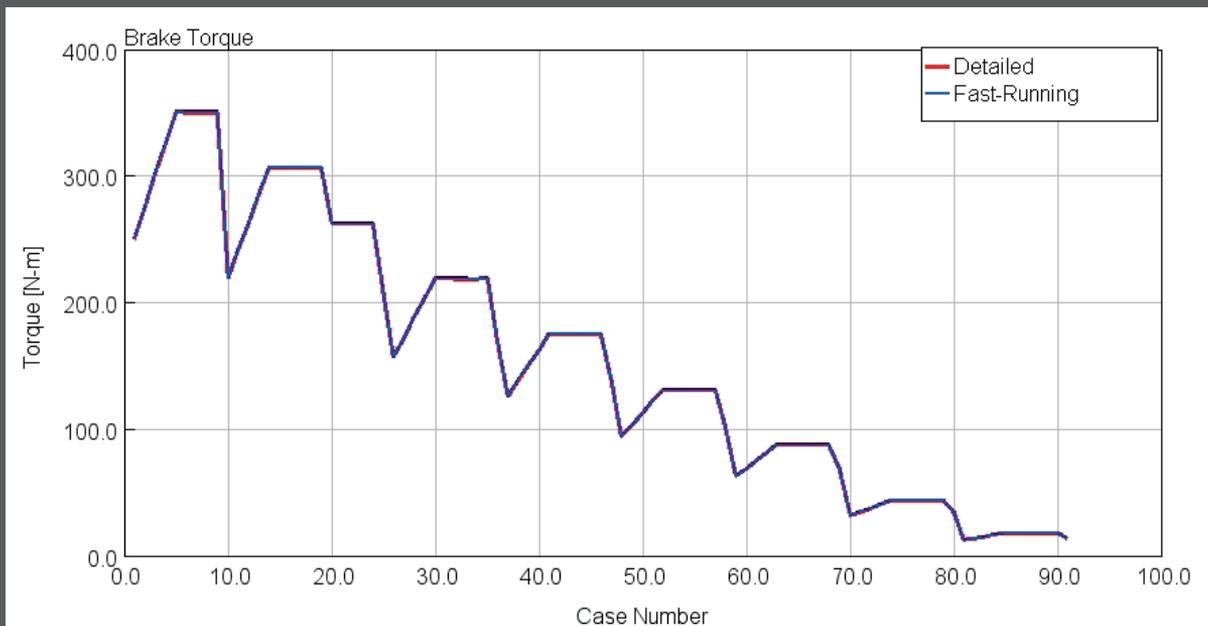


Figure 4.26 - Brake Torque, DETM vs FRM

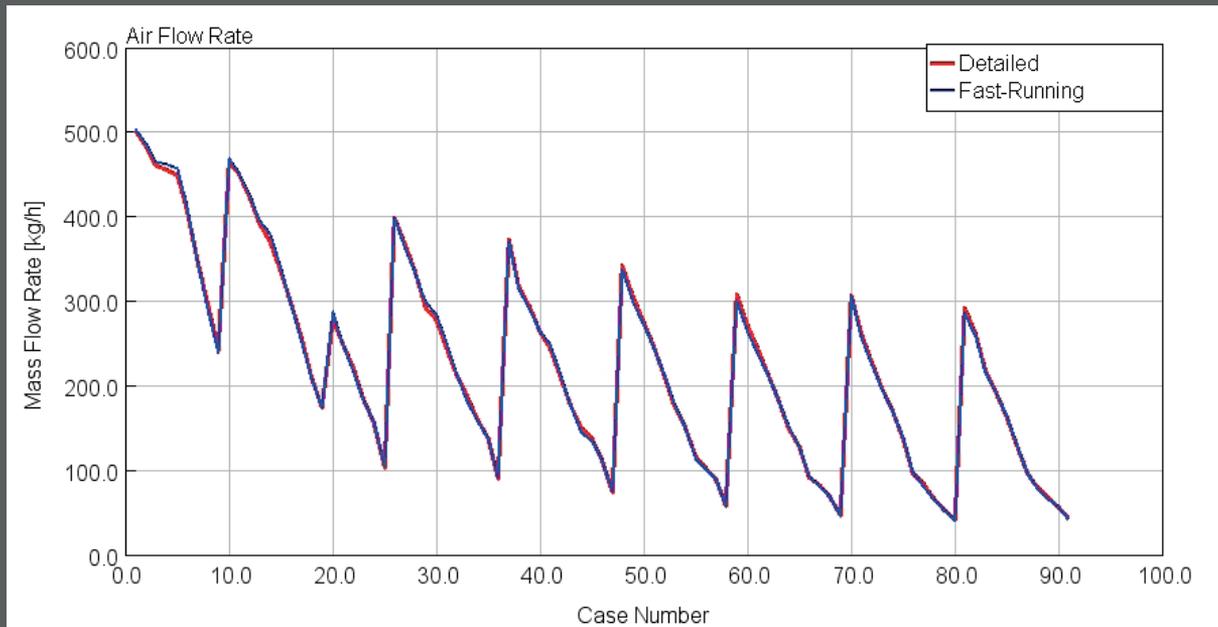


Figure 4.27 - Air Flow Rate, DETM vs FRM

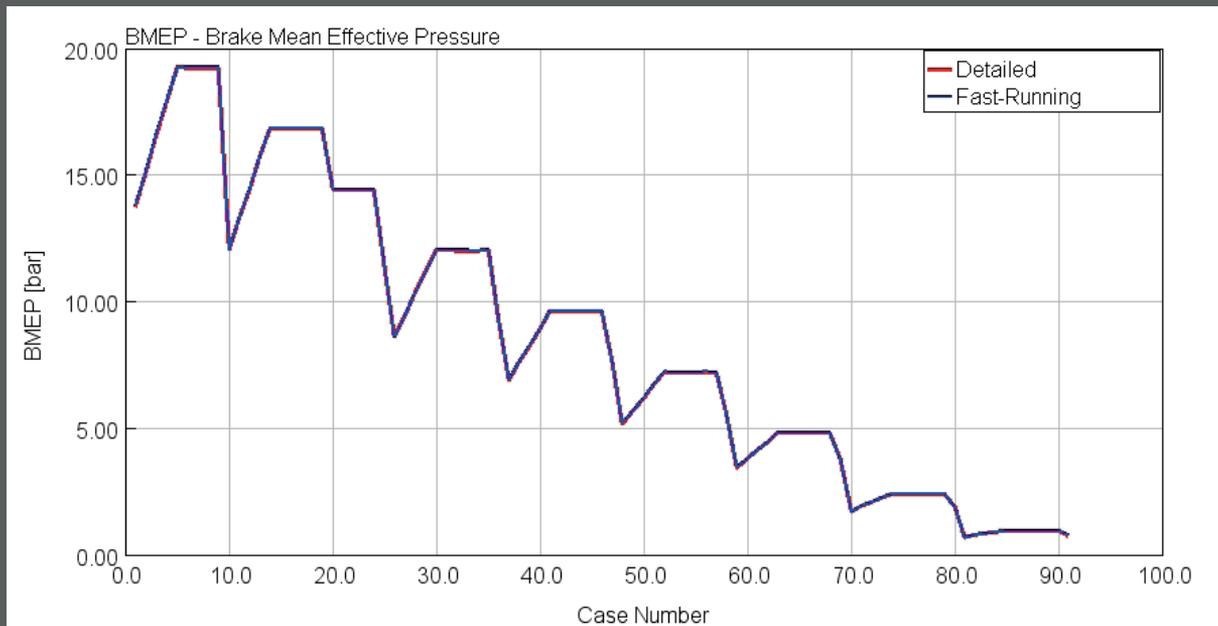


Figure 4.28 - BMEP, DETM vs FRM

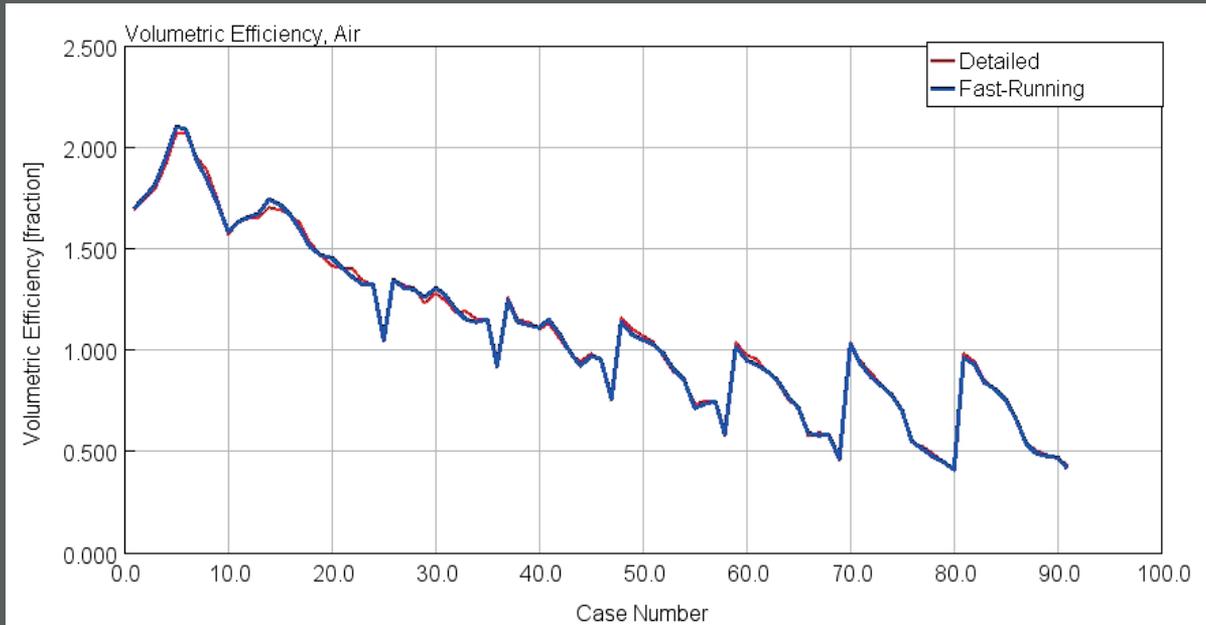


Figure 4.29 - Volumetric Efficiency, DETM vs FRM

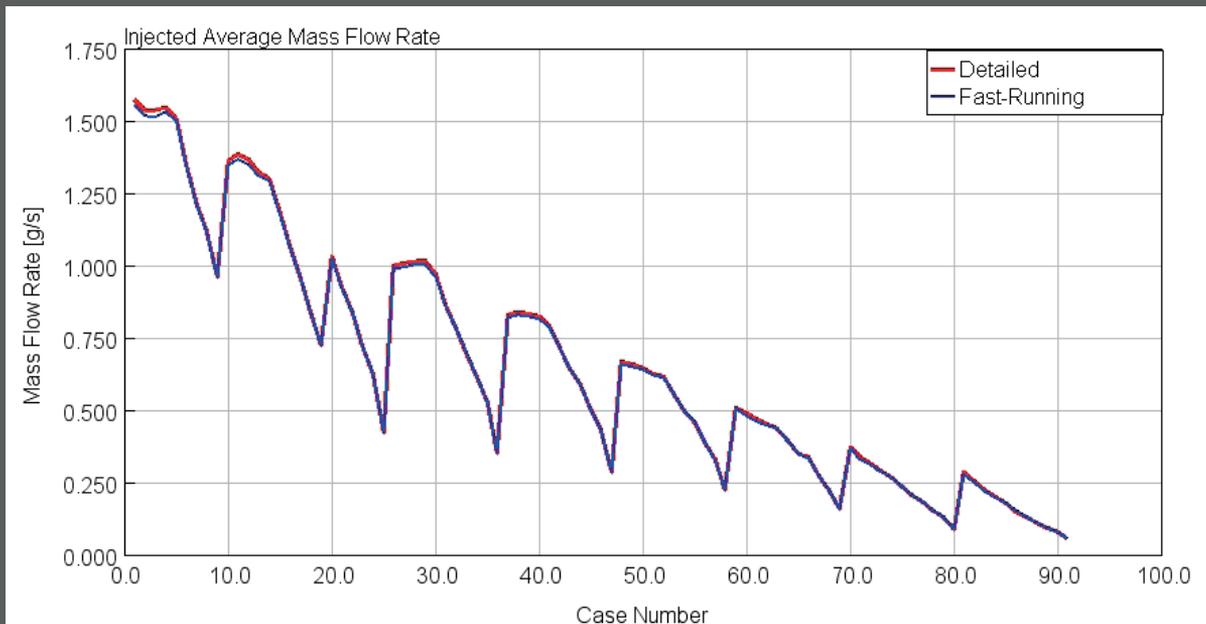


Figure 4.30 - Injected Average Mass Flow Rate, DETM vs FRM

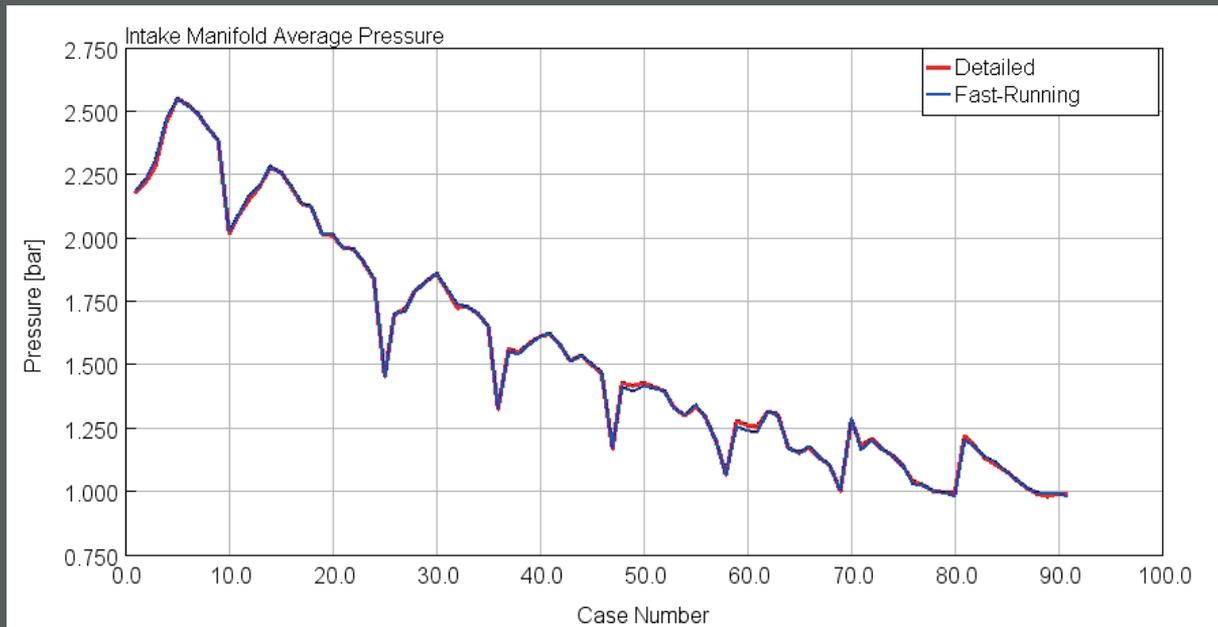


Figure 4.31 - Intake Manifold Pressure, DETM vs FRM

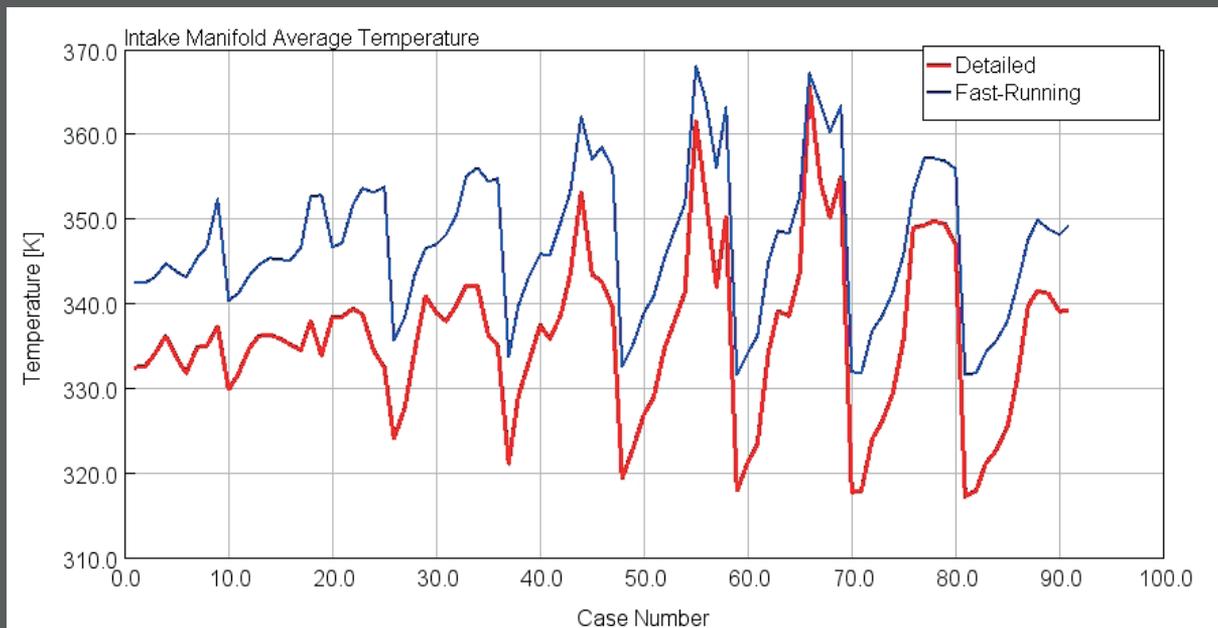


Figure 4.32 - Intake Manifold Temperature, DETM vs FRM

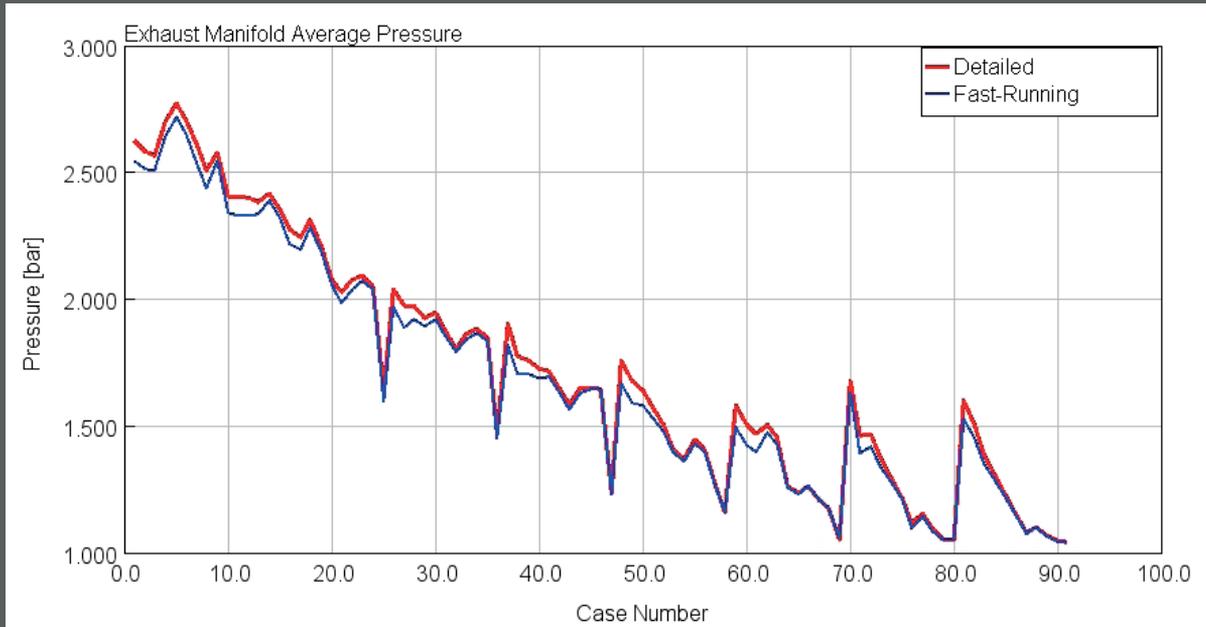


Figure 4.33 - Exhaust Manifold Pressure, DETM vs FRM

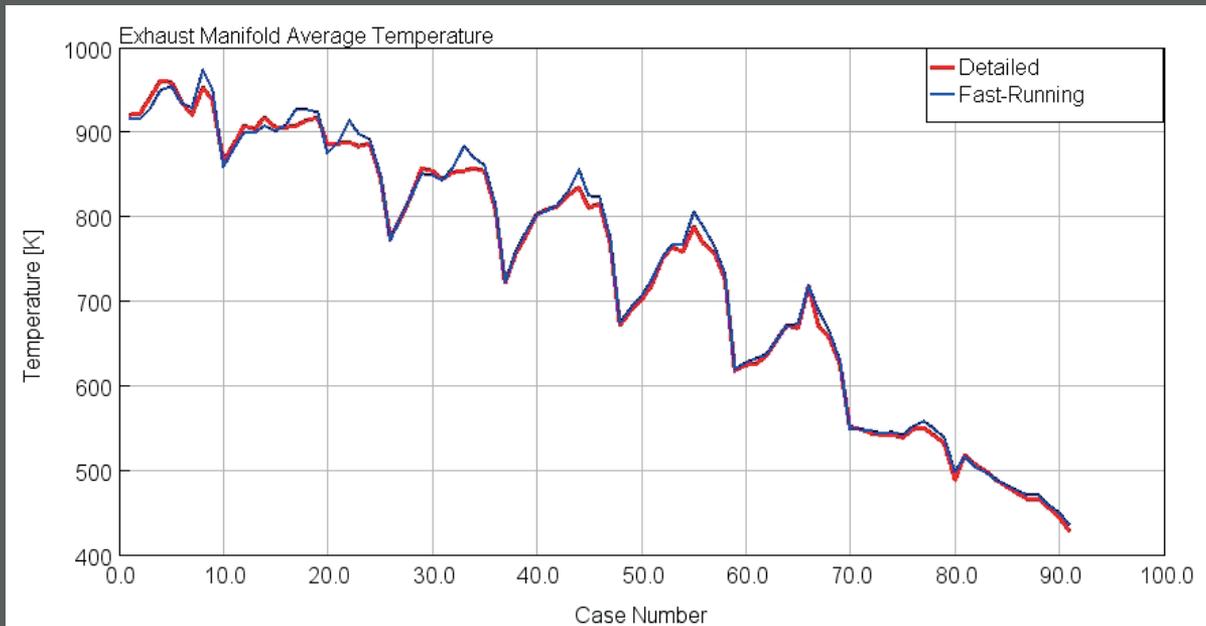


Figure 4.34 - Exhaust Manifold Temperature, DETM vs FRM

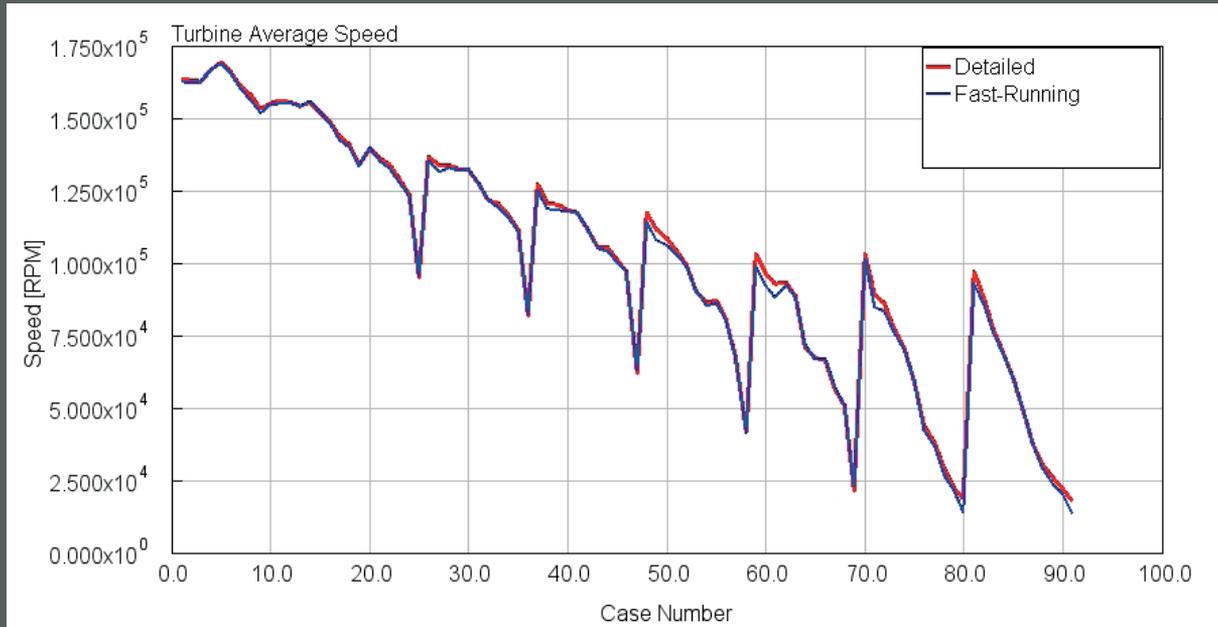


Figure 4.35 - Turbine Average Speed, DETM vs FRM

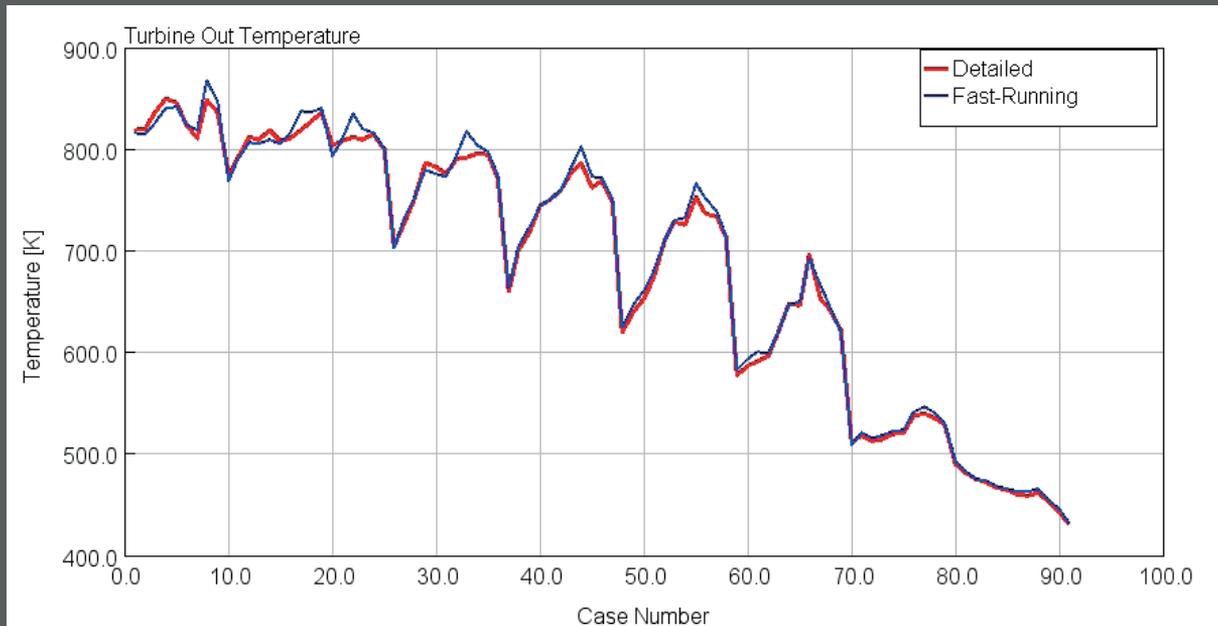


Figure 4.36 - Turbine Outlet Temperature, DETM vs FRM

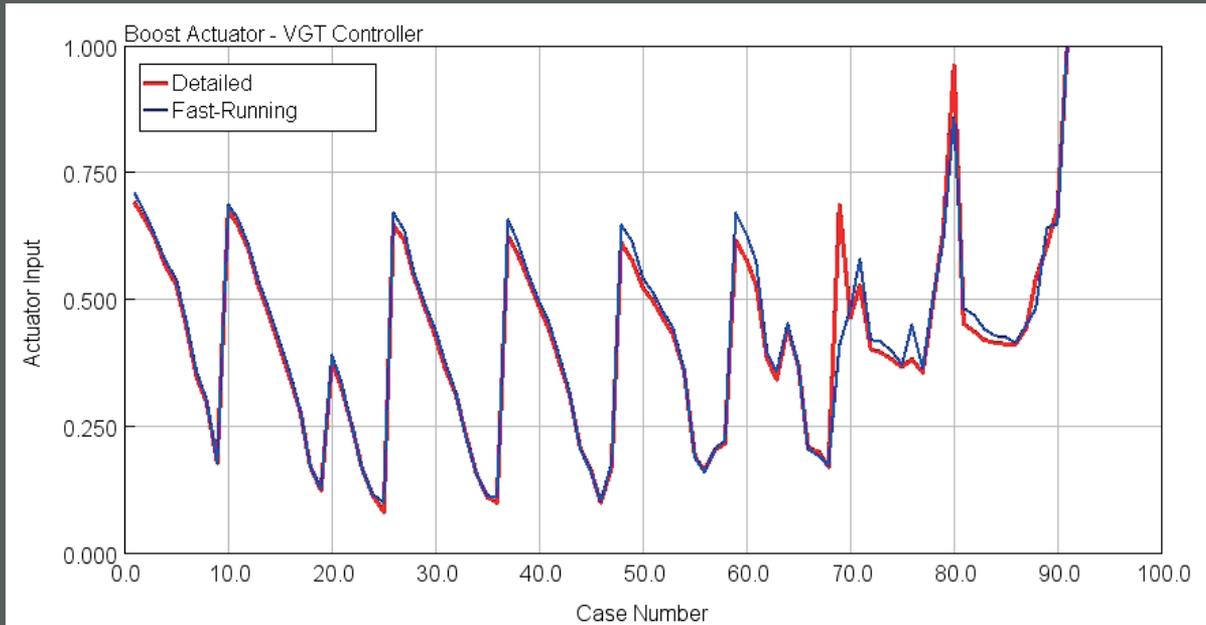


Figure 4.37 - VGT Controller (Boost Actuator), DETM vs FRM

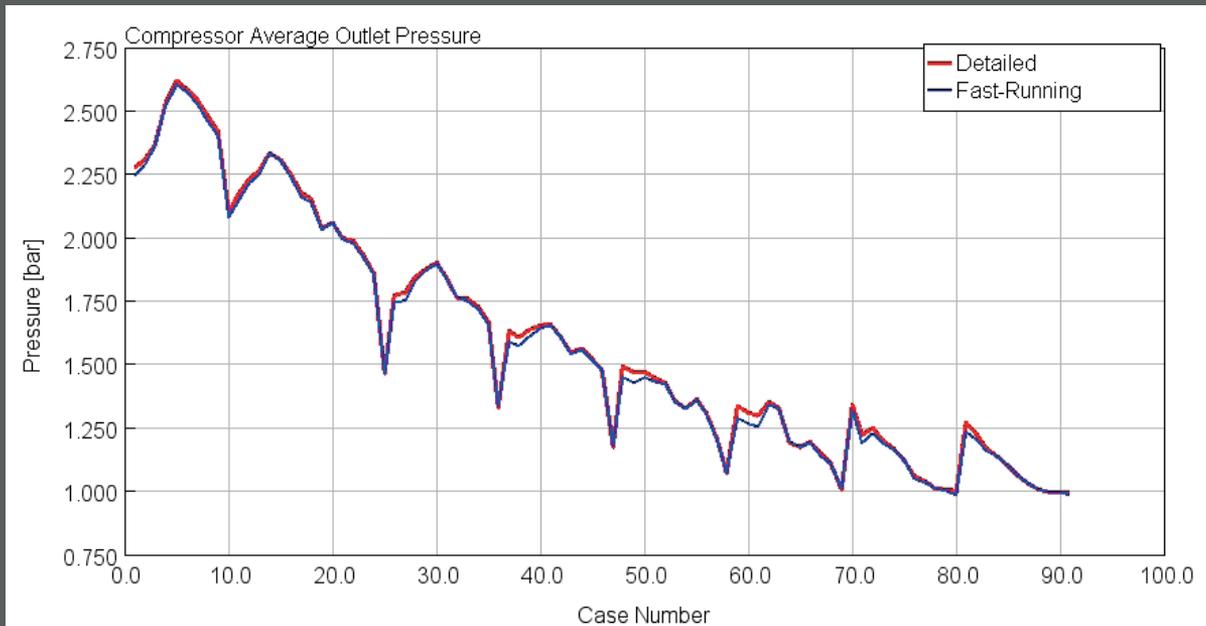


Figure 4.38 - Compressor Outlet Pressure, DETM vs FRM

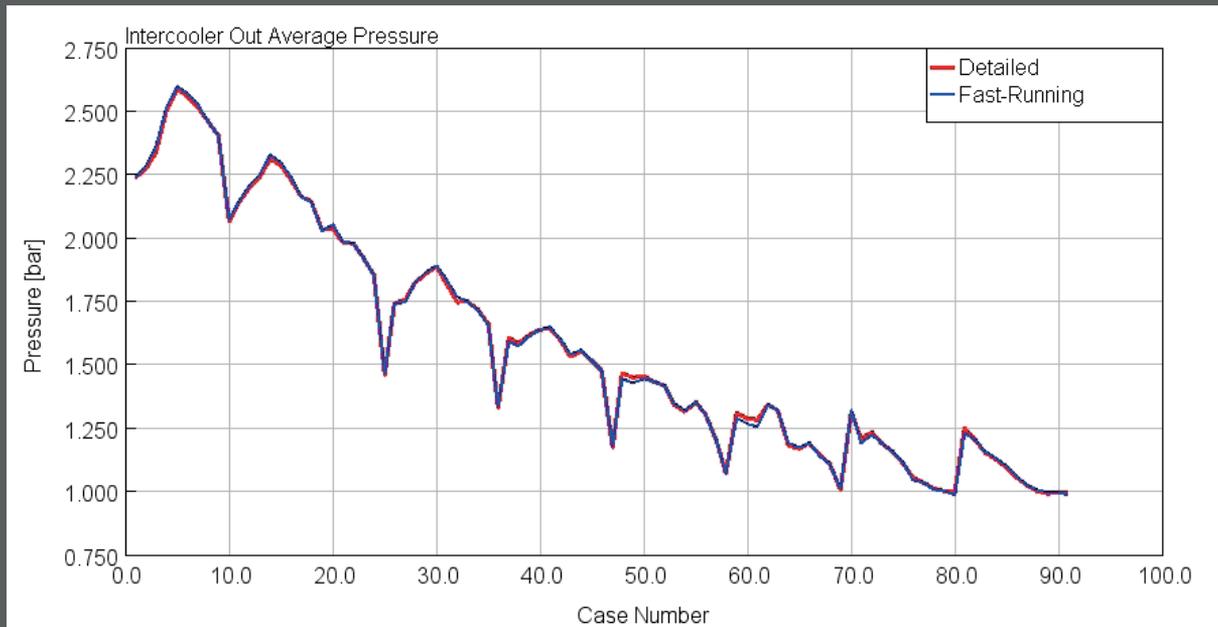


Figure 4.39 - Intercooler Outlet Pressure, DETM vs FRM

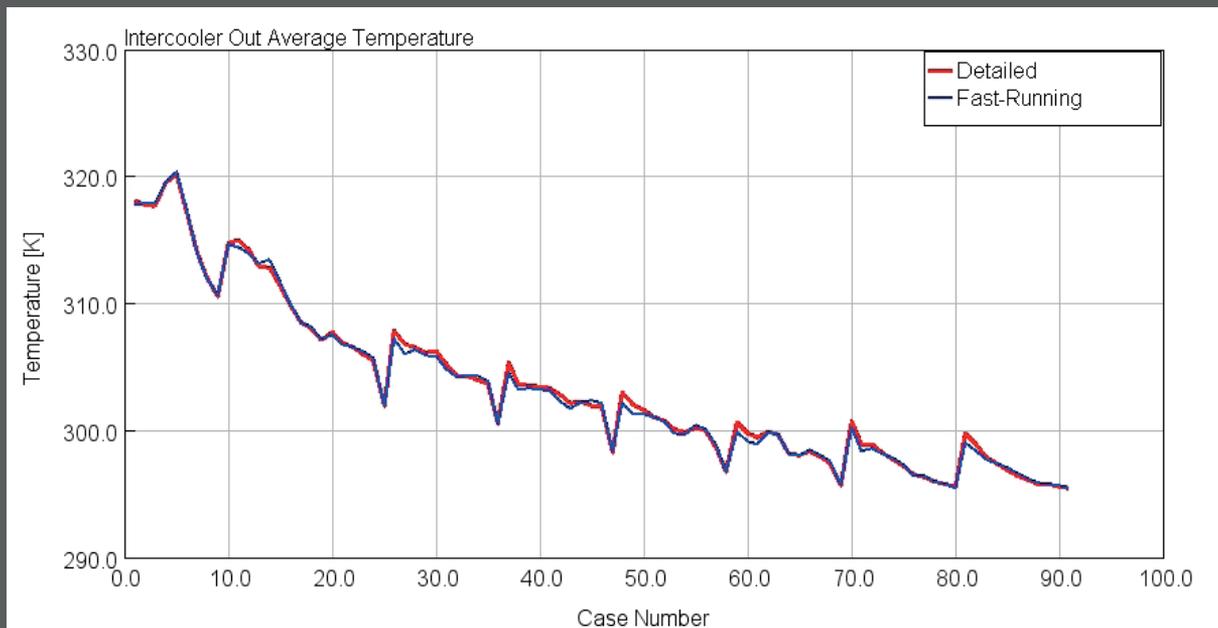


Figure 4.40 - Intercooler Outlet Temperature, DETM vs FRM

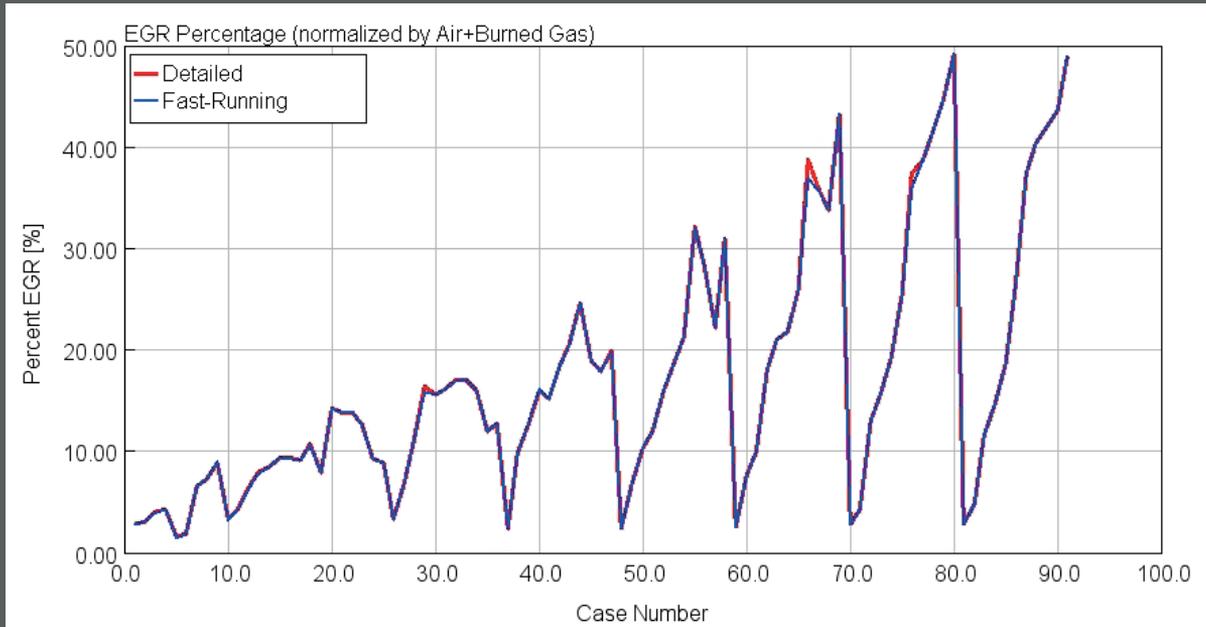


Figure 4.41 - EGR Rate, DETM vs FRM

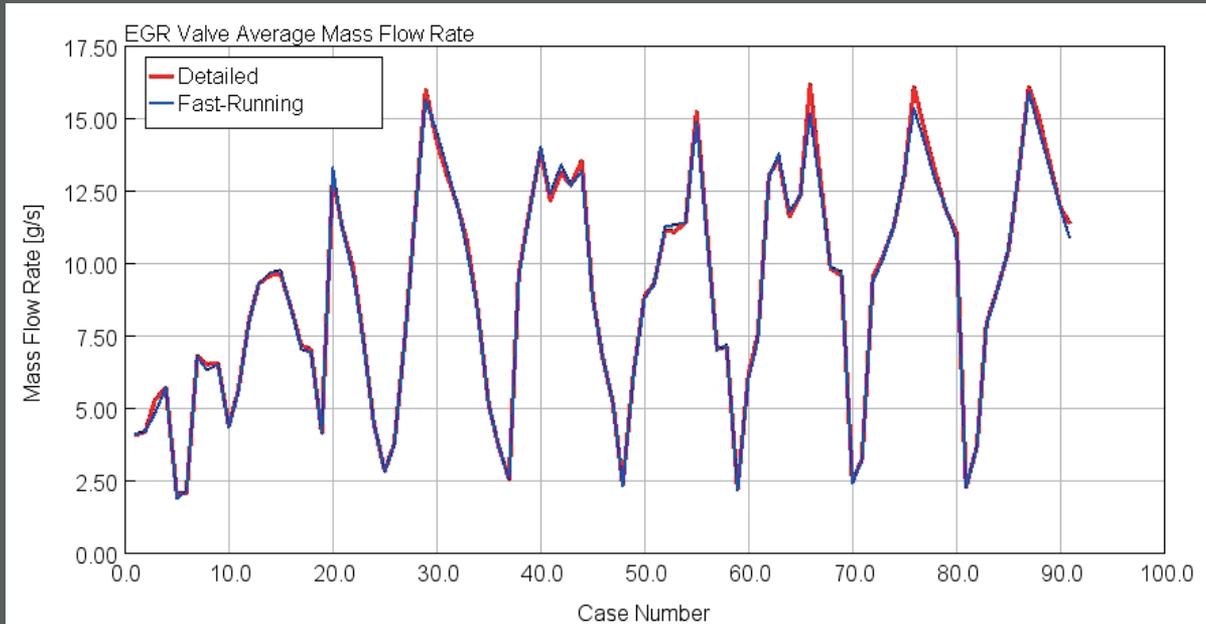


Figure 4.42 - EGR Valve Mass Flow Rate, DETM vs FRM

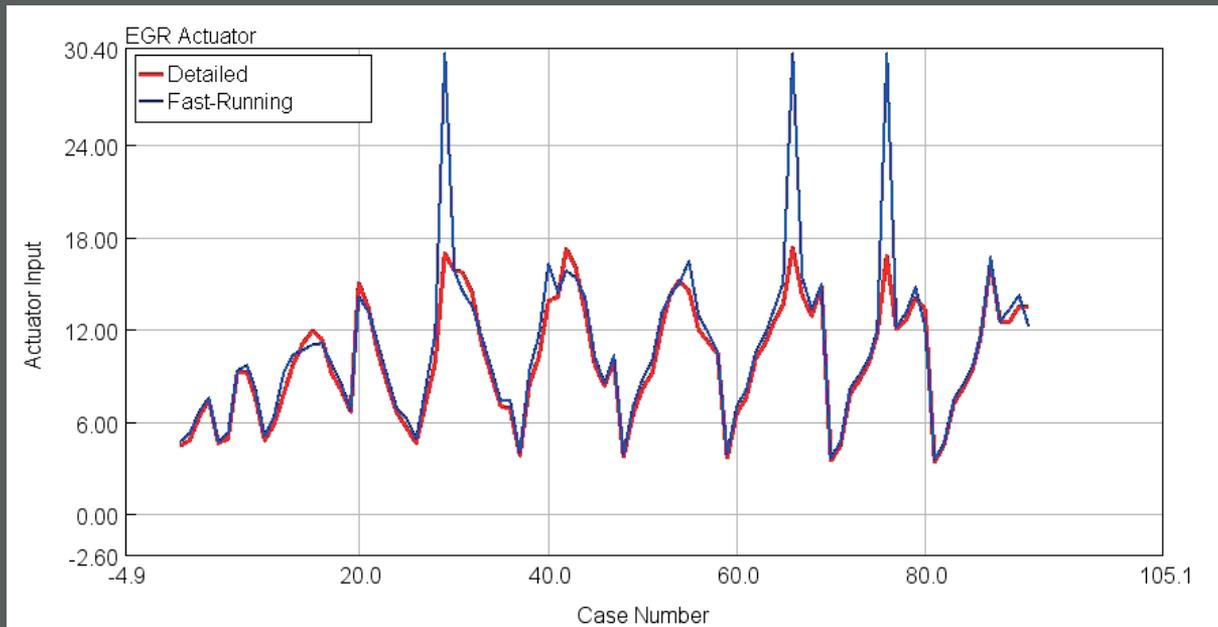


Figure 4.43 - EGR Controller, DETM vs FRM

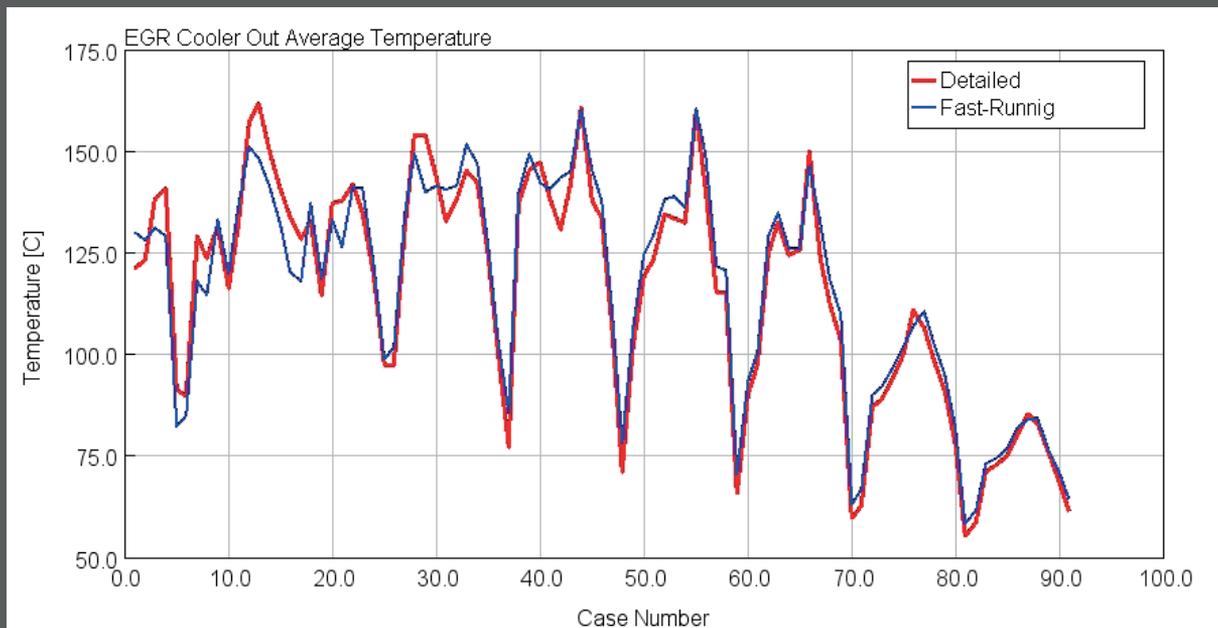


Figure 4.44 - EGR Cooler Outlet Temperature, DETM vs FRM



DESIGN OF CONTROL ALGORITHM

Control Theory

Combustion Process Control

Control Testing Procedures

The first task to perform after the development of the detailed model, is to develop the combustion controller in order to manage the pollutants emission and increase the engine efficiency. In addition, in the new combustion strategies, such as Low Temperature Combustion (LTC) and Premixed Charge Compression Ignition (PCCI), the combustion control is mandatory to counteract the instabilities of the process. The control and management is performed by the *brain* of the system which is the Electronic Control Unit (ECU).

The ECU works as follow: inside engine and transmission are present a lot of sensors measuring physical quantities required to manage the engine functioning, these amounts are converted into digital signals that will be sent to the engine control unit.

Then, the ECU processes and compares these input signals with a huge amount of look-up tables reporting target values of quantities to be controlled. Finally, the ECU sends the signals to the actuators in order to reach the desired output.

One of the key-parameter to control the combustion process is MFB50 (crank angle at which the 50% of fuel mass has burnt), as well as IMEP and maximum HRR. Gathering all these information, the ECU acts on the SOI parameter to regulate the combustion accordingly to the mission target.

This structure shows the advantage of a reduced amount of computational time required during engine functioning, but on the other hand, the calibration of these look-up tables required a huge effort and is very time and cost consuming.

Before analyzing the different techniques to control the combustion stability and quality, it worth mentioning the working principle of an general controller:

5.1 Control Theory

To control a dynamic system means to drive the system, also known as *plant*, with a command input u , so that the corresponding output y tracks any desired reference signal r :

In this way, it is possible to choose and impose the behavior of the system. Eventually, the target of a basic control problem is to find a controller C such that $y \approx r$ for a reference signal of interest.



Figure 5.1 - Scheme of a controller principle ^[2]

Control design can be seen as an inversion problem which means to invert a dynamic system, i.e. a set of differential equations. The two main approaches to control a dynamic system are:

- in the open-loop control, also known as feedforward, the controller is independent from the plant. For this reason, this controller is not very effective;
- in the closed-loop control, also known as feedback, the controller takes info from the plant, that is why this kind of controller is way more powerful and diffused.

5.1.1 Open-loop control

Open-loop control is mainly used whenever it is not possible to measure the output signal y . Major issues related to this strategy are based on the command input u which does not depend on the output y . Consequently, the controller is not aware of what actually is happening to the output y .

Since an open-loop system has no knowledge of the output condition, it cannot produce an automatic correction of the errors that it produces during its functioning despite the significant amplitude of the deviations with respect to the target value.

In conclusion, the open-loop control is not able to manage any disturbance that could occur during its functioning. For these reasons, it is not possible to stabilize an unstable system through the open-loop control approach.

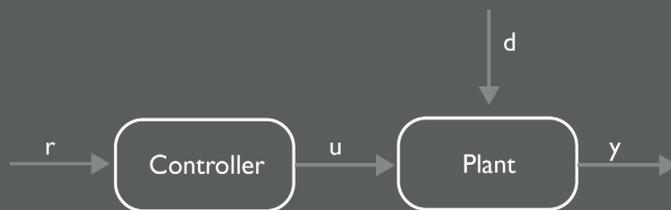


Figure 5.2 - Scheme of an open-loop controller ^[2]

The earliest open-loop controls evolved into the feed-forward controls. These last presents the capability to perform a reaction before that a hypothetical error will affect the system.

The objective of this kind of control, is to ability to do a *predictive control*, that is measuring or predicting any po-

tential open-loop disturbances and compensate for them before the controlled variable deviates too far from the original set point.

In a feed-forward system, the control variable adjustment is not error-based, but the on knowledge about the process to be controlled and knowledge about the process disturbances. The knowledge about the process is extracted from a mathematical model.

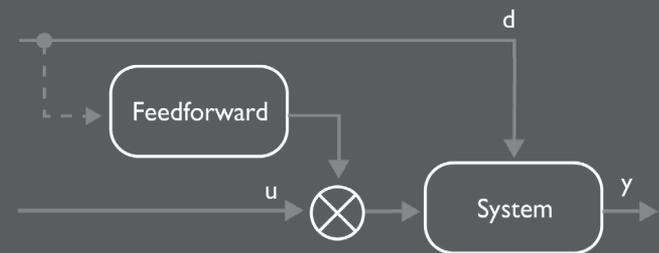


Figure 5.3 - Scheme of a feed-forward controller ^[2]

A reliable feed-forward control system requires that the effect of the output on the load is known and without variations.

The deepening in the study of feed-forward control systems was made possible by the invention of microprocessors, but it was seldom practiced due to the difficulty in the development of the mathematical models required.

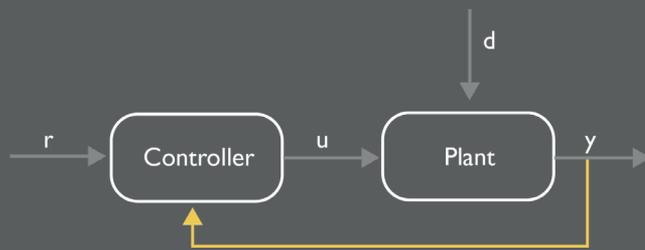
Control systems capable of adapting their mathematical model have become more practical as microprocessor speeds have increased.

To perform the feed-forward control it is required that the model is embedded inside the control algorithm. The model is used to determine the control actions accounting the known information regard the state of the system.

5.1.2 Closed-loop control

A different strategy is the Closed loop one, in which the control system has some pieces of information about the system state that has to be controlled. The required information are collected from the plant by sensors and then sent the controller, that performs a correction through a comparison between the system state and the required target value.

This means that In closed-loop the command u depends on the output y and, consequently, on the disturbance d . The controller corrects at each time the system behavior using the information coming from y .



v 5.4 - Scheme of a closed-loop controller [2]

A Closed-loop system is designed to reach and maintain the desired output by comparing it with the actual condition.

Differently, from Open-loop controller the Closed-loop one is able to manage the system disturbances because with this architecture the controller gets information regarding the output condition.

The error signal is the difference between the input signal and the feedback signal, and it is fed to the controller to reduce the system error and to bring the output close to the desired value.

Generally, the accuracy of the controller is influenced by the feedback path, that enables also the reduction of system's sensitivity to external disturbances, providing a more robust control and coherent output measured quantity.

5.2 Combustion Process Control

As previously stated, in-cylinder control strategies are needed for better exploitation of advanced combustion modes, such as PCCI.

The MFB50 is one of the most widely adopted metrics used in these types of controls as well as indicated mean effective pressure and maximum of heat release rate. Real time controls are divided into three big categories: map-based, pressure-based and model-based approaches and they will be discussed in the following sections.

The first control method is a pressure-based technique and it is based on a closed-loop approach: it uses the in-cylinder pressure to compute the corresponding value of MFB50 and then it performs a correction on the electric SOI of the main injection (SOI_{main}) in order to obtain a desired MFB50 target value.

The second control method is instead a model-based technique and it is based on a feed-forward approach: it uses a predictive model of the HRR to simulate the MFB50, then an optimal SOI_{main} to obtain the desired MFB50 target is identified by means of inversion of HRR model.

Eventually, the map-based control system is briefly presented. After discussing the different control techniques, the analysis then focuses on the control of the key-parameters such as the MFB50, torque and NOx control.

5.2.1 Pressure-based control

The pressure-based control is mainly used to control the MFB50 and it requires pressure sensor to be installed within each cylinder. This closed-loop control relies on the data provided by the sensor to correct the MFB50, instant by instant, by acting on the SOI_{main} . The center of gravity of the combustion is estimated basing on sensed pressure accordingly to the following mathematical model:

$$dQ_{net} = \frac{\gamma}{\gamma-1} p dV + \frac{1}{\gamma-1} V dp, \quad \gamma = \frac{C_p}{C_v}$$

in which p and V represent the pressure and volume within the combustion chamber. The values of SOI_{main} are corrected in real-time by the ECU with a cycle-to-cycle variation to match the target MFB50, basing on the scheme shown in Figure 5.5 [18].

The mathematical model return the trend of the Net Heat Release Rate (NHRR [J / °]) that gives, once normalized by its maximum, the of trend of the mass fraction burned. From this trend, one can extract the center of gravity of the combustion, that is, the MFB50.

In this way, the MFB50 is computed for the j -th cylinder at the i -th cycle, and such value is compared with the designed target. This technique returns the error expressed in the equation below:

$$Err_i^j = MFB50_{tgt,i}^j - MFB50_i^j$$

Thus, this error indicate how far the combustion is occurring from the designed conditions. Basing on this parameter, the ECU computes the value of the SOI of the following cycle ($i + 1$).

The SOI of the following cycle ($i + 1$) is corrected in accordance with the following equation:

$$Err_{main, i+1}^j = MFB50_{main,i}^j - K_{m,i}^j \cdot Err_i^j$$

K_m is a modulation factor needed to optimize the response of the controller and to maintain the system stability. Indeed, if K_m is constant, the system would experience high degree of instability.

More in detail, the factor K_m ranges in the set $[0.1, 1]$ and it is computed as function of the error :

$$K_{m,i}^j \in [0.1, 1]$$

$$\text{if } \text{sign}(Err_i^j) = \text{sign}(Err_{i-1}^j), \text{ then } K_{m,i}^j = K_{m,i-1}^j \cdot 2$$

$$\text{if } \text{sign}(Err_i^j) \neq \text{sign}(Err_{i-1}^j), \text{ then } K_{m,i}^j = K_{m,i-1}^j / 2$$

$$\text{if } Err_i^j > 3, \text{ then } K_{m,i}^j = 1$$

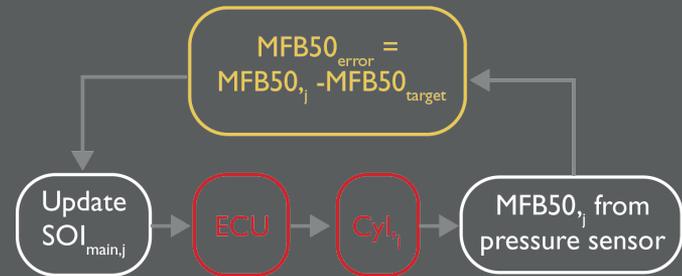


Figure 5.5 - SOI correction for pressure-based MFB50 control

5.2.2 Model-based control

The model-based control is an open-loop system that is based on the development of a mathematical model able to predict actual value of the quantities that have to be controlled. No additional hardware is required (i.e. lower cost) and main advantages rely on the reduction of time and calibration efforts.

With this kind of approach is implemented a predictive model that compute the required parameters instead of a direct or indirect measure of these quantities; for this reason it possible also consider the application of this control structure as virtual sensor especially for the emissions control tasks. Inside the control are then compared the output of our model with a reference value of the quantity to be controlled in order to evaluate an error quantity to be used for the correction of required quantities.

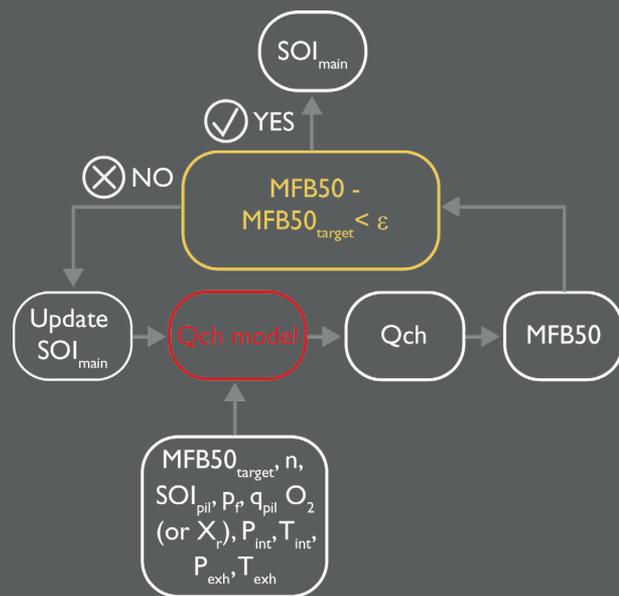


Figure 5.6 - SOI correction for model-based MFB50 control [18]

Concerning the combustion control, the control technique is based on the model inversion: by defining the optimal SOI value cycle-by-cycle, one obtains the designed value of the MFB50. To do this, the computational time is very low compared to CFD three-dimensional model.

This technique relies upon the Accumulated Fuel Mass model (AFM). This model tries to catch the link between the injection rate and the combustion itself.

This model is based on the assumption that the rate of released chemical energy Q_{ch} (real, gross HRR) is proportional to the energy associated with the fuel quantity available for combustion at the considered moment t .

Such energy can be computed at time t as the difference between the chemical energy Q_{fuel} associated with the fuel injected and the cumulative heat release Q_{ch} :

$$\frac{dQ_{ch}}{dt} = K[Q_{fuel}(t) - Q_{ch}(t)], \quad Q_{fuel} = \int_{t_{SOI}}^t \dot{m}_{f,inj} H_i dt$$

This means that the real HRR is proportional to the difference between the chemical energy associated to the injected fuel up to time t minus what has already been released. For instance, if the fuel has already released all the chemical energy, then the difference is nil and no more HRR is achievable.

The main advantage compared to combustion imposed approaches, is that the term $Q_{fuel}(t)$ embeds the information about the injection rate so the instantaneous HRR is analytically linked to the injection rate. Nevertheless, this model accounts for the coefficient K which is related to the velocity at which this difference between chemical energies is released.

The coefficient K was interpreted as a characteristic spray or atomization number. As K increases, the droplet size reduces and/or spray penetration is enhanced, thus the combustion rate is increased.

For a given spray, K is related to the swirl ratio or the turbulence, according to the assumption that the rate of the fuel oxidation is determined by the rate of mixing of the fuel vapor and fresh gas and thus by the local density of the turbulent kinetic energy. To sum up, the higher the K , the higher the velocity at which process occurs.

The AFM was updated later on in order to account for the multiple injections in Diesel engines according to the following equations:

$$\frac{dQ_{ch}}{dt} = K_1 [Q_{fuel}(t) - Q_{ch}(t)], \quad t_{SOC} \leq t \leq t_{SOCD}$$

$$\frac{dQ_{ch}}{dt} = K_2 [Q_{fuel}(t) - Q_{ch}(t)], \quad t > t_{SOCD}$$

The first equation refers to the Premixed combustion, whereas the second refers to the Diffusive combustion (SOCD stands for Start of Combustion - Diffusion).

More recently, the model was even more finalized (Patent GB2481364). Starting from the AFM approach, the atomization number was implemented for different injection events:

$$\frac{dQ_{ch}}{dt} = K_j [Q_{fuel,j}(t - \tau_j) - Q_{ch,j}(t)]$$

The first main difference is that the atomization number K must be split into the premixed and diffusive contribution, while the second difference is the reference time of $Q_{fuel,j}$ which is no more computed at time t but considering the delay $t - \tau_j$.

The chemical energy is no more proportional to the difference between the chemical energy released by the fuel at time t minus the total chemical energy released up to that time, because the chemical energy associated with the injected fuel is estimated earlier.

This because the fuel injected does not immediately release its own chemical energy owing to the ignition delay.

Previously, τ was embraced by the characteristic number K , whereas now the parameter $t - \tau_j$ considers such delay per each injection. Here, coefficient K is only influenced by the spray and turbulence effect.

To correctly set both τ_j and K , the model needs to be trained before to be applied. The calibration is performed for different points on the operating map.

In general, it is possible to find correlation between the AFM parameters (τ_{pilot} , τ_{main} , $K_{main,pre}$, $K_{main,diff}$, $\Delta t_{main,diff}$) and the specific engine operating parameters. Usually, 10 to 20 points are necessary to calibrate the model.

5.2.3 Map-based control

Currently, the control of Diesel engines is for the most map-based, i.e. quantities are interpolated from look-up tables that are mainly determined by the evaluated fuel injected quantity (which is related to the engine load), the engine speed and other measured quantities, such as the intake air mass and the boost pressure.

This control strategy determines the different actuations to achieve the desired target, however it requires a pervasive calibration.

In the last few years, the increasing computational performance of modern ECU has allowed the implementation of more and more sophisticated model-based approaches in IC engines. The main advantages from map-based to predictive model-based control are:

- lower calibration effort,
- wider range of operations,
- dynamic calibration,
- onboard real-time optimization of engine parameter oriented to pollutant emission reduction and fuel consumption saving.

The parameters for combustion controllers usually refer to MFB50, torque and NO_x control.

At this stage, it is required to merge all these information and realize an integrated controller able to follow the torque requested by the driver, still minimizing the fuel consumption without exceeding the nitrogen oxide level closed out by environmental legislations.

MFB50 Control

The closed loop pressured-based technique is an easy straight-forward way to compute and control MFB50. The pressure transducer determines the trend of the pressure as function of the crank angle basing on the single zone estimating function:

$$dQ_{net} = \frac{\gamma}{\gamma-1} p dV + \frac{1}{\gamma-1} V dp, \quad \gamma = \frac{C_p}{C_v}$$

where dQ_{net} is the apparent heat release rate. Once dQ_{net} has been computed, the MFB50 is the gravimetric center of the overall combustion.

However, the limits of closed loop controllers about MFB50 are:

- high cost of transducers,
- sensors deterioration over time.

The lower cost and the reliability of current models are restoring trustworthiness to open loop controllers. Indeed, the open loop control based on the AFM model (patent GB2481364) combined with a single-zone approach for pressure prediction, has proved to be suitable for the implementation in the ECU for feed-forward combustion control (real-time simulation) called LT-AFM.

This model-based approach directly estimates the rate of chemical energy and thus MFB50 provided the combustion rate coefficient K and the ignition delay τ .

With the pressure-based approach, the MFB50 of each cylinder matches that of the target with very low cylinder-to-cylinder variation. This result is possible because in the map-based control, the SOI is fixed and equal for all

cylinders in opposite to the closed loop which takes action on the individual SOI to match the MFB50. In general, the pressure-based approach enables a control of MFB50 cylinder by cylinder acting on the SOI of each cylinder. In this way, the MFB50 variation from cycle-to-cycle is dramatically reduced with respect to the map-based approach.

Considering the comparison between a model-based and a map-based controller, this last cannot regulate cylinder by cylinder, or rather, it can correct cylinder by cylinder, but the model does not distinguish one cylinder from the other. At nominal conditions one expects that the model-based approach behaves like map-based controller. The main advantage is that whether the environmental conditions change, the model-based controller takes care of these differences, and consequently adjusting the SOI in a way much more easy and flexible rather than the map-based controller.

	Map-based	Pressure-based	Model-based
Control loop	Open-loop	Closed-loop	Open-loop
Mechanism	Calibrated 2D Map	Pressure measurement	0D Predictive Combustion Model
Additional Costs	No	Pressure Sensors	No
Cylinder-wise MFB50 Control	No	Yes	No
Control Delay	No	At least one cycle	No
Computational Power Demand	Low	Low	High
Control Accuracy	Medium	High	Medium
MFB50 Target	Fixed	Flexible	Flexible
Calibration Effort	High	Very Low	Low

Table 5.1 - Comparison among MFB50 Control technique

Torque Control

Again, the torque control is usually map-based translating the position of the accelerating pedal into torque request. The pedal position information is transferred to the torque-to-fuel map which acts on the energizing time of the injectors.

If the system is provided of a pressure transducer, it is also used to estimate, in addition to MFB50, also the IMEP which is related to the torque. This because the pressure is still the input coming from the direct measurement:

$$\text{IMEP} = \frac{\int_0^{720} p dV}{V_{\text{cyl}}}$$

In an open loop model-based control, the MFB50 was regulated accordingly to the AFM that provides the estimation of the chemical energy released. The pressure trace is reconstructed from the trend of MFB50 through the single zone.

So, as for all combustion models, the chemical energy represents the input and the pressure is the output. In this way the pressure trace is obtained and consequently the torque is derived.

$$dp = \left(\frac{\gamma-1}{\gamma} \right) \left(dQ_{\text{net}} - \frac{\gamma}{\gamma-1} p dV \right)$$

Briefly, the torque control, generally performed acting on the injected fuel quantity, can substitute the traditional map-based approach. Using the closed loop approach pressure trace can be read from sensors inside the combustion chamber. Using open loop approach, the pressure must be virtually estimated/reconstructed.

NOx Control

NOx control allows flexible regulation that influence of course also engine efficiency. This flexibility gives advantages in real driving condition, especially if a energy management system is present. Closed-loop control requires NOx sensors which are different from those implemented in the aftertreatment systems as for SCR devices.

For this reason, and because the model-based NOx prediction are evolving, currently OEM are interested in an open-loop comeback. However, open-loop approach needs to be very accurate and NOx production modelling is not easy to manage due to their sensitivity to a greater extent of boundary conditions. Also, the quality of the input data (kept from ECU) that the model uses affect the overall results.

One opportunity is the simplification of the thermal mechanism of NOx, however the calculation time is still long for real ECU. This model is based on the estimation of NOx with respect to MFB50 and rail pressure.

This semi-empirical model still requires calibration which is engine-specific, but the approach is quite general and easily extended from one engine to the other.

The maps can be obtained experimentally provided the good prototype of the engine, or by modelling using the AFM together with the thermodynamic multizone and thermal NOx model run in commercial software such as GT-Power Engine Simulation Software, and then refined through the prototype.

5.3 Control Testing Procedures

The control testing procedure refers to the development and implementation of the controllers. The target is to test the designed control algorithms from all virtual environments up to real engines with physical ECUs.

The design of control algorithms actually means to develop the logic of the ECU. This process presents a high level of complexity. One of the most critical points is the definition of computational time required for the application of new and heavy control's strategies.

This parameter is discriminant for the final application of the code because must be compared with the hardware computational power available.

The principal steps required by the procedure that allows a correct and robust design are divided in three phases:

1. MiL Model in the Loop
2. HiL Hardware in the Loop
3. RP Rapid Prototyping

The demand of high robustness is caused by the high complexity reached by vehicle control systems. The number itself of ECUs is dramatically increased during the last years; another demanding point, considering an industrial case, is that the development process is frequently spread across the supply chain.

So it is possible to face problems of integration between dissimilar codes developed through different software environment.

5.3.1 Model in the Loop

In this first stage, the models of both the controller and the engine are simulated, i.e. they are virtual. The controllers are developed and implemented in Matlab & Simulink whereas the engine is modeled in GT-Power.

Later, the Simulink model are coupled to the GT-Power model. This last must be a Fast-Running engine model (discussed in paragraph 4.3) which is the simplified detailed model needed to reduce the computational time.

The Simulink information are exchanged in GT-Power differently depending on the strategy of the controller: pressure-based or model-based.

In the first case, the controller receives the information concerning the MFB50 for the specific operating point on GT-Power and realizes the tuning of the SOI_{main}^j according to the equations:

$$Err_i^j = MFB50_{tgt,i}^j - MFB50_i^j$$

$$Err_{main,i+1}^j = MFB50_{main,i}^j - K_{m,i}^j \cdot Err_i^j$$

and sends the new SOI information back to GT-Power.

In the second case, the Simulink block receives the data for the definition of the HRR, estimates the SOI_{main}^j to obtain the optimized MFB50 and sends back its value to GT-Power.

MiL phase tests the functionality and effectiveness of the controller, improving its response offline (money saving reducing experimental tests) and verifying the safety (reduction of the risk of engine damage for improper/unexpected behavior).

5.3.2 Hardware in the Loop

Hardware in the Loop phase provides that the engine model runs in a engine emulator hub, and the virtual controller is substituted by a real ECU which still receives the Simulink model as an input by an external device.

Thus, in this phase, the goal is to interface the virtual engine model with the properly tuned ECU to see if the controller can run on a real hardware.

This test phase includes the following steps:

- development of the **GT-Power Mean Value Model**;
- integration of the control software developed in Matlab Simulink on the rapid prototyping device (for example, ETAS ES910) using INTECRIO software;
- integration of the Mean Value Model on a PXI (PCI eXtensions for Instrumentation), which is a device constituted by a controller and many input/output acquisition boards, so as to reproduce the I/O signals of the sensors installed in the real engine;
- coupling of the rapid prototyping device with the PXI, which represents a real-time engine simulator.

The Hardware in the Loop has the aim of testing the rapid prototyping device and its functionalities by coupling it with a real-time engine mode.

When the HiL test is validated, the engine model is substituted with a real engine no doubts that the controller can run correctly in a real engine.

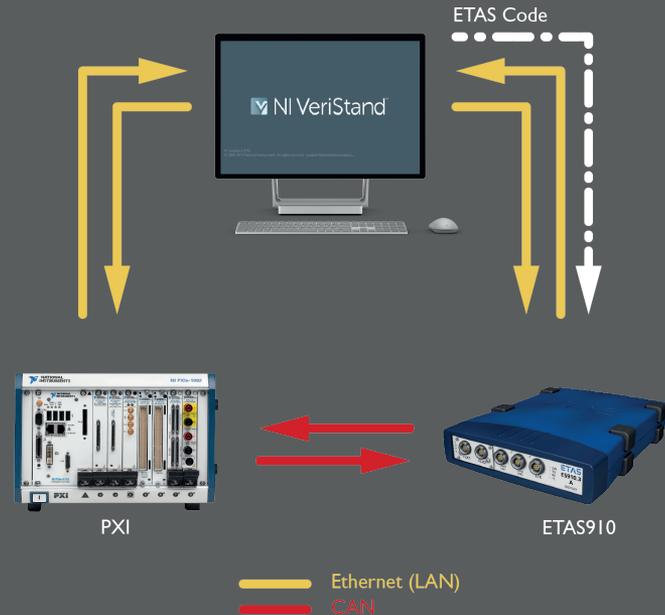


Figure 5.7 - Hardware in the Loop set-up ^[18]

5.3.3 Rapid Prototyping

The last phase of control implementation is Rapid Prototyping. In this phase, the rapid prototyping device is connected to the engine control unit, in order to by-pass the standard ECU functions (es. SOI_{main}) with those deriving from the models implemented in the RP device. This the output from the HiL phase is now connected to the real engine. This aims at verifying the new control functionalities in the real engine, it allows to test the controller without the need to implement it in a real ECU and to exploit the standard ECU calibration for all the other functionalities.

Phase	ICE	Controller
MiL	GT-Power FR Model	Matlab-Simulink
HiL	GT-Power RT Model	ETAS RT code
RP	Real Engine	ETAS RT code

Table 5.2 - Control testing procedures

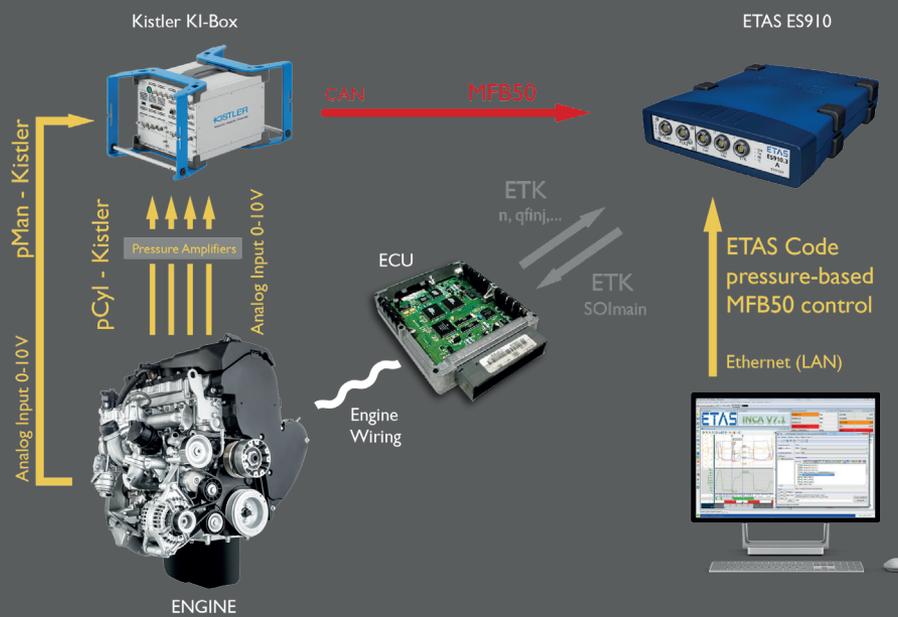


Figure 5.8 - Rapid Prototyping set-up for pressure-based controller [3]

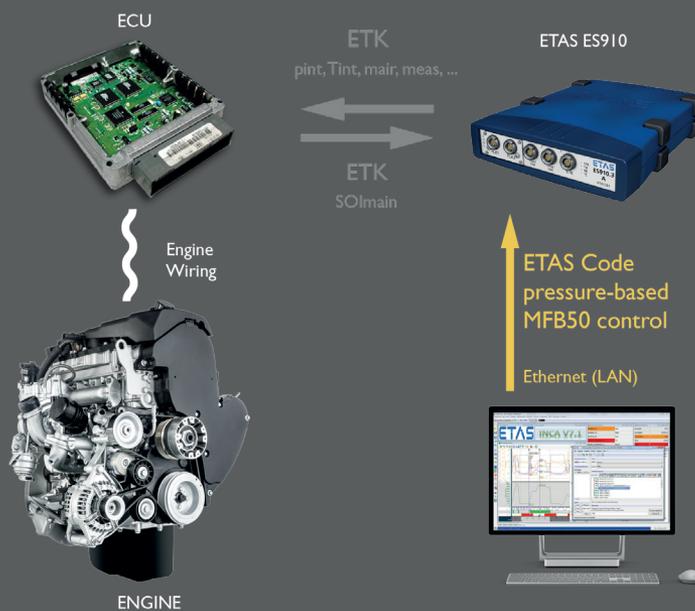


Figure 5.9 - Rapid Prototyping set-up for model-based controller [3]

A large, bold, yellow number '6' with a dark grey outline, positioned on the left side of the slide.

THE MEAN VALUE MODEL

Design of Experiment - DoE

Training of the Neural Network

From FR to MV Model

Integration of MVM in HiL testing

While Fast-Running Models (FRMs) are the preferred option to have a fast running engine model for system level simulations (MiL), it can be necessary to further reduce the fidelity of a model, but in turn, to increase the response time of the simulation. Indeed, the main purpose of a Mean Value Model (MVM) is to reduce the Factor of Real Time (FRT), that is the ratio between the simulation time and the real time. A FRT lower than the unity is necessary for the implementation of the model on the PXI hardware to perform real-time HiL testing. The core of this entire work is the development of such Mean Value Model.

Mean value engine models are useful for certain types of modeling where simulation speed is of primary importance, the details of wave dynamics are not critical, and bulk fluid flow is still important (for example, the modeling of the turbocharger lag.).

MVMs contains a map based cylinder model that is computationally faster than a regular (detailed) cylinder. The simulation speed can be increased further by combining multiple detailed cylinders into a single mean value cylinder.

In addition, many of the other flow components from the detailed model can be combined to create a simplified flow network of larger volumes.

This single mean value cylinder is the basis of the MVM and it is simply a map-based cylinder. The three maps that determine the mean value cylinder performance are:

- Volumetric Efficiency,
- Indicated Mean Effective Pressure - IMEP,
- Exhaust Gas Temperature.

Each of these three quantities is an input to the mean value cylinder and is imposed by the cylinder during a simulation. So to build a realistic MVM, it is necessary to define each of these three quantities as a function of other variables. These variables are named *input variables*. Therefore, the first step in building a MVM is to determine on which input variables the MV cylinder quantities will depend on.

Generally, volumetric efficiency should nearly always be a function of engine speed (unless a constant speed engine) and intake manifold pressure. For specific purposes, also the intake manifold temperature, exhaust manifold pressure, valve event timing and profile, EGR rate may be considered as input variables.

IMEP and Exhaust Gas Temperature will usually be dependent on the same variables as they are both related to distribution of fuel energy. These variables are the trapped air mass in the cylinder (and so containing at a minimum the same dependencies as volumetric efficiency), injected fuel mass (or A/F) and injection timing (for direct injection).

For the analysis of the FIA PCCI engine, the input variables are:

- Engine speed,
- Intake manifold pressure,
- Intake manifold temperature,
- Exhaust manifold pressure,
- Fuel mass rate,
- EGR fraction,
- SOI of the main pulse.

6.1 Design of Experiment - DoE

In general, the experiments are carried out to study the performances of systems and processes. A process may be intended as a combination of calculations that transforms input data into an output affected by one or more variables. The targets of an experiment usually are:

- to determine the most affecting parameters on the output y ,
- to establish a method to determine such parameters so that the error of y is as small as possible, also causing minimum disturbances of the uncontrollable factors on the output y ,

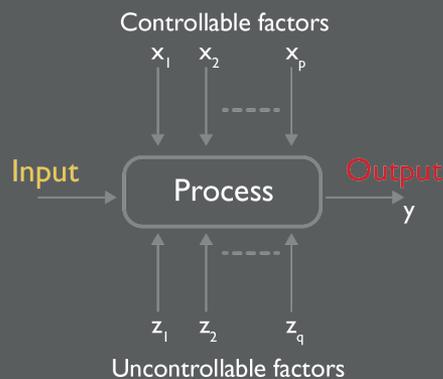


Figure 6.1 - Scheme of the variable in a generic process ^[2]

The Design of Experiment (DoE) is the technique to optimize the experiments. The experimental design is used to create suitable experimental conditions to provide the best results. Thus, the target is manage the input variables so as to provide a certain number of experiment each of them with different condition. The solution of the DoE determines a map needed to train the Neural Network that feeds the MV cylinder: v

Usually, the experimental measurements and results are subjected by disturbances. For this reason, it is preferable to engage statistical methods.

The core principles of statistical methods are:

- repetition,
- randomization,
- tests clustering.

On one hand, it is important that the tests are repeated, because in this way the results are more robust and the estimated error is more accurate, but on other hand, it is crucial that these tests are performed randomly to ensure that a certain result does not depend on the previous test.

Apart these two aspects, the clustering of similar tests is required to isolate well-known disturbances.

To realize the DoE it is necessary to define the problem and the variables, also named *factors* or *parameters*. Also, the problem domain must be defined.

The domain is a region of interest, a range fluctuation range for each factor. In general, the values that the variables may assume in the DoE is limited.

However, about the domain, it is necessary to distinguish between qualitative discrete quantities and quantitative discrete quantities.

Concerning the first, the characteristics of the solution space are unknown, meaning that the optimal design may undergo to enlarged space. Once understood the error, one can fix the problem accordingly to the project specification.

The DoE techniques, and the number of levels, must be chosen depending on the experiments number.

The *levels* represent the number of different values that a variable assumes accordingly to the adopted discretization scheme: usually the number of level is the same for all the variables, even though some DoE techniques allow to have different levels per variable.

There are several design of experiments techniques. The most remarkable are:

- randomized complete block design - RCBD
- Latin Square,
- full factorial,
- central composite,
- Latin Hypercube.

The technique which has been used in the DoE for the development of the FIA MVM is the Latin Hypercube.

6.1.1 Latin Hypercube

In the Latin Hypercube technique, the domain is divided into an orthogonal grill, in which the number of elements is equal to that of the parameters N . From this grill it possible to recognize N sub-volumes.

Figure 6.2 ^[19] represents the classical graphical layout of Latin Hypercube technique. From the grill, one chooses the sub-volumes (in black) and inside the sub-volume, one chooses randomly the sample.

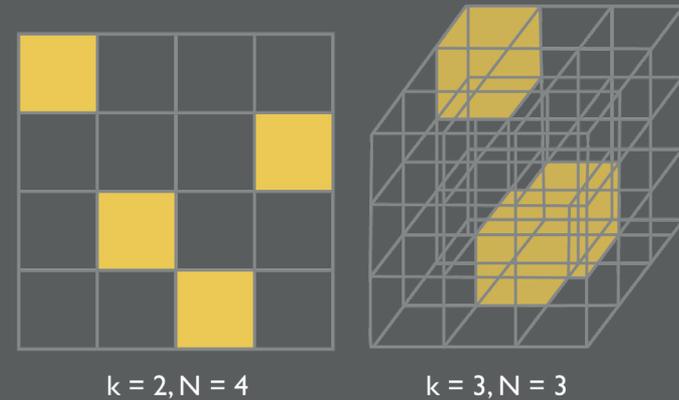


Figure 6.2 - Example of Latin Hypercube technique

It is important that the selection of sub-volumes do not have spurious correlations between the dimensions in order to space all the domain. For example, consider to pick-up the sub-volumes all along the grill diagonal, this is for sure correct from a conceptual perspective, however this strategy shows a strong correlation between the disposition of the sub-volumes, meaning that most of the domain would remain unexplored. Suppose now to have k parameters and N samples: to design th experiments according to the Latin Hypercube method, it is required to build up the following matrices:

$$Q_{N \times k} \quad \text{and} \quad R_{N \times k}$$

in which the columns of Q are given by the random permutations of the integer numbers from 1 to N , while the elements of R are randomly distributed over $[0, 1]$. By setting the range $[0, 1]$ per each parameter, the sampling map S is given by:

$$S = \frac{1}{N} (Q - R)$$

If the elements are distributed over \mathbb{R}^k in accordance with a certain distribution function, each element of S is mapped by the matrix X accordingly with the cumulative distribution function D :

$$x_{i,j} = D_j^{-1}(s_{i,j})$$

The distribution function D may be chosen among several types. The normal Gaussian distribution stands for:

$$D(x) = \frac{1}{2} \left(1 + \operatorname{erf} \left(\frac{x-\mu}{\sigma\sqrt{2}} \right) \right)$$

in which μ is the average value, and σ is the standard deviation. If the parameters are distributed uniformly over the range $[0, 1]$, then $X = S$.

A further solution is adopted to reduce the correlations that map the elements of Q . These last are first divided by $N+1$ through a Gaussian distribution, and then decomposed through the Choleski decomposition, as reported in Figure 6.3 ^[19].

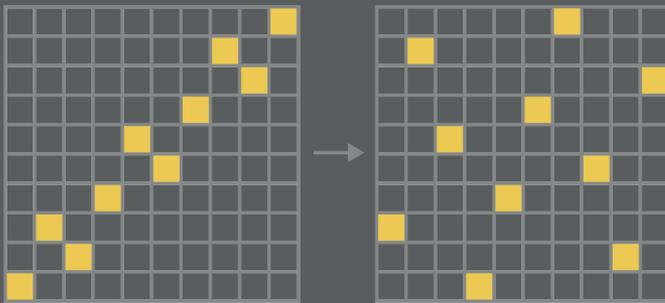


Figure 6.3 - Reduction of correlations in the Latin Hypercube technique with $k = 2, N = 10$

6.1.2 Setting of the DoE

In setting the Design of Experiment in the GT-Power engine model must be modified to comply with the mission, that is, to create optimized experiments according to the Latin Hypercube technique. Eventually, the purpose of DoE is to generate data than be used to define the relationships between the input factors and the mean value map quantities.

For this purpose, the first task to do is to remove any controller (EGR, injection, VGT) because their reference values will be imposed by the input factors.

However, the input variables that have been presented at the beginning of the chapter; refer to the physical engine, not the virtual one. This actually means that all those input variables must be translated into virtual parameters.

Normally, the physical variables such as the intake manifold pressure and the exhaust manifold pressure, are not simulation input but rather a result of the simulation.

To impose different pressures in the DoE, the turbocharger is cut off to impose the compressor outlet and turbine inlet pressure.

This is necessary to make it possible to model pressure ratios across the engine that would occur during a transient, but that cannot be achieved at steady state with the turbo in place (because it would either accelerate or decelerate away from that operating point). The pressure in each ambient will be the parameter that is varied in the DOE. Since there is no turbocharger in the model, again, there is also no need for the VGT control system.

The table below summarizes the process to convert the physical quantities of the engine into input factor of the GT-Power engine model. In addition, the table shows to which GT-Power template the input factors refer to.

More in detail, the EGR rate is controlled through the EGR valve diameter. For this reason, the control system to actuate the diameter is no longer needed.

Concerning the injection, it will be necessary to change the fuel injection control system so that the full range of injected mass can be achieved regardless of the engine speed and boost pressure.

To do so, it is required to control independently the mass flow rate and the timing.

Considering the fuel injected mass flow rate, the GT-Power parameter [FUELMASS] only refers to the quantity of the main pulse, whereas the pilot and pre injection remain unchanged. Note that all these GT-Power parameters will be swept in a user-defined range, and the combinations, accordingly to the Latin Hypercube method, will represent the experiments.

The timing indeed, has been developed outside the GT-SUITE environment and implemented in GT-Power later on. This because the parameter to be controlled is the Start of Injection of the main pulse (SOI_{main}), and by changing it, also the SOI of both pilot and pre injection must be consistent. To ensure this to happen, it is crucial to make the SOI of pilot and pre events dependent on the SOI of

Engine Physical Variable	GT-Power Parameter	GT-Power Part
Engine Speed	[Engine_Speed]	Engine
Intake Manifold Pressure	[PBOOST]	amb-in
Intake Manifold Temperature	[TBOOST]	amb-in
Exhaust Manifold Pressure	[PBACK]	amb-out
EGR Fraction	[d_EGR]	EGR_VALVE
Fuel Mass Rate	[FUELMASS]	InjMultiProfileConn
SOI of the Main Injection	[SOI_MAIN]	SOI_main

Table 6.1 - GT-Power input factors

the main pulse. Consider the focus on the injection system in Figure 6.4. Starting from the maps of SOI of the pilot, pre and main injection, it was possible to retrace the maps of the Dwell times (see paragraph 1.6.2) between the Pilot-Pre injections and between the Main-Pre injections.

As seen from Figure 6.4, the *SOI_{main}* part represents the input for the injectors, as well as for the *SOI_{Pre}* component. In addition, the second input for the *SOI_{Pre}* component is the Dwell-Time between the main and pre injection.

Actually, the *DT_{Main-Pre}* is the map that considers the engine conditions (BMEP and speed) to deliver the angular delay in degrees before top dead center ($^{\circ}$ BTDC) between the two injections.

Similarly, the pilot injection, that is represented by the component *SOI_{Pil}*, receives as first input the timing of the pre injection, and as second input, the Dwell-Time between the Pre and Pilot injection from the *DT_{Pre-Pil}* map.

In this way, the fuel quantity of main injection is set as a parameter by [FUELMASS] and the fuel injection of pre and pilot injections remains unchanged compared to the detailed engine model.

Furthermore, also the timing of the main pulse is a parameter set by [SOI_{MAIN}] and due to the previously mentioned system, the timing of the pre and pilot injections are dependent on the timing of the main pulse, accordingly to the Dwell-Time maps.

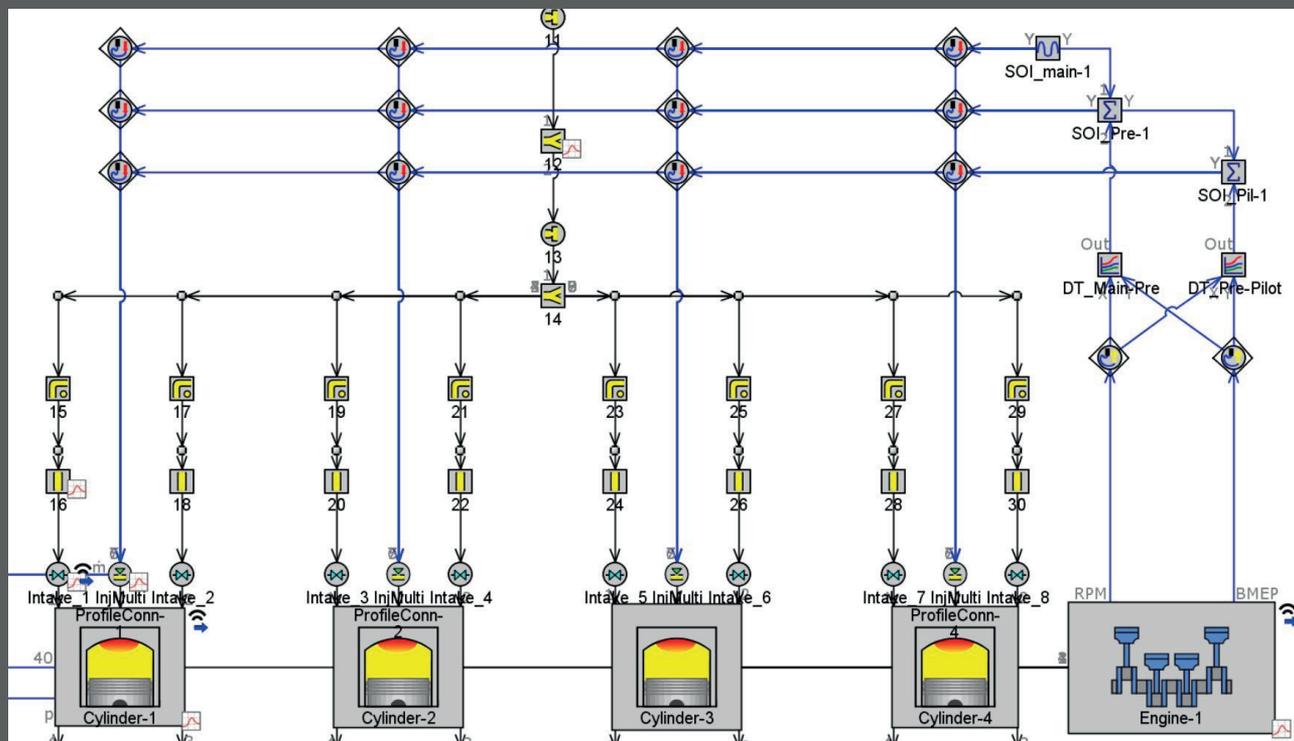


Figure 6.4 - System Injection in the DoE set-up

		Engine speed [RPM]												
		1000	1250	1500	1750	2000	2250	2500	2750	3000	3250	3500	3750	4000
BMEP [bar]	0	-7,43	-8,50	-6,98	-7,57	-7,82	-8,40	-8,65	-9,61	-10,53	-11,42	-12,24	-13,08	-13,93
	2	-8,47	-8,91	-9,26	-8,87	-8,41	-9,00	-8,79	-9,75	-10,92	-12,54	-13,30	-14,29	-15,26
	4	-9,05	-8,89	-8,48	-8,83	-9,22	-9,56	-10,21	-10,94	-11,71	-12,52	-13,37	-14,37	-15,30
	6	-9,12	-8,74	-8,74	-8,84	-8,98	-9,15	-9,94	-10,84	-11,74	-12,60	-13,44	-14,26	-15,27
	8	-9,88	-8,73	-8,71	-8,81	-8,91	-9,05	-9,78	-10,79	-11,72	-12,46	-13,15	-13,85	-8,65
	10	-9,41	-8,78	-8,69	-8,65	-8,77	-8,84	-9,64	-10,85	-11,68	-12,36	-7,83	-8,38	-8,64
	12	-9,44	-8,70	-8,61	-8,46	-8,40	-8,71	-9,67	-10,82	-11,47	-7,33	-7,57	-8,11	-8,65
	14	-9,44	-8,69	-8,32	-8,16	-8,01	-8,69	-9,84	-10,67	-11,12	-7,06	-7,58	-8,10	-8,65
	16	-9,44	-8,69	-8,32	-8,13	-7,98	-8,69	-9,81	-10,51	-6,89	-7,07	-7,54	-8,11	-8,65
	18	-9,44	-8,69	-8,32	-8,13	-7,96	-8,70	-9,81	-10,51	-6,65	-7,05	-7,57	-8,10	-8,65
	20	-9,44	-8,69	-8,32	-8,12	-7,97	-8,70	-9,81	-10,50	-6,67	-7,06	-7,57	-8,10	-8,65
	22	-9,44	-8,69	-8,32	-8,13	-7,98	-8,70	-9,82	-10,51	-6,66	-7,06	-7,57	-8,10	-8,65
	24	-9,44	-8,69	-8,32	-8,13	-7,98	-8,70	-9,82	-10,51	-6,66	-7,06	-7,57	-8,10	-8,65
	26	-9,44	-8,69	-8,32	-8,13	-7,98	-8,70	-9,82	-10,51	-6,66	-7,06	-7,57	-8,10	-8,65
28	-9,44	-8,69	-8,32	-8,13	-7,98	-8,70	-9,82	-10,51	-6,66	-7,06	-7,57	-8,10	-8,65	

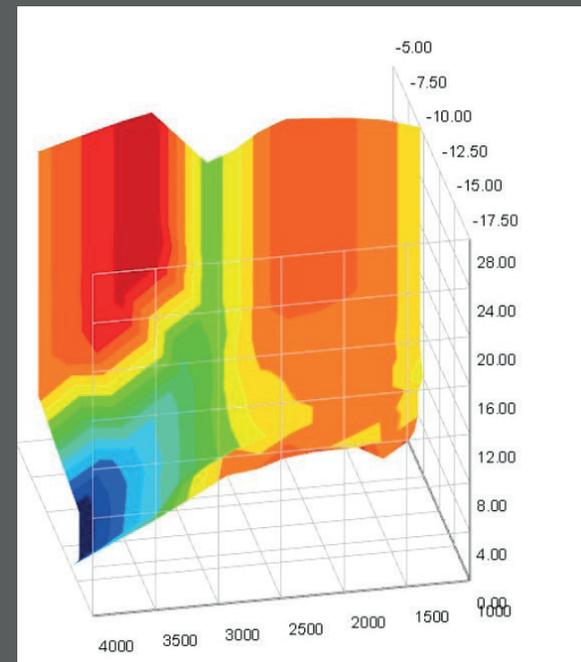
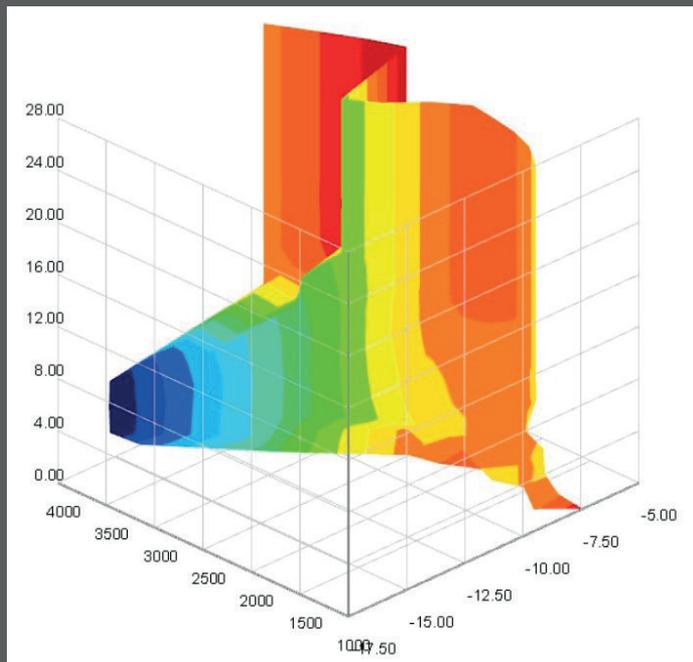
Table 6.2 - Dwell-Time map [$^{\circ}$ BTDC] between Main and Pre injection

Figure 6.5 - Views of the Dwell-Time map between Main and Pre injection

BMEP [bar]	Engine speed [RPM]													
	1000	1250	1500	1750	2000	2250	2500	2750	3000	3250	3500	3750	4000	
0	-7,81	-10,66	-7,73	-9,17	-9,39	-9,94	-10,46	-10,97	-11,04	-13,83	-14,75	-15,17	-10,98	
2	-8,39	-8,96	-9,67	-9,12	-7,67	-8,31	-8,95	-9,51	-9,81	-14,18	-15,10	-16,07	-11,36	
4	-9,00	-9,59	-10,90	-10,49	-10,81	-11,06	-11,73	-12,51	-13,39	-14,37	-15,40	-16,42	-11,77	
6	-9,82	-10,59	-10,86	-11,20	-11,24	-11,38	-11,98	-12,57	-13,22	-14,24	-14,83	-11,03	-11,76	
8	-10,30	-10,65	-10,93	-11,19	-11,25	-11,38	-11,93	-12,43	-13,09	-9,56	-10,29	-11,03	-3,12	
10	-10,31	-10,69	-10,99	-11,19	-11,19	-11,31	-11,81	-12,34	-9,02	-9,56	-3,09	-3,31	-3,12	
12	-10,35	-10,71	-10,98	-11,32	-11,17	-11,19	-9,87	-8,60	-8,82	-2,84	-2,72	-2,93	-3,12	
14	-10,35	-10,71	-10,95	-11,26	-11,07	-8,00	-8,18	-8,42	-8,82	-2,50	-2,72	-2,93	-3,12	
16	-10,35	-10,71	-10,94	-11,24	-11,03	-7,84	-8,10	-8,42	-2,49	-2,50	-2,73	-2,92	-3,12	
18	-10,35	-10,71	-10,93	-11,25	-11,05	-7,83	-8,11	-8,42	-2,17	-2,50	-2,72	-2,93	-3,12	
20	-10,35	-10,71	-10,93	-11,24	-11,03	-7,84	-8,11	-8,42	-2,16	-2,50	-2,73	-2,93	-3,12	
22	-10,35	-10,71	-10,93	-11,24	-11,03	-7,84	-8,10	-8,42	-2,16	-2,50	-2,73	-2,93	-3,12	
24	-10,35	-10,71	-10,93	-11,24	-11,03	-7,84	-8,10	-8,42	-2,16	-2,50	-2,73	-2,93	-3,12	
26	-10,35	-10,71	-10,93	-11,24	-11,03	-7,84	-8,10	-8,42	-2,16	-2,50	-2,73	-2,93	-3,12	
28	-10,35	-10,71	-10,93	-11,24	-11,03	-7,84	-8,10	-8,42	-2,16	-2,50	-2,73	-2,93	-3,12	

Table 6.3 - Dwell-Time map [°BTDC] between Pre and Pilot injection

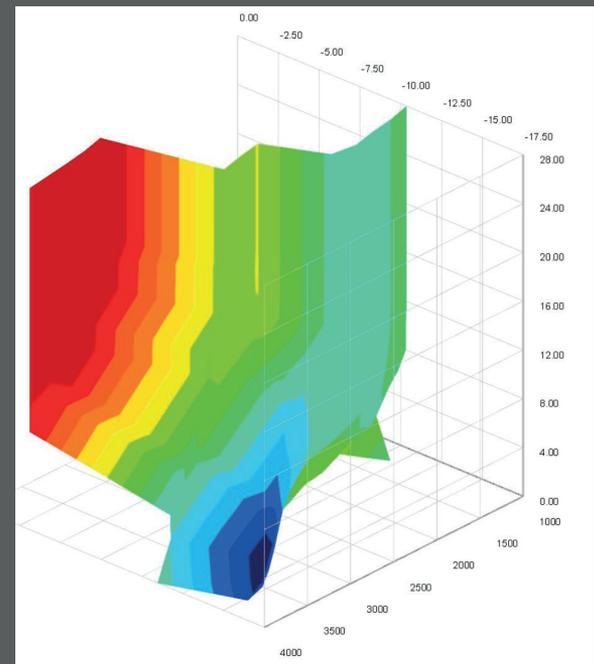
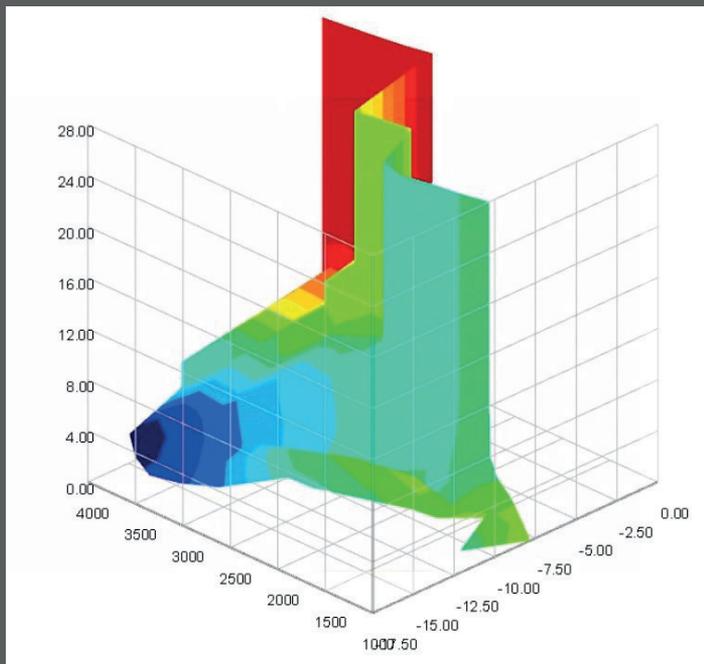


Figure 6.6 - Views of the Dwell-Time map between Pre and Pilot injection

Once all the necessary parameters have been defined and implemented in the GT-Power engine model, then it is required to define the range in which these parameters must sweep. To this purpose, Table 6.4 reports the minimum and maximum values of the parameters.

This DoE includes 2000 experiments. From literature, this value is good trade-off between computational accuracy and time (running time of the simulation about 28 h).

The minimum number of periods for the simulation was set to 25, and the maximum to 50 in order to speed up the calculation (convergence is in general achieved quickly, as the turbocharger is removed from the model).

After the simulation is complete, the stored outputs returns a file of about 400 MB. The complete model of the engine for the DoE is represented in Figure 6.7

Factor	Unit	Min	Max
[Engine_Speed]	RPM	1000	3750
[PBOOST]	bar	1.0	2.5
[TBOOST]	K	303.0	453.0
[PBACK]	bar	1.0	3.5
[d_EGR]	mm	0.0	27.0
[FUELMASS]	mg	0.0	100.0
[SOI_MAIN]	-	-25.0	25.0

Table 6.4 - Parameters sweep settings for DoE

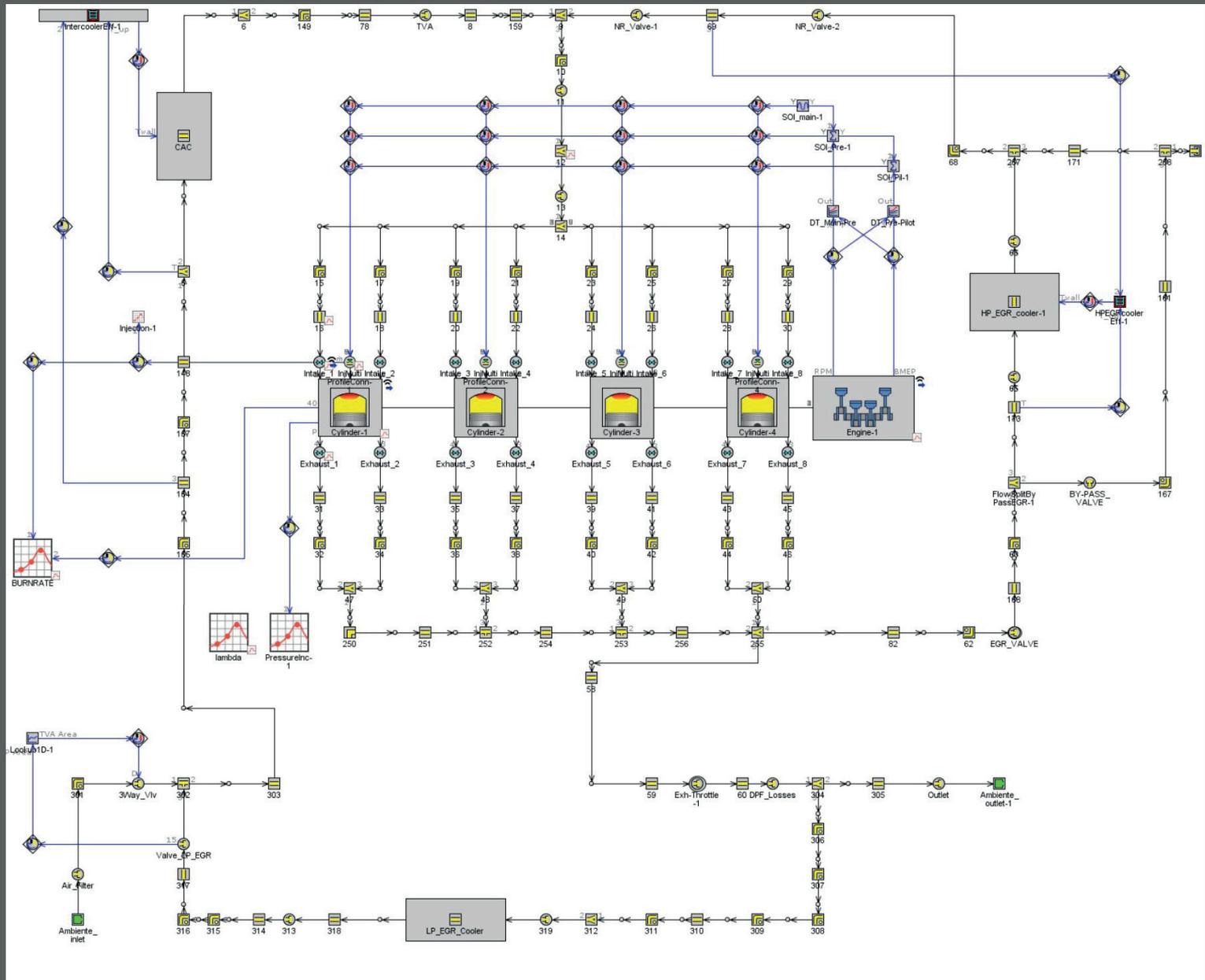


Figure 6.7 - GTPower FIA engine model for the Design of Experiment

6.2 Training of the Neural Network

As in Figure 6.8, the DOE results contain the information of the MV quantities spread in 2000 different engine conditions (experiments), in accordance with the 7 input parameters.

This information is what is needed by these factors (volumetric efficiency, IMEP and exhaust gas temperature) to create their own map which shall be supplied to the MV cylinder. Since these dependencies are complex (7 factors), a lookup table is not suitable, consequently, it is necessary to use neural networks.

An artificial Neural Network (NN) is an information processing system inspired by the way the human brain works, which implies a parallel computing architecture.

Its basic purpose is therefore similar to that of a simple lookup table or map, however the method by which a neural network calculates its output is quite different.

Indeed, the Neural Network is a control component that calculates an output based on multiple inputs using one of several available neural network algorithms. A neural network must be defined to control each of the mapped functions of the mean value cylinder: volumetric efficiency, IMEP, and exhaust temperature. Each neural network must be trained using the data generated in the DOE. A Neural Network has several potential advantages over a simple lookup table. First, it can handle more inputs (up to fifteen). Second, Neural Networks can typically determine outputs faster than a lookup table because there are no interpolation routines.

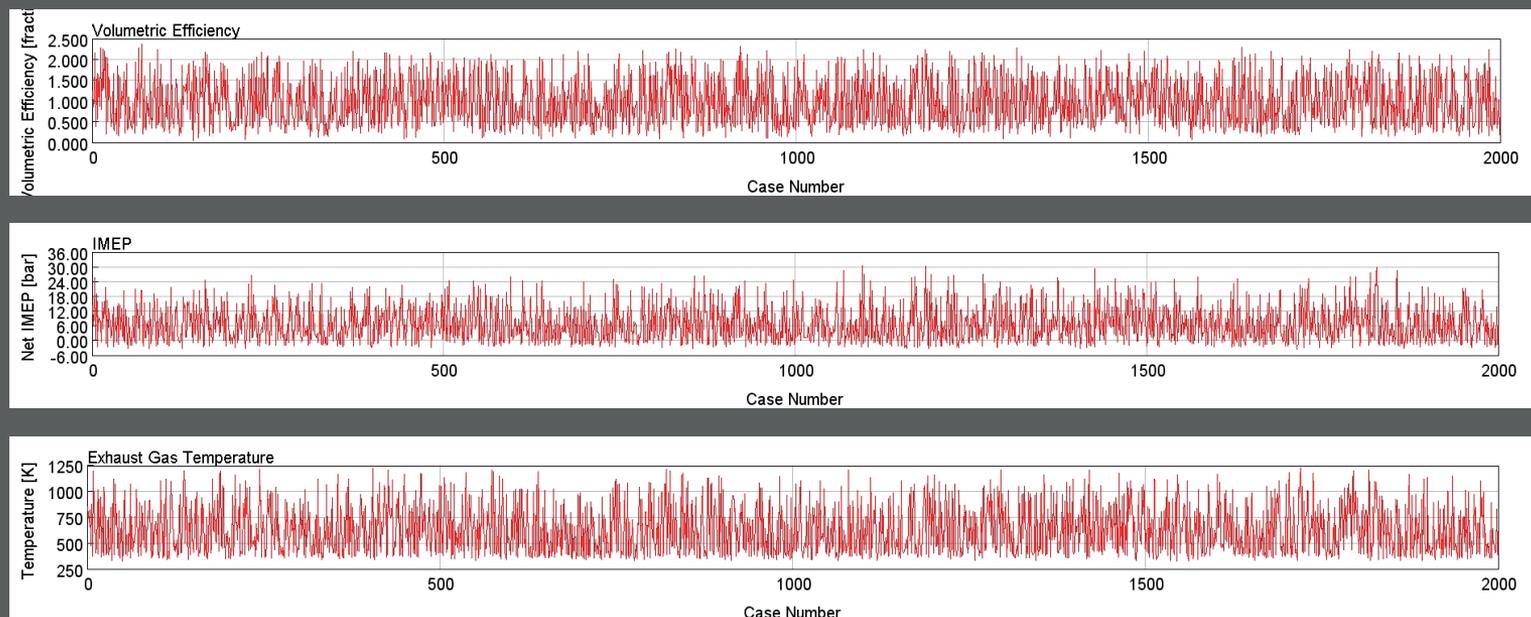


Figure 6.8 - Volumetric Efficiency, IMEP and Exhaust Temperature information after 2000 experiments of the DoE

Eventually, neural networks can potentially do a better job at fitting data where linear interpolation may not be sufficient (therefore requiring less points to adequately define the relationship between inputs and outputs). In general, they can also be referred to as *black box* or *gray box* approaches, as they do not require any detailed physical knowledge of the investigated process and are able to capture complex nonlinear system behavior by means of relatively simple mathematical operations. However, in order to be trained, they usually require a high number of experimental tests, and they are not reliable outside the pre-fixed calibration range.

In GT-Power, the tool *Neural Network Wizard* allows to train the neural networks of interest. Provided that the input of the Neural Network are still the 7 factors, the

DoE Parameter	RLT Variable	Reference Part
[Engine_Speed]	avgrpm	Engine
[PBOOST]	pavl	10
[TBOOST]	tavl0	10
[PBACK]	pavl	58
[d_EGR]	egrin	Engine
[FUELMASS]	injrateav	InMultiProfileConn
[SOI_MAIN]	output l	SOI_main

Table 6.5 - Input RLT variables for NN training

NN admits the result of the DoE, not its same inputs. For this purpose, the Table 6.5 shows the input ResuLT (RLT) variables of the NN, obtained from the DoE simulation.

Some data were excluded for the Neural Network training. In particular, data with EGR rate < 0 or MFB50 cyl1 $> 50^\circ$ ATDC and exhaust mass flow rates < 0 were excluded as unrealistic. When using the *InjMultiProfileConn* object for the injectors, the parameter *Angle at Start of Injection* in the outputs (i.e., parameter *thinj:injector l*) indicates the SOI of the most advanced pulse (pilot in this case), which is not usually the one that is needed for training the neural networks. Indeed, the reference quantity is the start of the injection of the main pulse, that is, SOI_{main} .

It is therefore necessary to include, among the output data to be stored in the DoE, the object which imposes the SOI_{main} in the model (typically a signal generator).

The results of a NN training consist of a file *.nno* that is actually the map of the quantity of interest. The consistency of such map is validated by the regression plots.

These plot show the results of the best network by plotting the predicted values compared to the experimental. The more the predicted values lie on the angular bisector, the higher the accuracy of the model.

Note that because the Friction Mean Effective Pressure (FMEP) will be implemented in the Mean Value cylinder, then an additional NN has been developed. The resultant regression plots are represented from Figure 6.9 up to Figure 6.12

Volumetric Efficiency

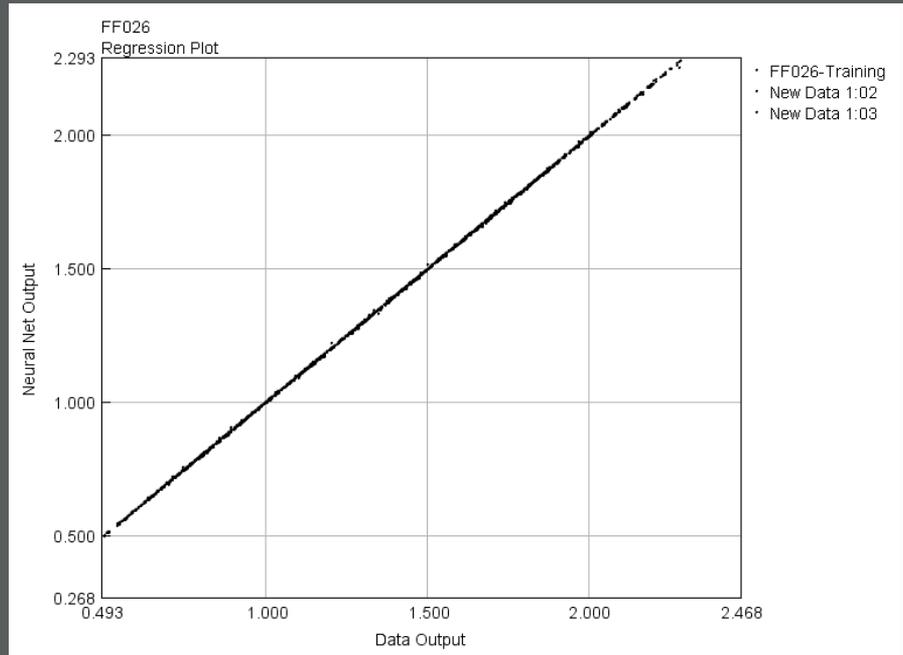


Figure 6.9 - Regression plot of the Volumetric Efficiency

IMEP

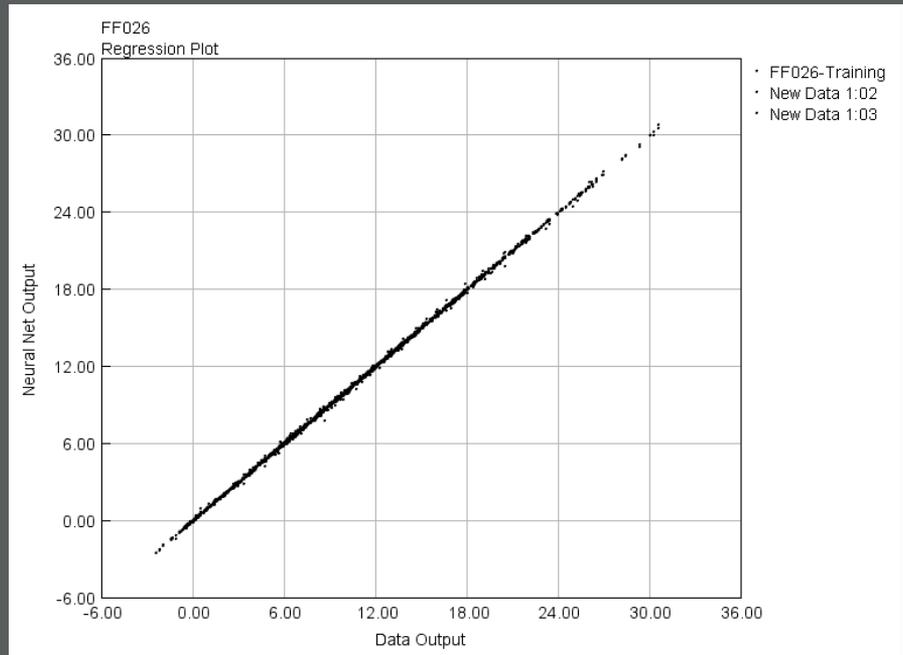


Figure 6.10 - Regression plot of the Indicated Mean Effective Pressure (IMEP)

Exhaust Gas Temperature

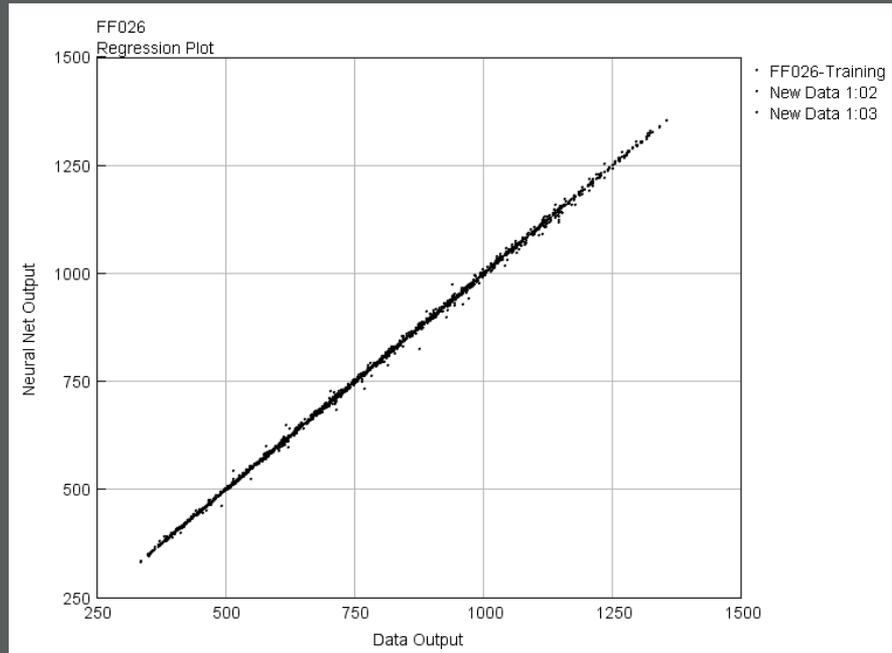


Figure 6.11 - Regression plot of the Exhaust Gas Temperature

FMPE

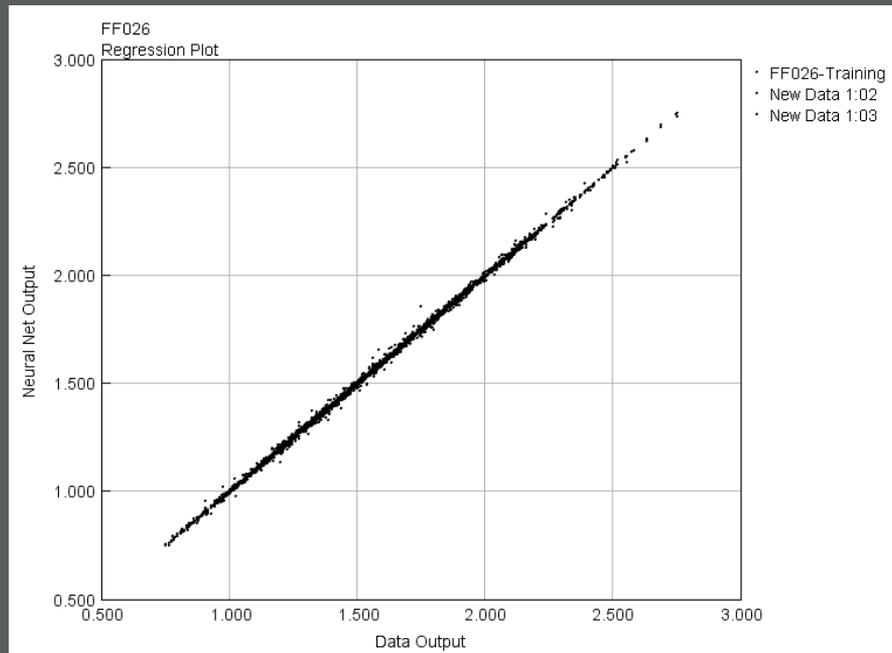


Figure 6.12 - Regression plot of the Friction Mean Effective Pressure (FMPE)

6.3 From FR to MV Model

Once the Mean Value quantities have been trained through the Neural Network, it is now necessary to make comply the overall engine model with the Mean Value Requirements.

The first step is to simplify the flow system to minimize the number of volumes in the model. The main component that, starting from the detailed model, are combined into lumped volumes are:

- InMan_MV
all the components and pipes of the intake manifold, runner and ports from the outlet of the compressor to the cylinder;
- ExMan_MV
all the components and pipes of the exhaust manifold, runner and ports from the outlet of the cylinder to the inlet of the turbine,
- Heat Exchangers
In the detailed model, a semi-predictive model is used for each of the heat exchangers (intercooler and EGR cooler). For each heat exchanger, a block of controls components calculates the desired outlet temperature based on the inlet mass flow rate, temperature and efficiency.

In Fast-Running and Mean Value model, the template changes into *HeatExchangerConn* since it is specifically intended to run at real-time, when it is necessary to impose the gas temperature at a particular point in a flow system

Actually, most of the flow system simplification has been already obtained in the Fast-Running model development. Indeed, a FR model drops down the computation time of the simulations by reducing the sub-volumes of the correspondent Detailed model.

What really converts a Fast-Running into a Mean Value Model is the implementation of a MV cylinder, a MV injector and the Neural Network.

The MV cylinder represents an engine cylinder through mapped functions of volumetric efficiency (i.e. airflow), indicated mean effective pressure, and exhaust gas temperature (i.e. exhaust energy). This component substitutes all the cylinder of the engine. As usual, it is mechanically connected to the engine and to the manifolds, whereas the information are exchanged with the Neural Network only.

The MV injector only works with the MV cylinder and considers the mass of fuel injected per cycle. Even though the engine is a multicylinder engine, the MV injector only requires the specifications of a single cylinder.

Finally, the previously trained neural networks is now implemented through *controls components* that sense the neural network input quantities, calculate the neural network outputs, and then actuate the output quantities. The inputs that the NN require are:

- static pressure, temperature and burned gas mass fraction variables from the intake manifold,
- engine speed from the engine,
- mass flow rate from the injector,
- static pressure from the exhaust manifold.

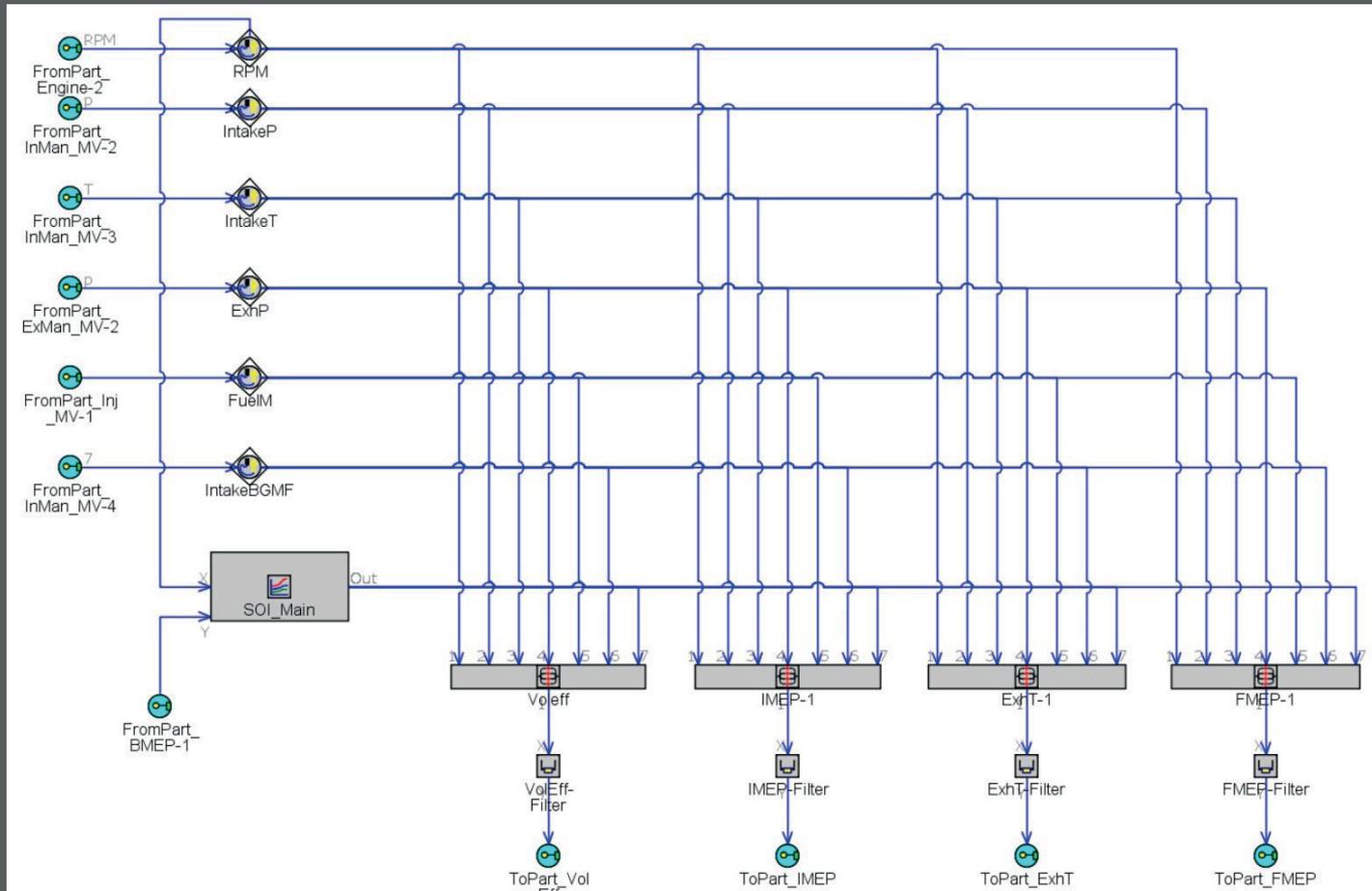


Figure 6.14 - Detail of the Neural Network layout

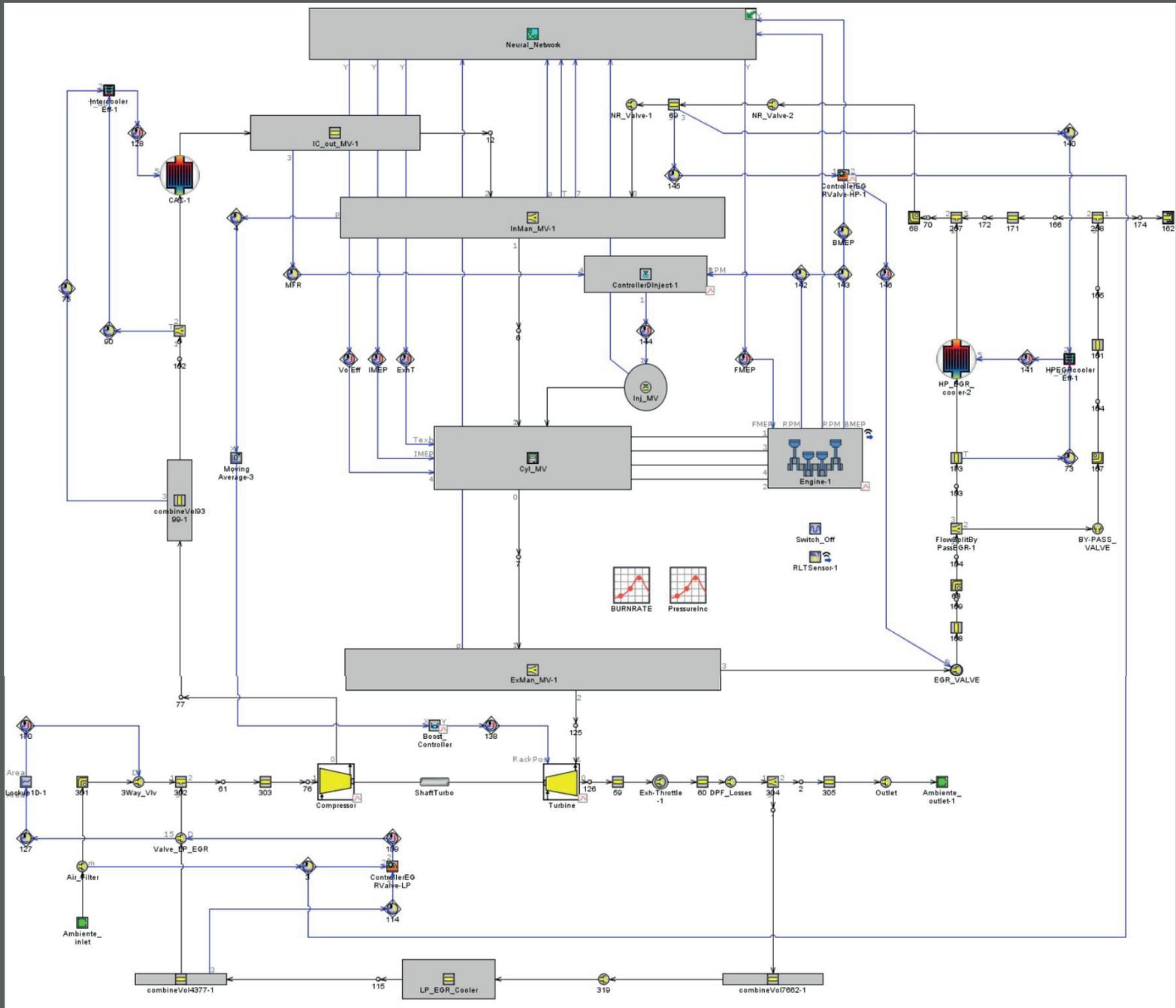


Figure 6.15 - Mean Value Model of the FPT FIA engine

6.4 Comparison between DETM, FRM and MVM

The comparison between the Detailed, the Fast-Running and the Mean Value model refers to engine map discussed in the paragraph 4.3 and described through the Table 4.2 and Figure 4.20. Provided the fidelity of the model, the most remarkable result is that the Factor of Real Time (FRT) has decreased on average up to 2.8. Again, the models have been tested to the following quantities:

1. Factor of Real Time	Turbine Average Speed .11
2. Brake Torque	Turbine Outlet Temperature .12
3. Air Flow Rate	VGT Controller .13
4. BMEP	Compressor Outlet Pressure .14
5. Volumetric Efficiency	Intercooler Outlet Pressure .15
6. Injected Average Mass Flow Rate	Intercooler Outlet Temperature .16
7. Intake Manifold Pressure	EGR Rate .17
8. Intake Manifold Temperature	EGR Valve Mass Flow Rate .18
9. Exhaust Manifold Pressure	EGR Controller .19
10. Exhaust Manifold Temperature	EGR Cooler Outlet Temperature .20

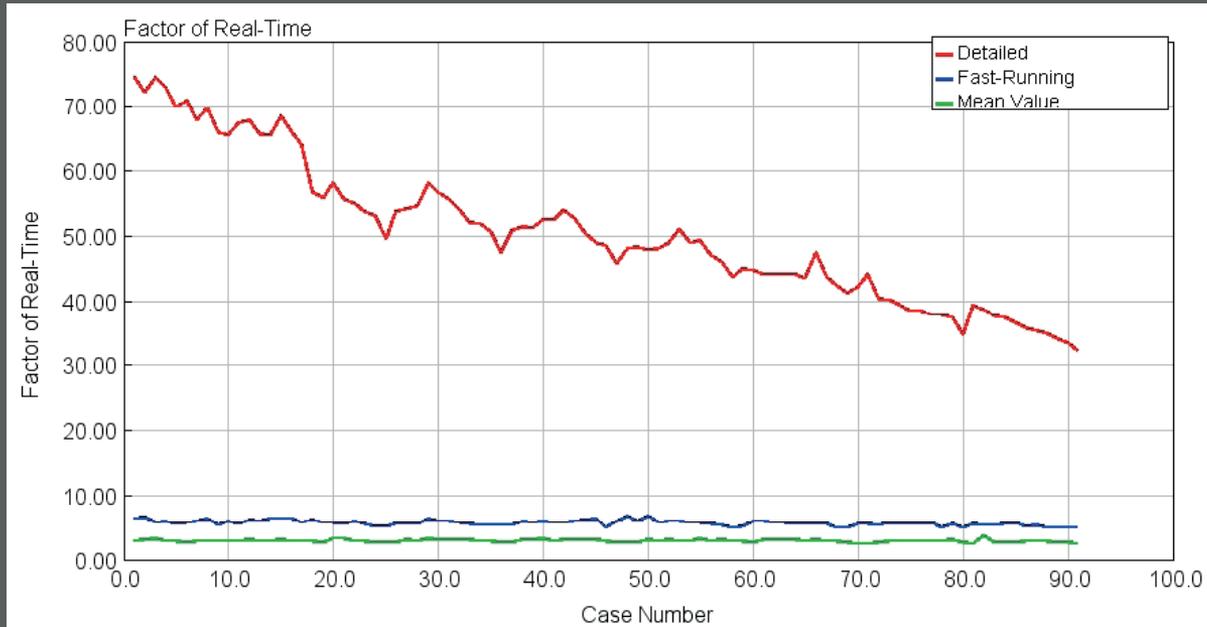


Figure 6.16 - FRT trend, DETM vs FRM vs MVM

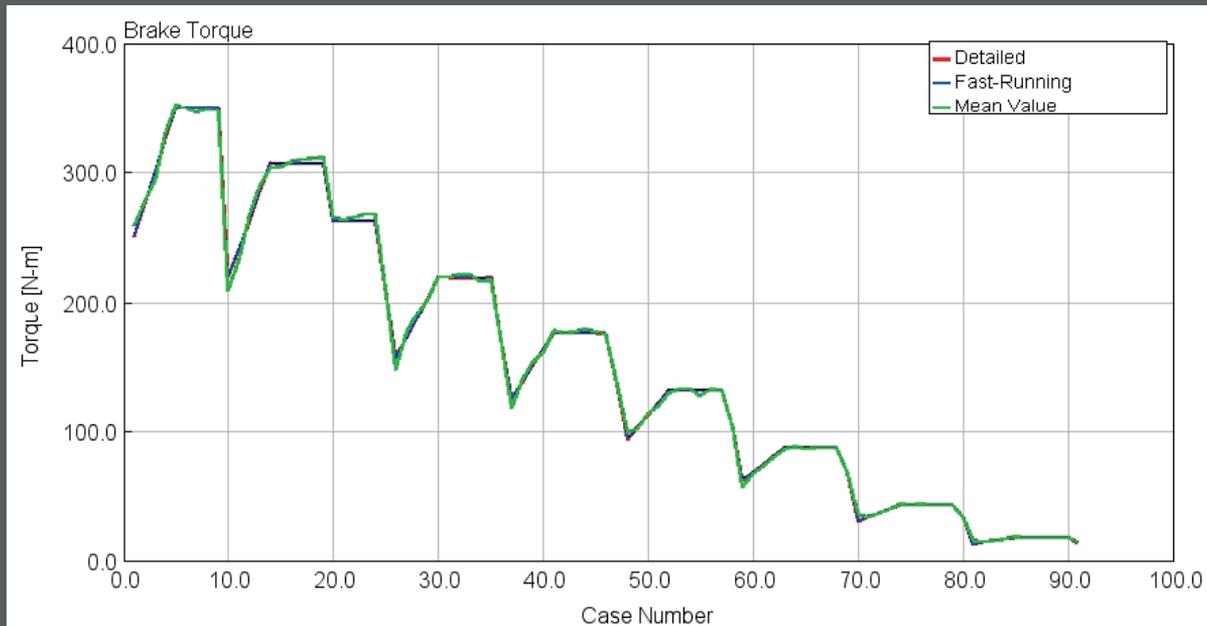


Figure 6.17 - Brake Torque, DETM vs FRM vs MVM

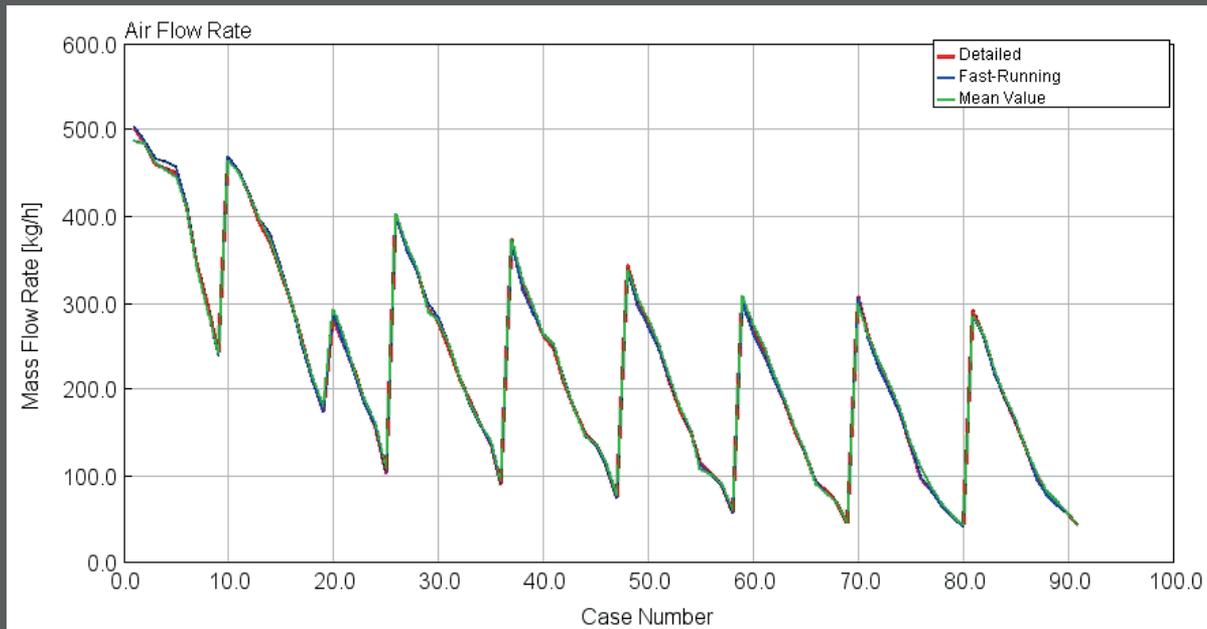


Figure 6.18 - Air Flow Rate, DETM vs FRM vs MVM

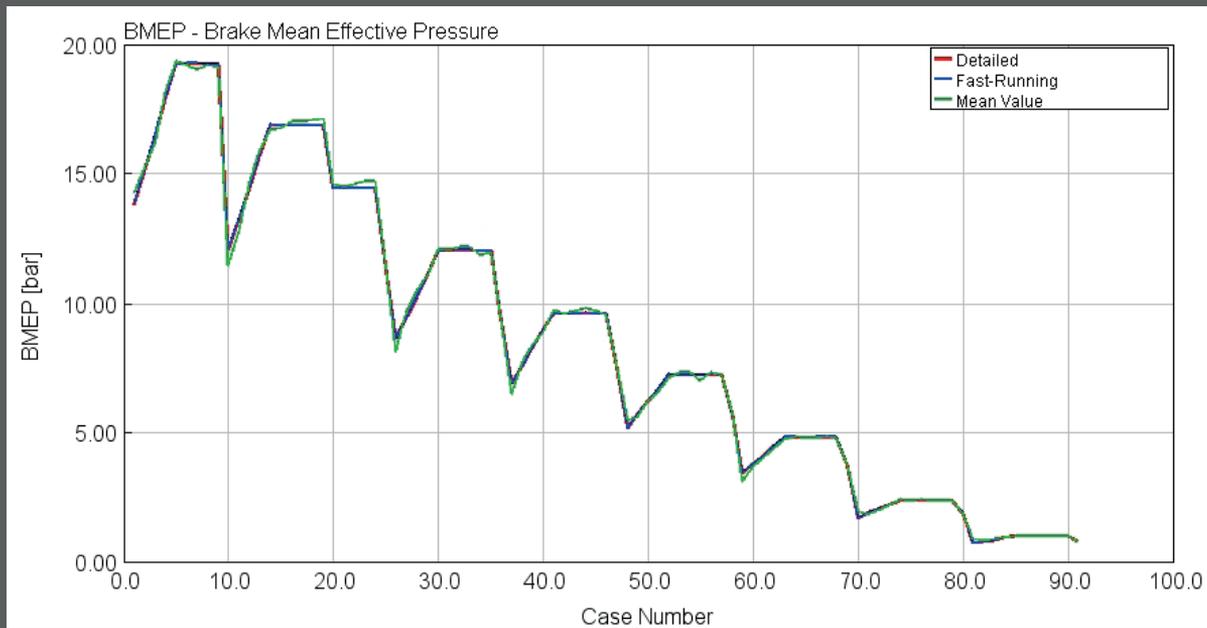


Figure 6.19 - BMEP, DETM vs FRM vs MVM

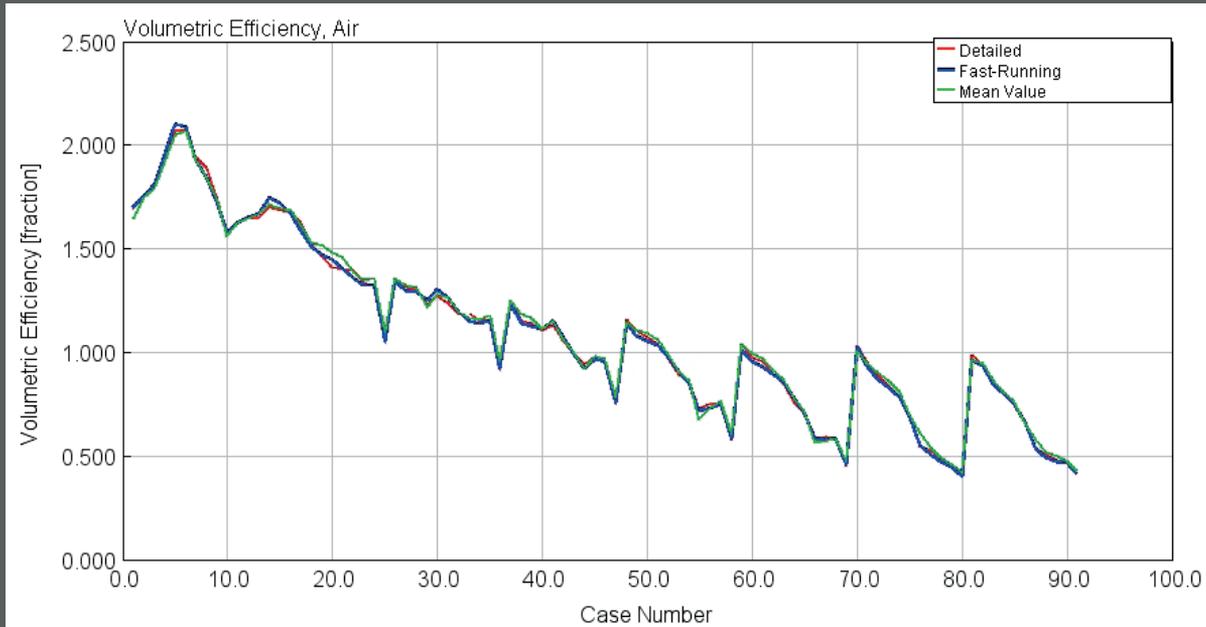


Figure 6.20 - Volumetric Efficiency, DETM vs FRM vs MVM

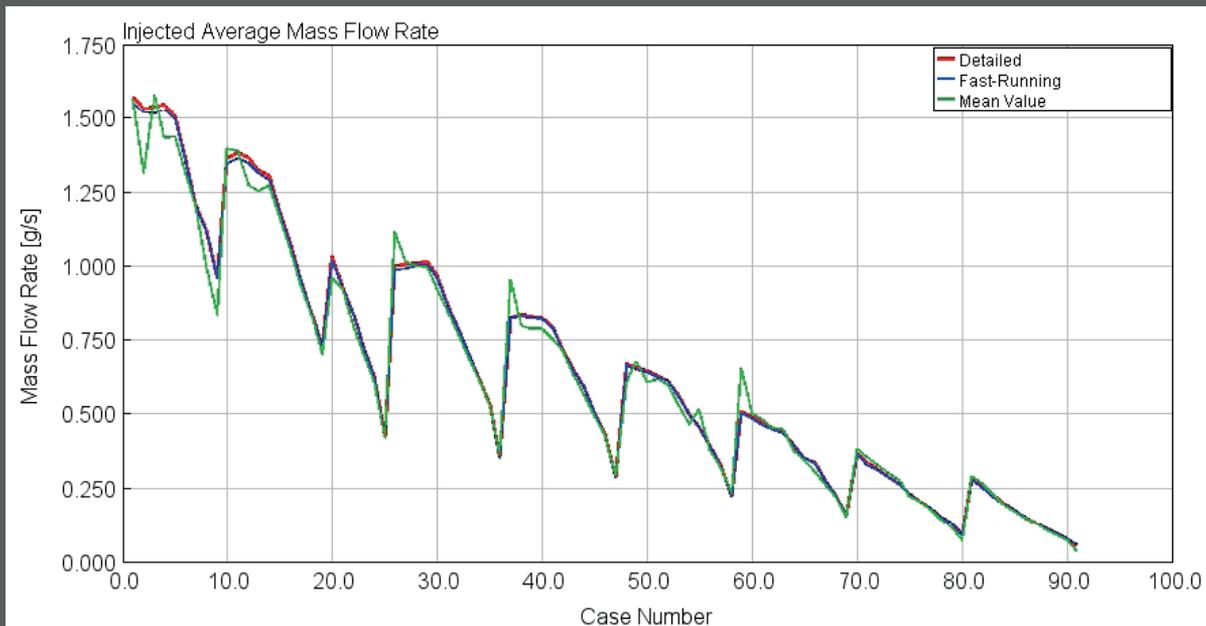


Figure 6.21 - Injected Average Mass Flow Rate, DETM vs FRM vs MVM

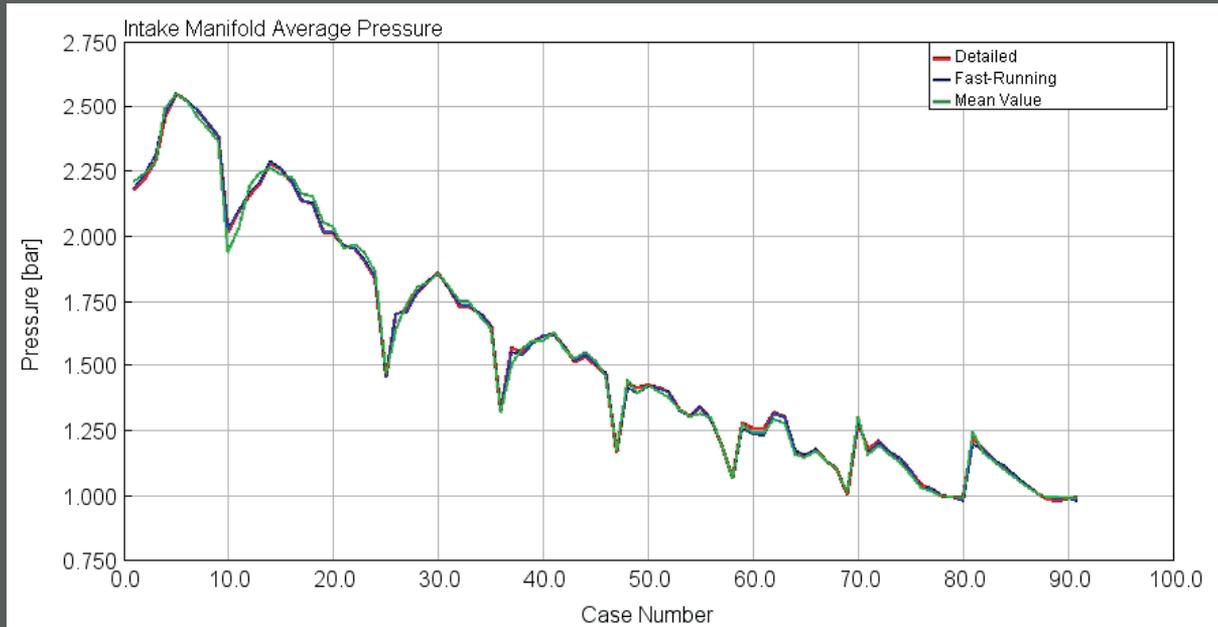


Figure 6.22 - Intake Manifold Pressure, DETM vs FRM vs MVM

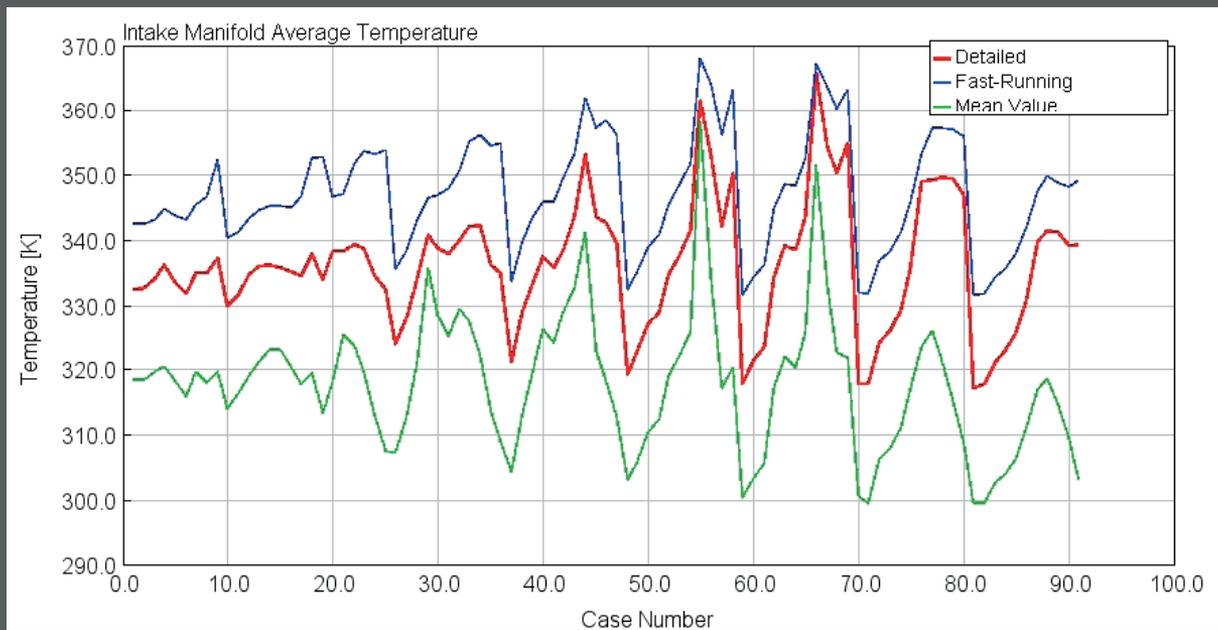


Figure 6.23 - Intake Manifold Temperature, DETM vs FRM vs MVM

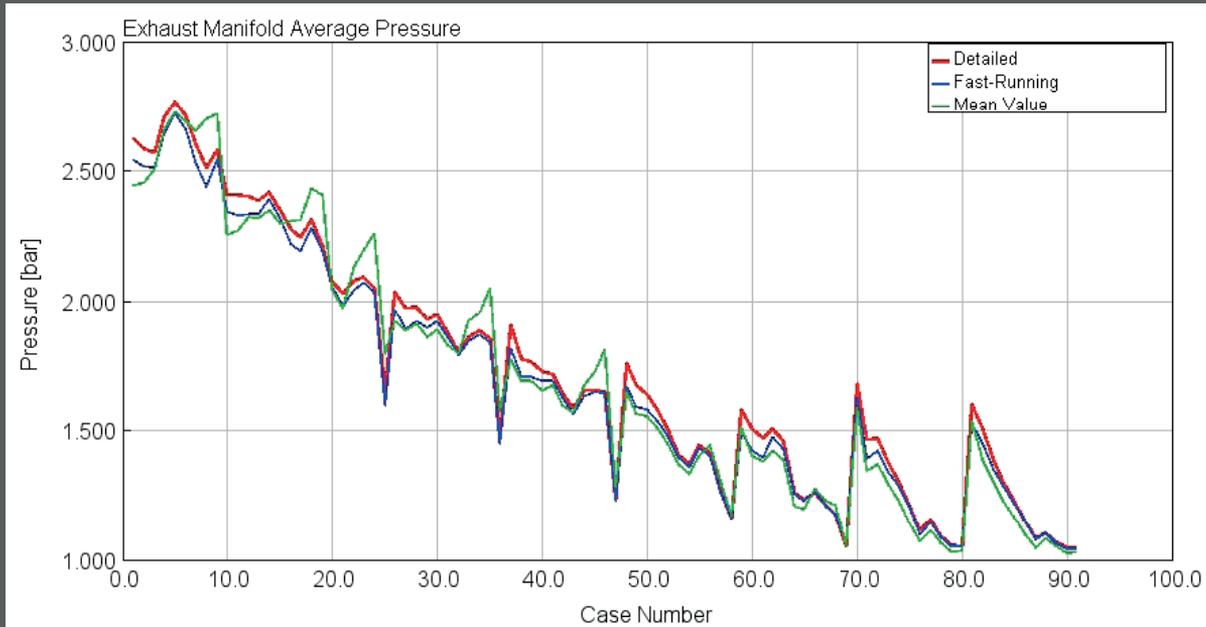


Figure 6.24 - Exhaust Manifold Pressure, DETM vs FRM vs MVM

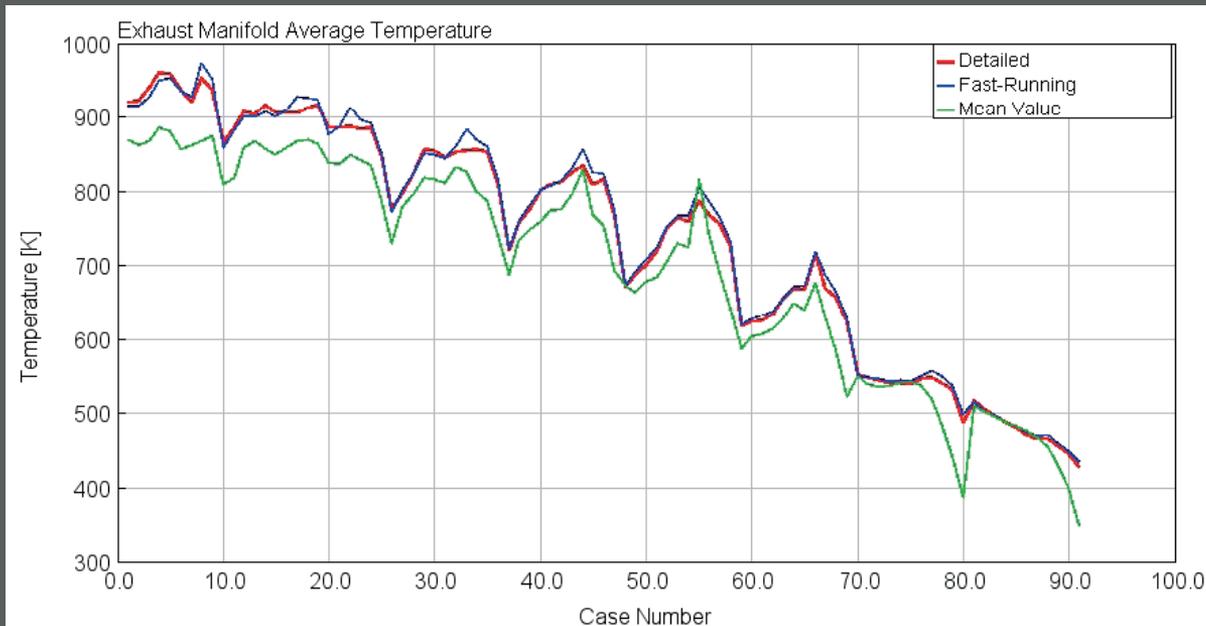


Figure 6.25 - Exhaust Manifold Temperature, DETM vs FRM vs MVM

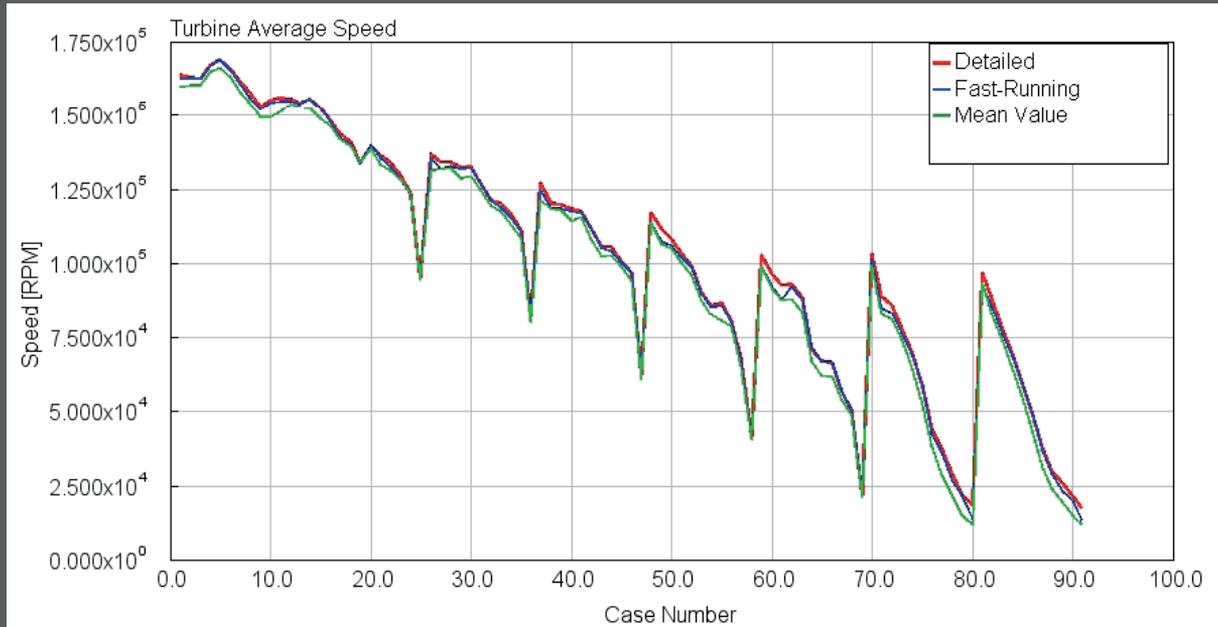


Figure 6.26 - Turbine Average Speed, DETM vs FRM vs MVM

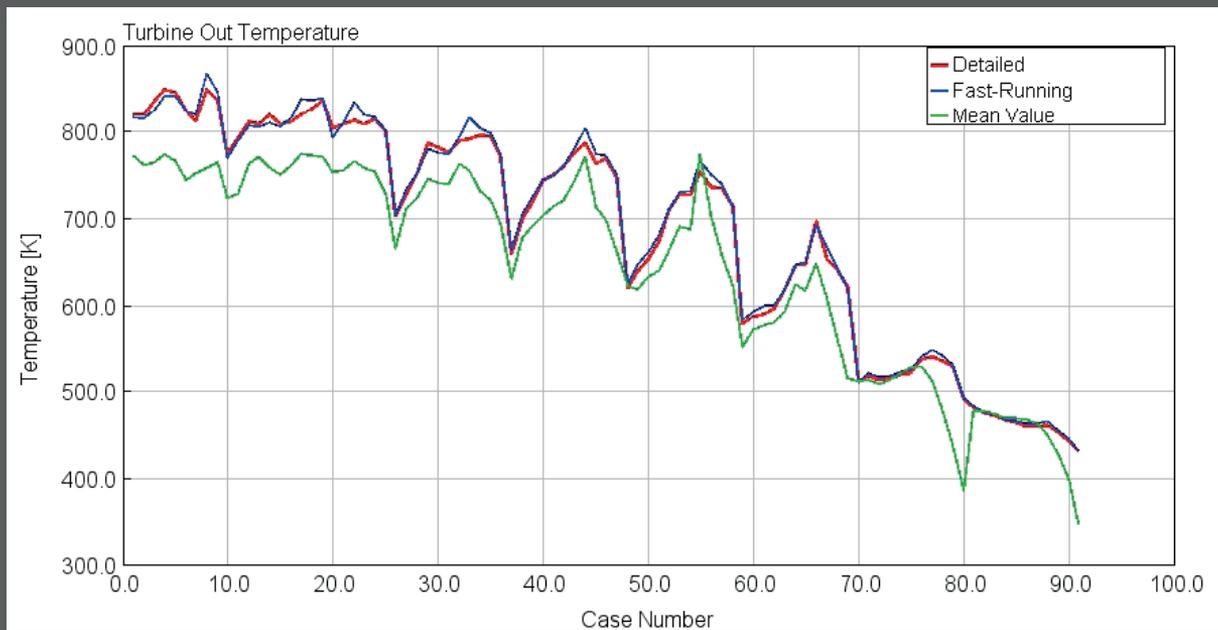


Figure 6.27 - Turbine Outlet Temperature, DETM vs FRM vs MVM

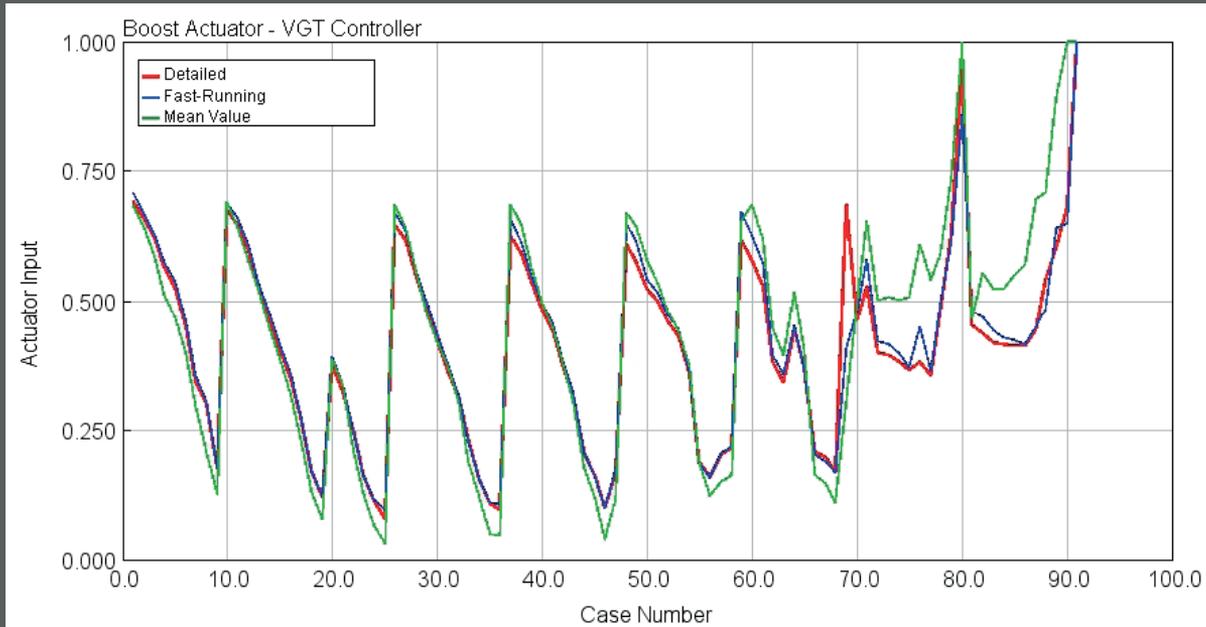


Figure 6.28 - VGT Controller (Boost Actuator), DETM vs FRM vs MVM

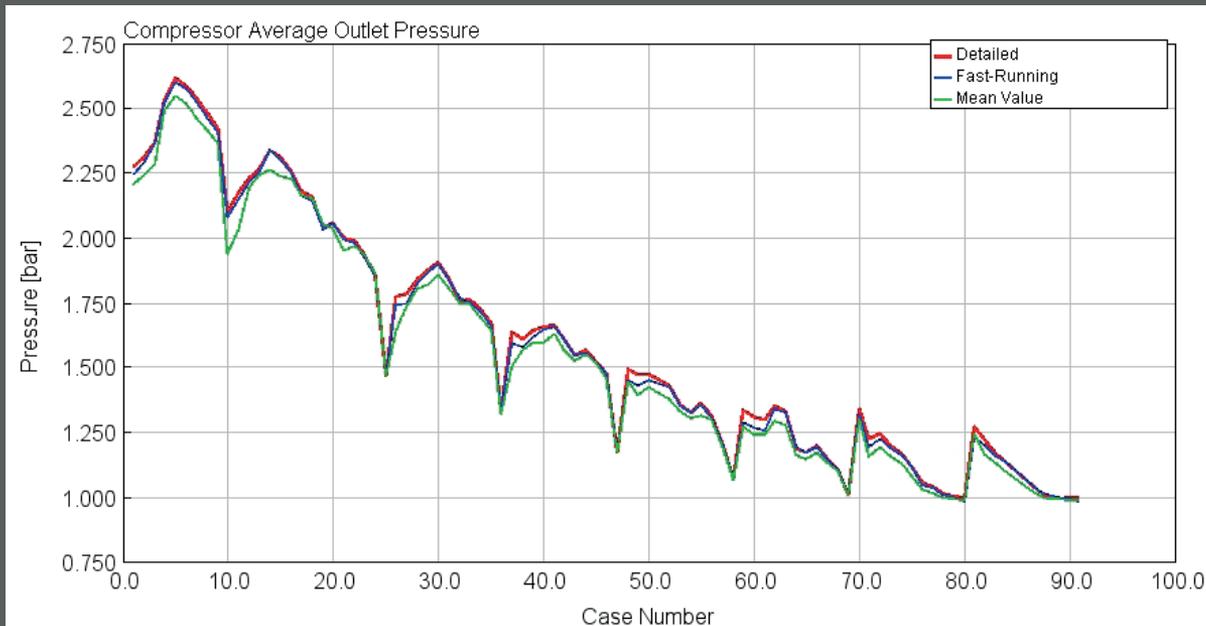


Figure 6.29 - Compressor Outlet Pressure, DETM vs FRM vs MVM

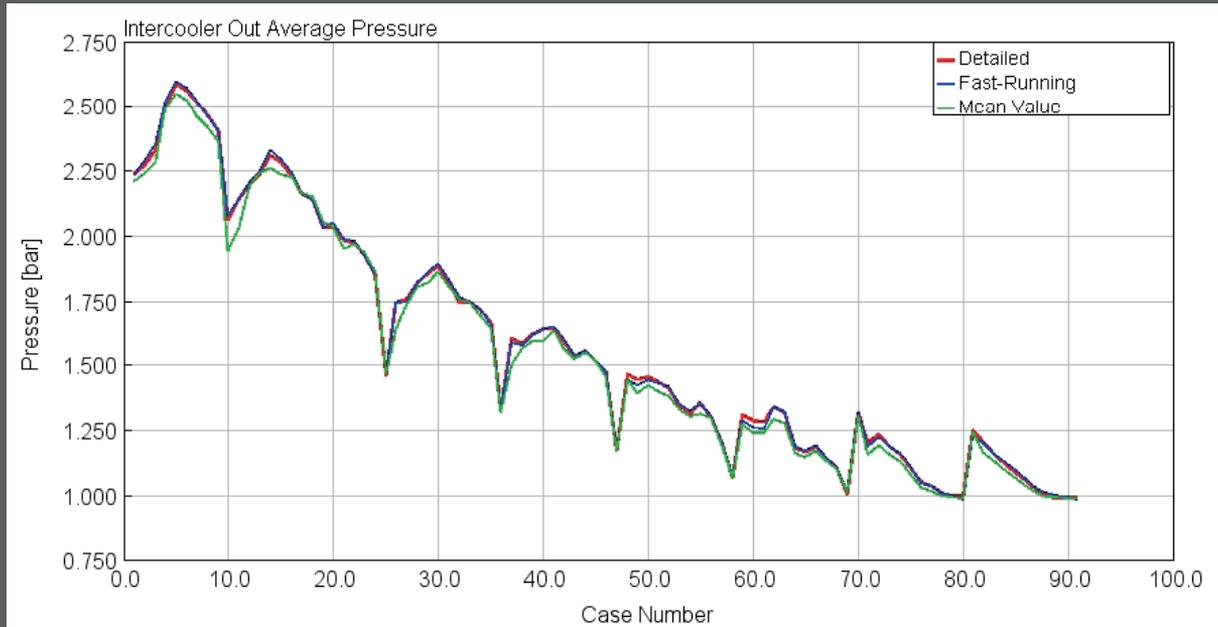


Figure 6.30 - Intercooler Outlet Pressure, DETM vs FRM vs MVM

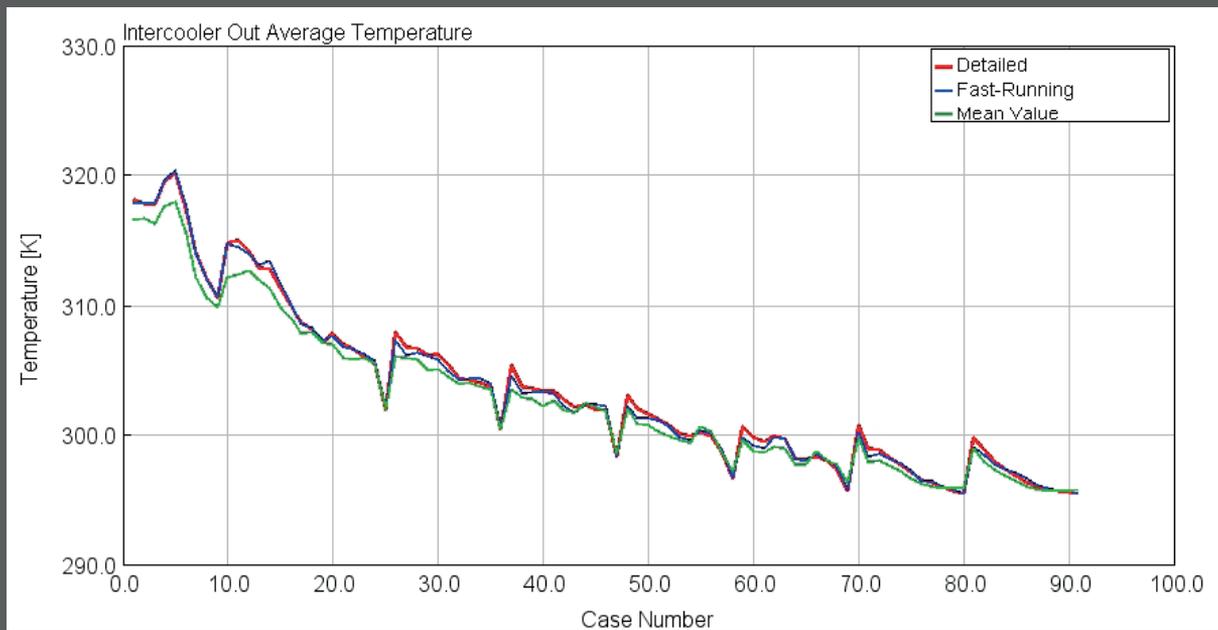


Figure 6.31 - Intercooler Outlet Temperature, DETM vs FRM vs MVM

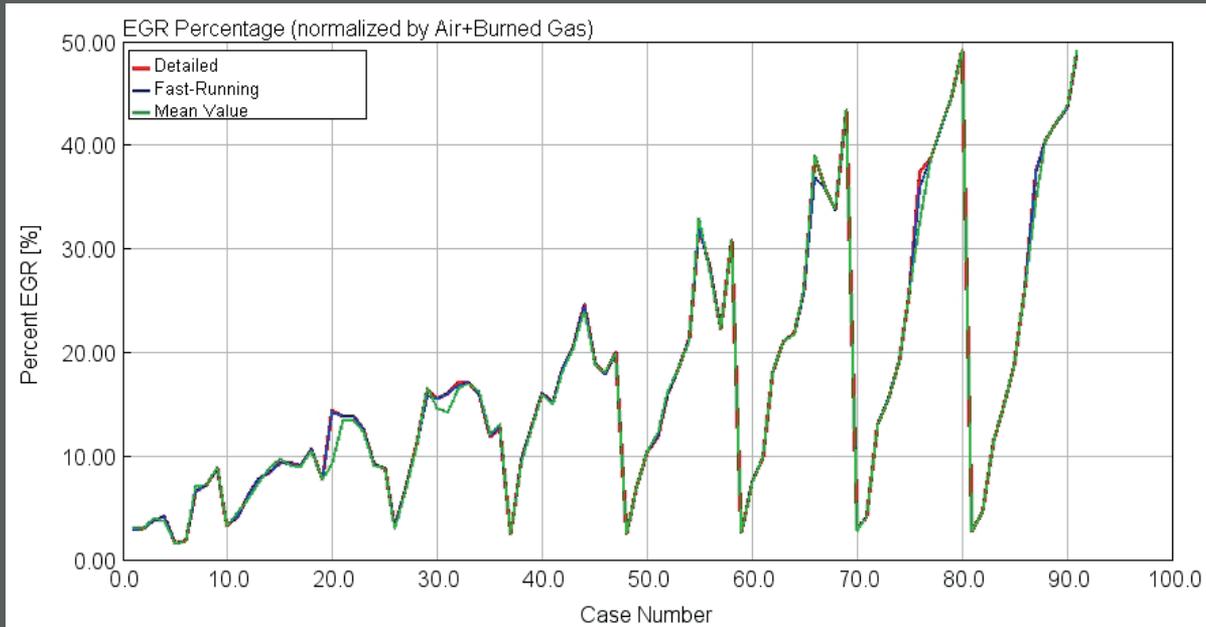


Figure 6.32 - EGR Rate, DETM vs FRM vs MVM

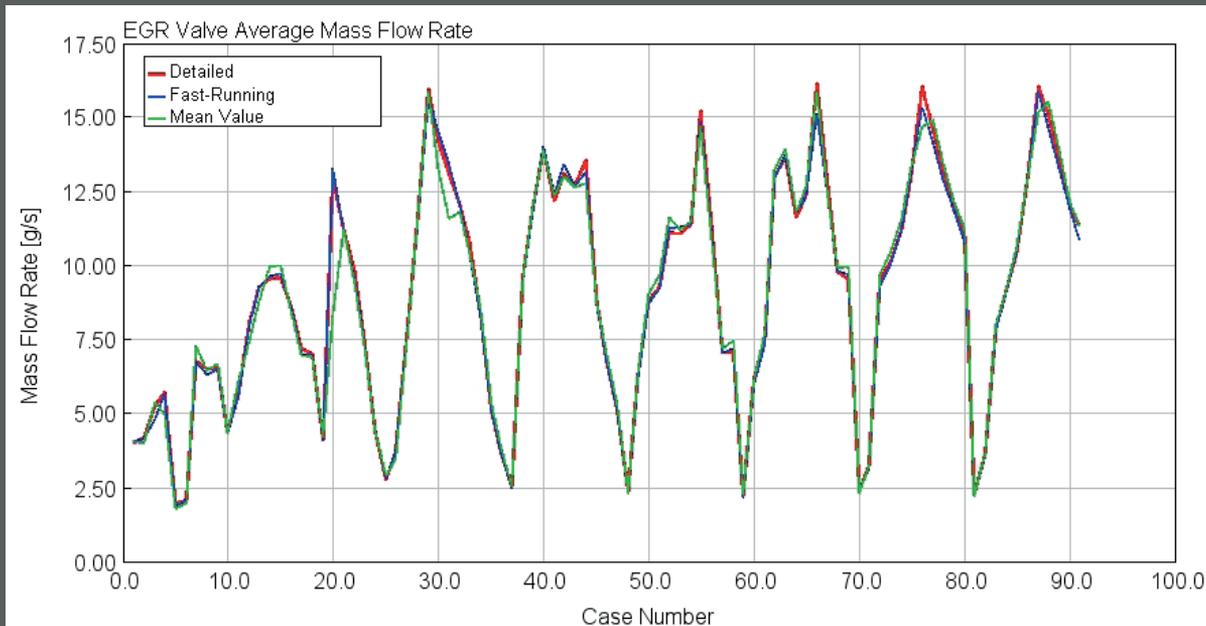


Figure 6.33 - EGR Valve Mass Flow Rate, DETM vs FRM vs MVM

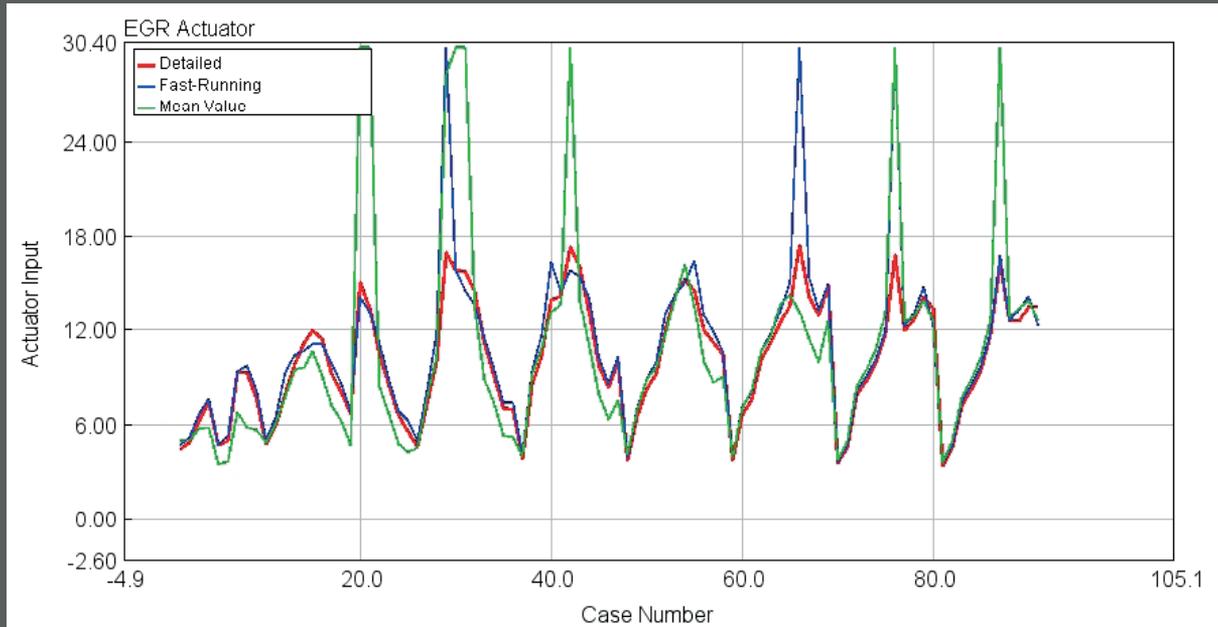


Figure 6.34 - EGR Controller, DETM vs FRM vs MVM

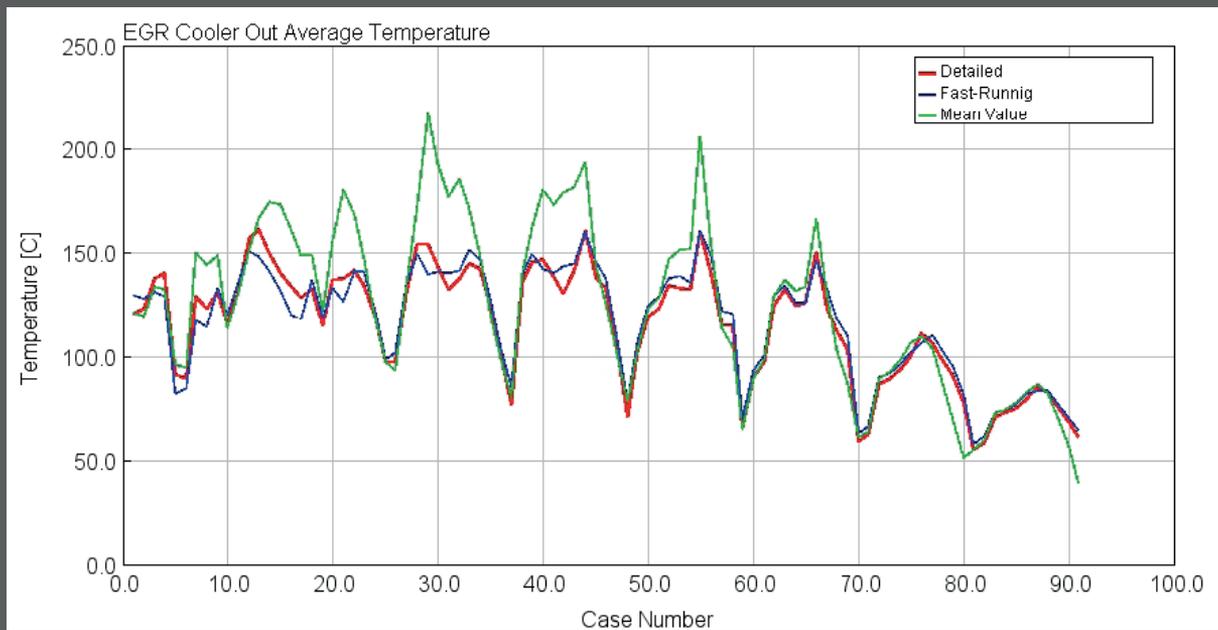


Figure 6.35 - EGR Cooler Outlet Temperature, DETM vs FRM vs MVM

6.5 Integration of MVM in HiL Testing

In the early stages of testing, the design of engine controllers is both developed and assessed in a fully virtual environment. Recalling paragraph 5.3, in the Model in the Loop testing, the engine model is developed in GT-Power and it is the Fast-Running model, whereas the control algorithm is developed in Matlab-Simulink environment. The code is then tested by coupling the Fast-Running engine model and the Simulink controller by means of the block *Simulink harness* that enables the communication between the two software. The simulation is run in GT-Power, while the results are processed in GT-Post.

Passing from Model to Hardware in the Loop, the virtual degree of the test procedure is reduced meaning that the calculations are no more made by the PC hardware, but rather by specific components that simulate the engine and the controller. Still, the GT-Power model and the Simulink code supply the information to the hardware but then the engine emulator and the controller emulator, called Prototyping Interface, are connected each other to communicate.

The Mean Value Models are required to perform Hardware in the Loop testing procedures owing to the lower computational time compared to Fast-Running Models.

So, in HiL testing, the hardware and software to perform the assessment of the algorithm control are:

- Engine Emulator;
- Real Time OS and Engine Emulator Software,
- Prototyping and Interface Module.

The Engine Emulator - NI PXI

An engine Emulator is a device devoted to the simulation in Real-time of the tested plant. In Politecnico di Torino, the one used for tests is the NI-PXIe 1082 chassis, in which it is inserted the controller e-8135. PXI stands for PCI eXtensions for Instrumentation that is a rugged PC-based platform that offers a high-performance, low-cost deployment solution for measurement and automation systems. The NI PXI and the GT-Power Mean Value Model represent the engine emulator. However, to integrate the GT-Power and the PXI, it is required to convert the MVM into a compiled model (.dll file) in a dedicated Engine Emulator Software.

PXI Type	NI PXIe-8135
Controller	Intel Core i7-3610QE 2.3 GHz
I/O Interfaces	2-port CAN, High-Speed/FD
Real-Time OS	Pharlap OS
Execution Mechanism	Veristad Engine
Engine Model	GT-Power Mean Value Model

Table 6.6 - PXIe-8135 data sheet

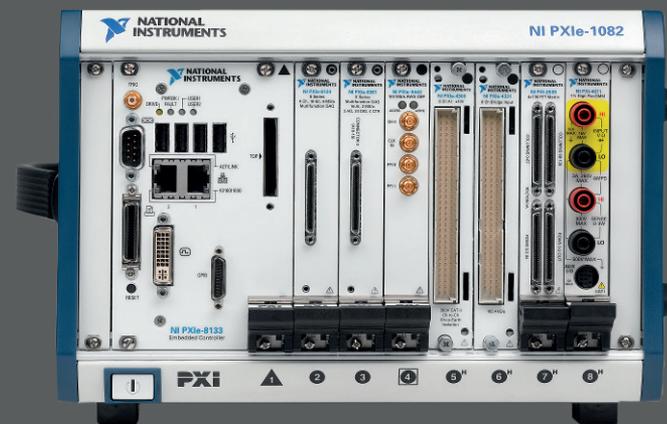


Figure 6.36 - NI-PXIe-1082 from National Instruments

The Engine Emulator Software - NI Veristand

A Real-time test requires a Real-Time Operative system (RTOS). The real-time simulation does not imply a lower duration of the performed test, but rather that the test performs an in-time simulation. Microsoft Windows is not a RTOS, so the NI-PXIe8135 is equipped with *Phar-lap* operative system that guarantees the correct duration of the test under investigation. The real-time test allows a cost reduction for the development, a shortening of time to market and an improvement of quality.

The Engine Emulator is managed by *NI-VeriStand*. This software is the environment for configuring the real-time testing as well as the virtual interface between the engine emulator and the controller emulator. VeriStand works with the execution mechanism represented in Figure 6.37⁽⁵⁾. This mechanism is applied to control the timing of the entire system and also to manage the communication between the execution target and host computer. The VeriStand mechanism is composed of multiple loops that accomplish specific tasks with specific priority. The technique to transfer data between loops is the real-time FIFOs (*First In First Out*), meaning that the first received signal is also the first to be evaluated.

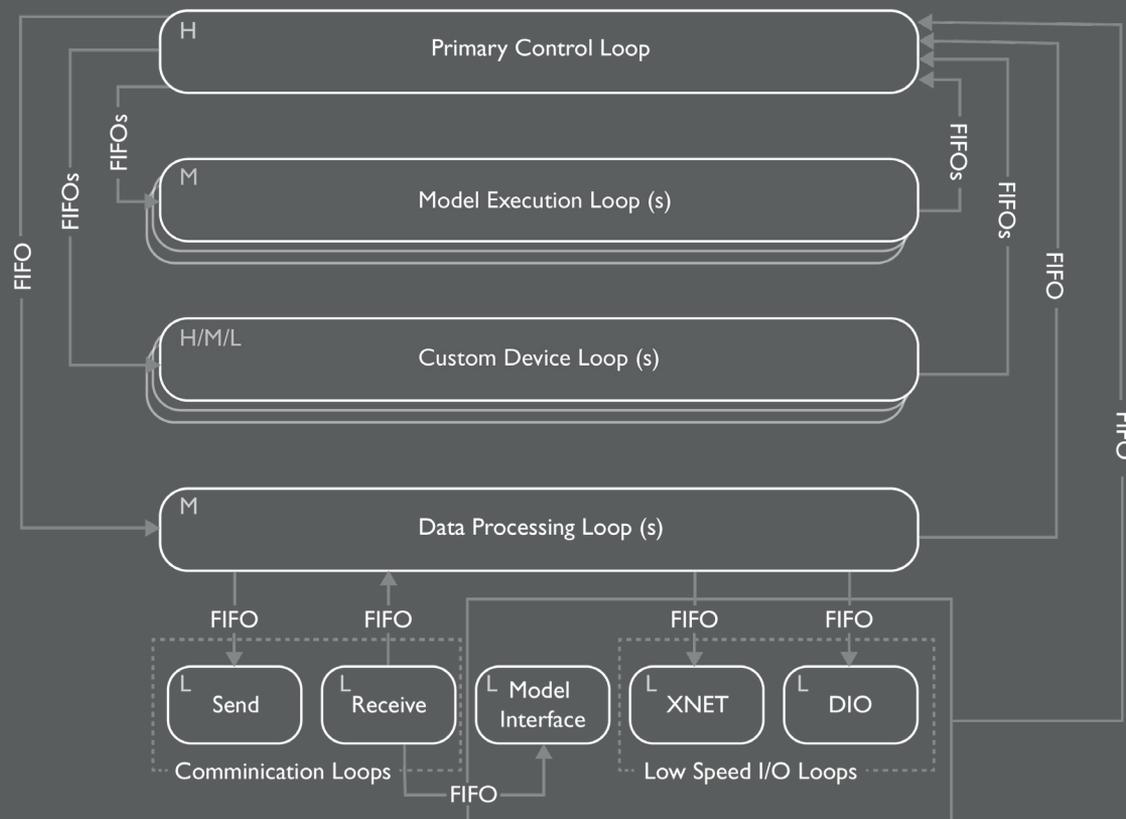


Figure 6.37 - Veristand execution mechanism

In NI VeriStand the loops are subjected to the following priorities:

1. Low priority L
2. Medium priority M
3. High priority H

The priority of loops is fixed it is possible only to modify the execution rate of the model. Still referring to Figure 6.37, note that all the loops are subordinated the *Primary Control Loop*. This, is responsible for the execution of the external *stimulus profile*, which is the test executive that can call real-time sequences, open and close NI VeriStand projects, and perform data-logging and pass/fail analysis. The stimulus also connects real-time sequences to system definition files to bind channel data within the system definition file to variables in the real-time sequence.

The Primary Control Loop is also encharged to send data to *Data Processing Loop*, *Model Execution Loop*, to create the mapping connections and to prompt the *Data Pocessing Loop* and *Model Execution Loop*. Since the Primary Loop Control is the loop of reference, if the time step of the Mean Value Model is higher than the time step of the primary loops, this may cause the overrun of the system. The root cause stems from the non-synchronization between the model and the real-time target.

Although it is not possible to change the priority or primary tasks of engine loops, the execution rate of the loop can be customized. Therefore, the target execution rate need to be chosen in order to benchmark the execution speed (i.e. the maximum target rate) without increasing the loop counts.

The Prototyping Interface - ETAS ES910

The Prototyping Interface, also known as Rapid Prototyping (RP) device, embraces the controller code that communicate with the Mean Value Model through NI VeriStand. The available Prototyping Interface at Politecnico di Torino is the ETAS ES910.

First, the control code must be converted from the Simulink code to an ETAS code through the software named *INTECRIO*. This is the prototyping software-tool that performs the configuration of the ES910 module in a simplified way, is able to manage different modeling tools at the same time, in a complete virtual prototyping or rapid prototyping system.

The ETAS ES 910 is provided with CAN and LIN interfaces for the connection of the RP module to vehicle and PXI buses. It is possible to configure the two line of CAN in a separate way by imposing a different rate of data exchange.

ETAS ES910 can also be implemented for the rapid prototyping phase for the development of an ECU.



Figure 6.38 - ETAS ES910 Rapid Prototyping Device

Coupling the PXI with the ETAS RP device

The coupling between the PXI (Mean Value/Real-Time engine model) and the ETAS ES910 (the control code) is made possible provided that the control algorithm has been translated into an ETAS Real-Time (RT) code through the software INTECRIO.

The physical connection between the hardware is represented in Figure 6.40. The core of the Hardware in the Loop testing procedure is to verify the control algorithm computational time on the rapid prototyping device (ETAS910) which has processing speed of 800 MHz.

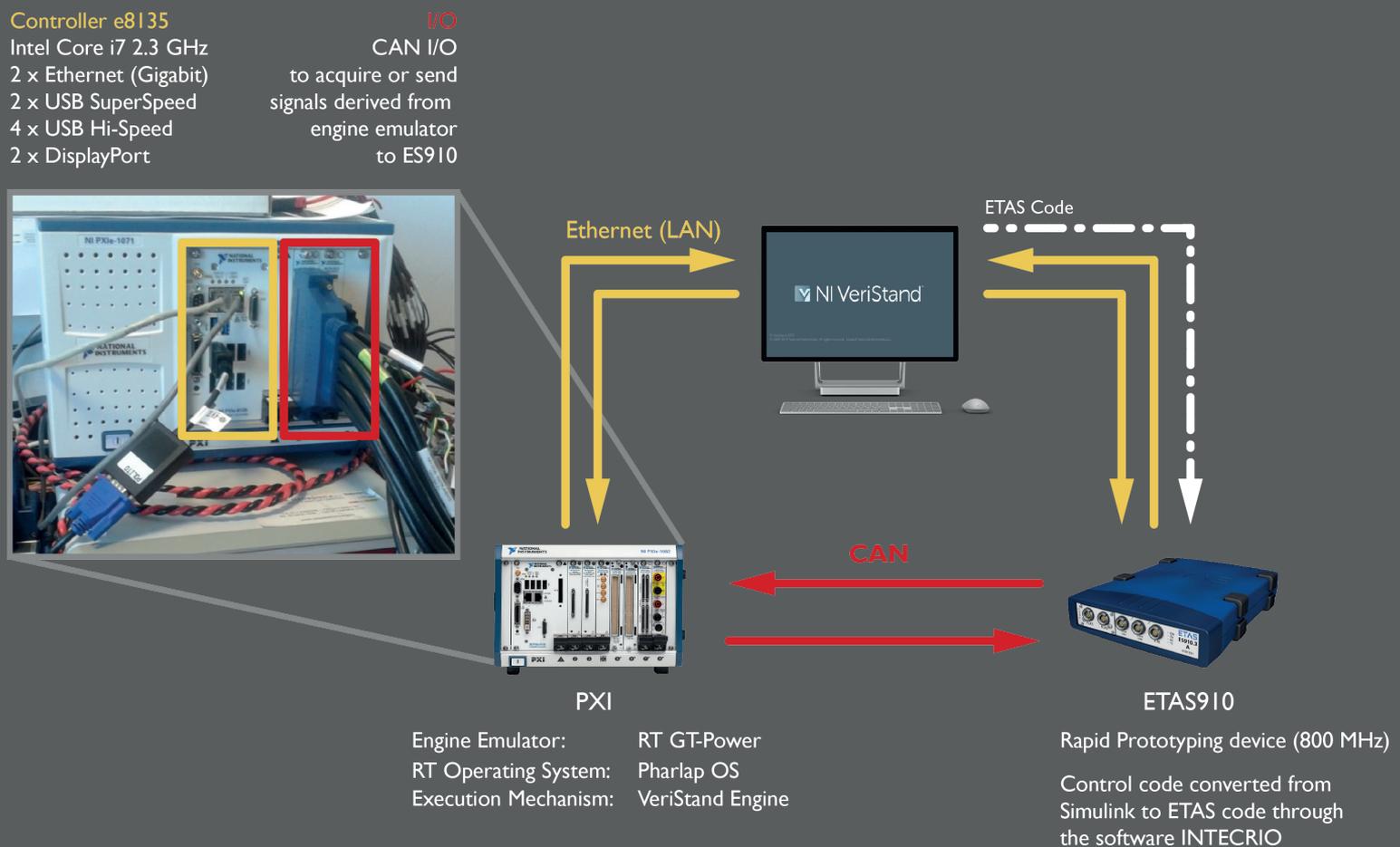


Figure 6.40 - Hardware connections in the HiL testing procedure

7 *THE REAL TIME MODEL*

Development of the RT Model

From FR to MV Model

Integration of MVM in HiL testing

7.1 Development of the RT Model

The Real Time Model (RTM) represents a step further regarding the decreasing of the model Factor of Real Time. Actually, the RTM is the evolution of the MVM, which is also used for Hardware in Loop testing procedure. In GT-Power, it is possible to run the simulation at real time assuming that the RT hardware and software are used. The RT simulation is the best practice for HiL testing.

The first main difference from MVM to RTM is that the real time simulation requires a GT-SUITE-RT product license and solver files. GT-SUITE-RT does not store RLT variables in order to save computation effort and to ensure deterministic turnaround times throughout the entire cycle. Such condition may be changed whether specifically requested by the user. In this specific case, no results have been saved so that the simulation run as fast as possible.

Using the RT license, the RTM is built up by adapting the MVM through components named *RLTCreatorRT*. This template allows to create RLT variables that can be used with GT-SUITE-RT. The minimum variables to be stored through *RLTCreatorRT* are engine speed and BMEP. In this way, these two quantities will be stored so that they are accessible by the rest of the model. Figure 7.1 represents this adjustment of the RTM.

The second part of the development consists in creating the interface for communicating with Simulink. Provided that the RT model can either run standalone or coupled with Simulink, in case of coupling the output of interest must be passed to Simulink through the *SimulinkHarness* part. However, this implementation has not been requested. The focus of the development was the lower FRT.

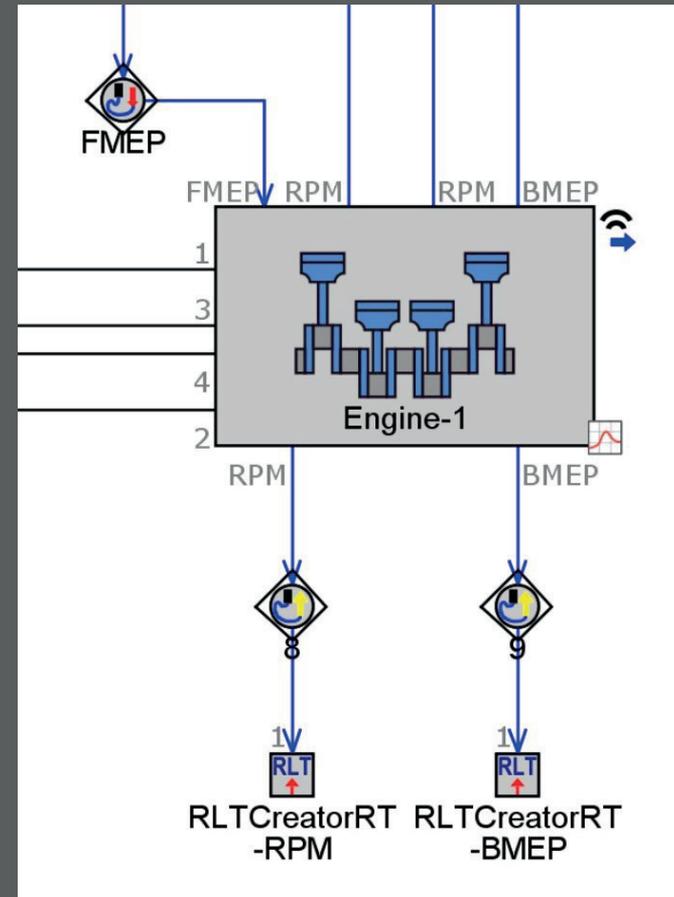


Figure 7.1 - Real-Time Model conversion parts

7.2 Comparison between the DET, FR, MV and RT Models

The comparison between the Detailed, Fast-Running, Mean Value and Real-Time models still refer to the engine map provided in Table 4.2 and Figure 4.20. As expected, the RT model returns the lowest Factor of Real Time. Table 7.1 summarizes the trend of the FRT. Since the GT-SUITE-RT does not store the variables by default settings, the available plots are the followings:

- | | |
|--------------------------------|---------------------------------|
| 1. Factor of Real Time | Exhaust Manifold Pressure .7 |
| 2. Injected Mass Flow Rate | Exhaust Manifold Temperature .8 |
| 3. Brake Torque | Intercooler Out Pressure .9 |
| 4. BMEP | Intercooler Out Temperature .10 |
| 5. Intake Manifold Pressure | VGT Controller .11 |
| 6. Intake Manifold Temperature | EGR Actuator .12 |

GT-Power Engine Model	Average Factor of Real Time	Gain
Detailed	50,428	
Fast-Running	5,674	- 88,75 %
Mean Value	2,858	- 49,62 %
Real-Time	2,251	- 21,24 %

Table 7.1 - FRT results for DET, FR, MV and RT models

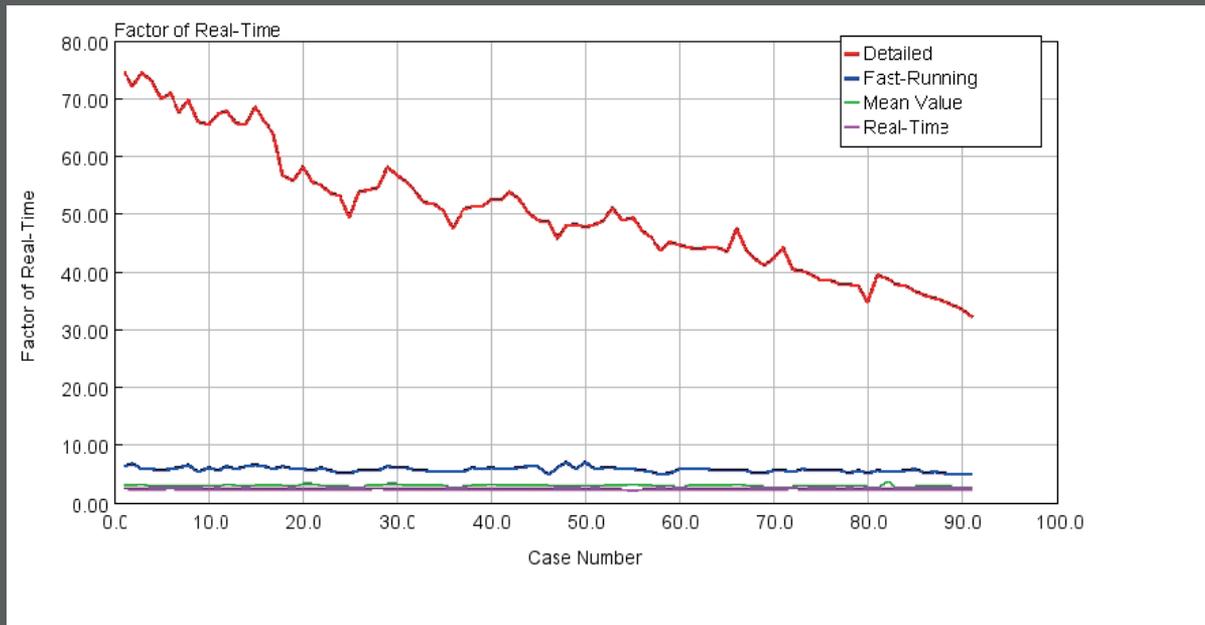


Figure 7.2 - Factor of RealTime, DETM vs FRM vs MVM vs RTM

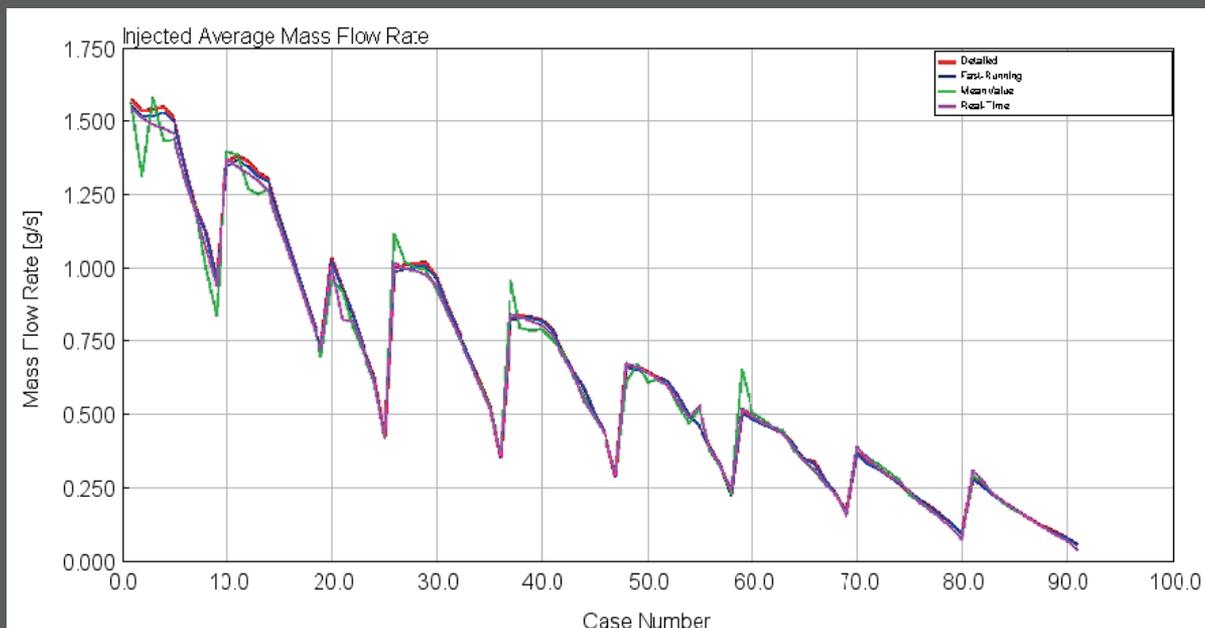


Figure 7.3 - Injected Average Mass Flow Rate, DETM vs FRM vs MVM vs RTM

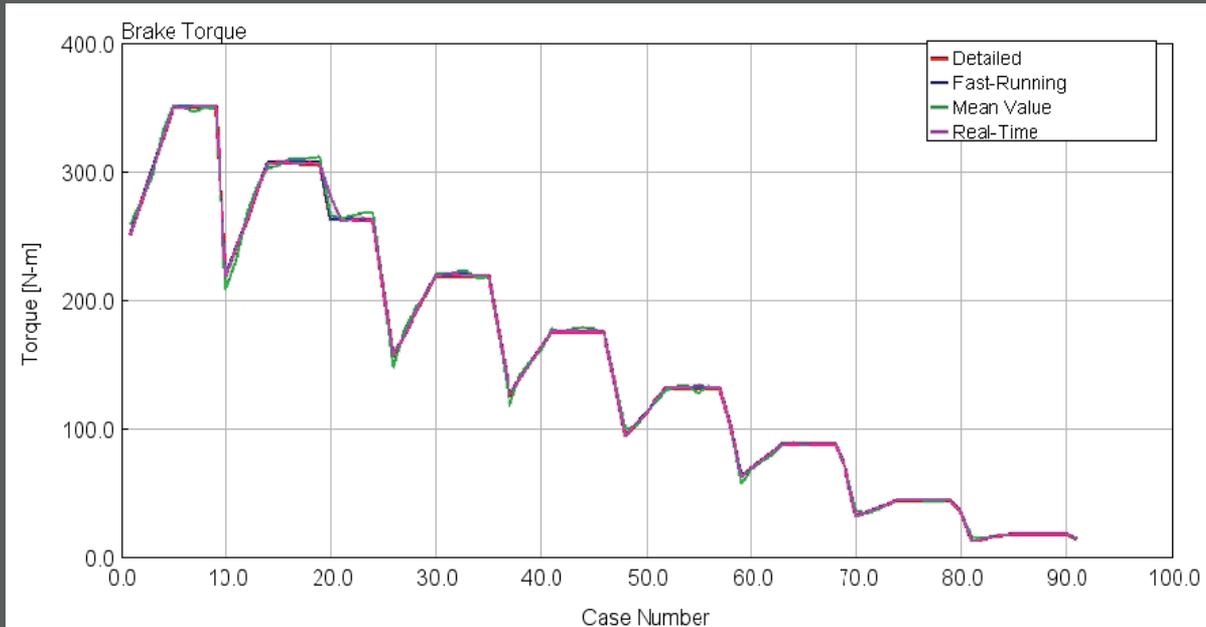


Figure 7.4 - Brake Torque, DETM vs FRM vs MVM vs RTM

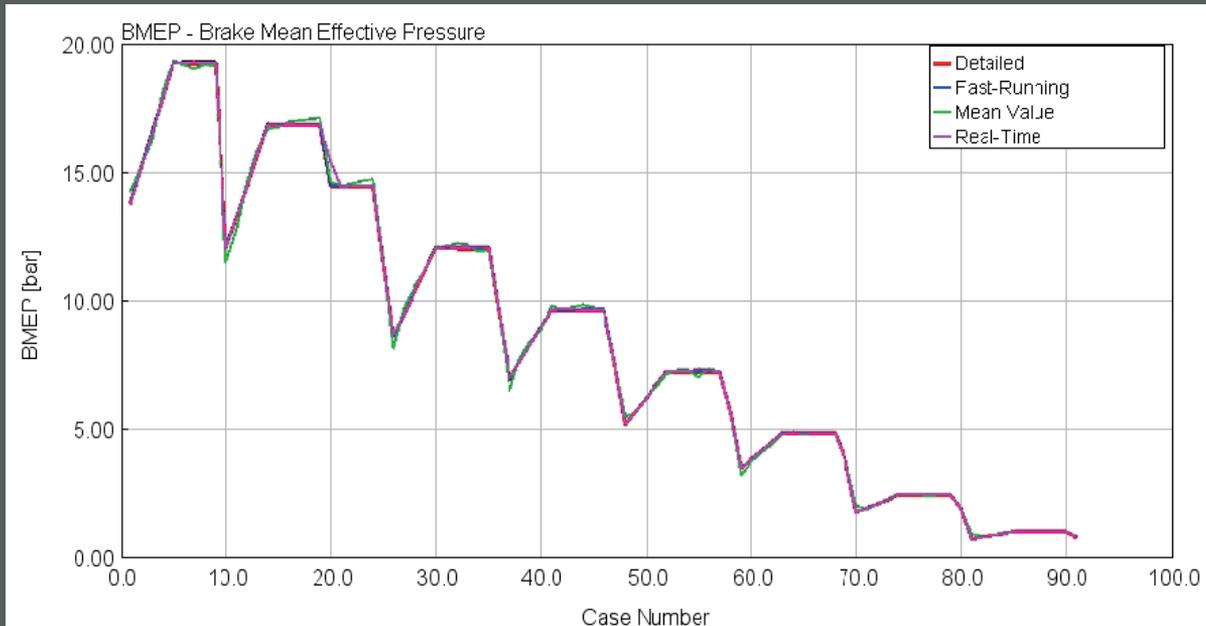


Figure 7.5 - BMEP, DETM vs FRM vs MVM vs RTM

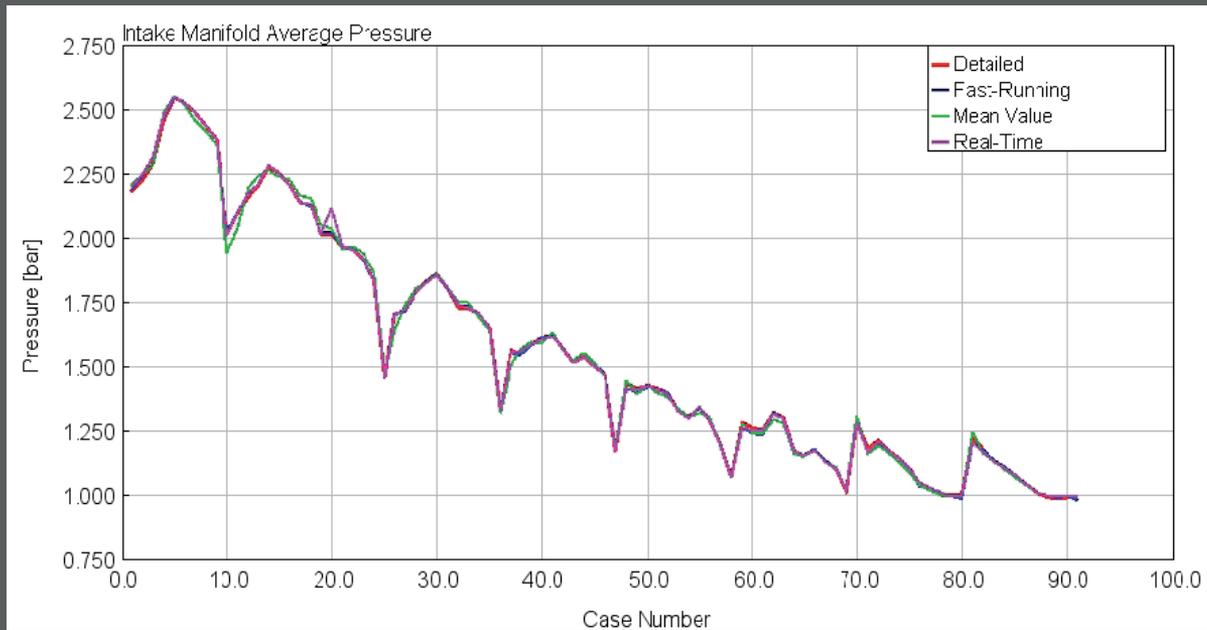


Figure 7.6 - Intake Manifold Pressure, DETM vs FRM vs MVM vs RTM

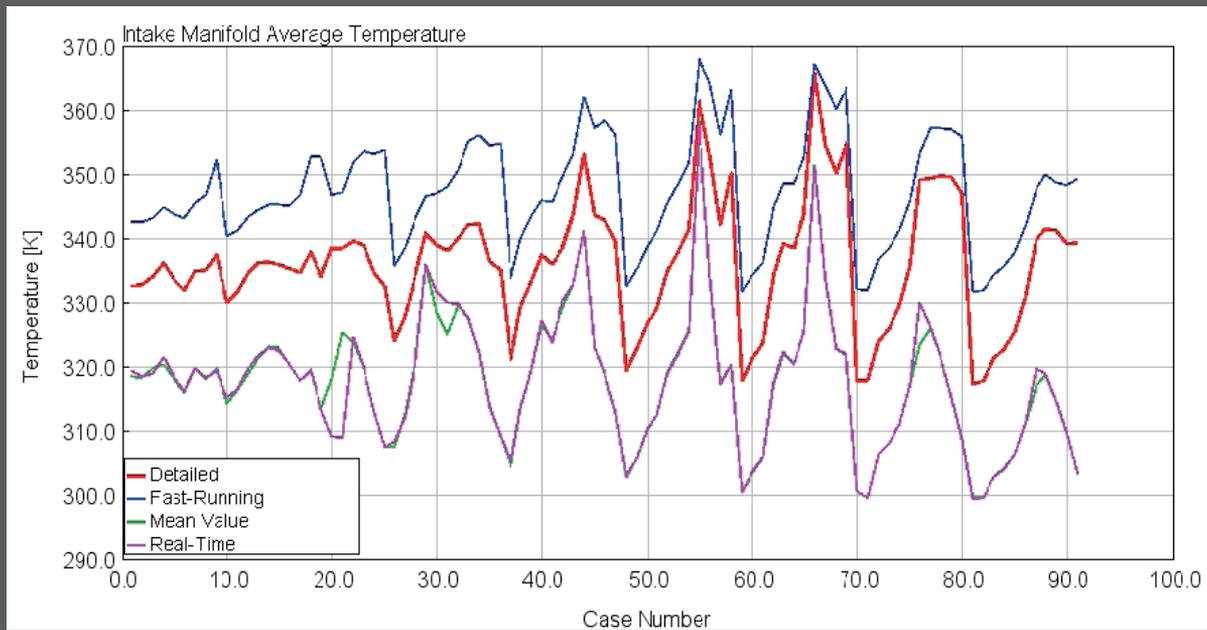


Figure 7.7 - Intake Manifold Temperature, DETM vs FRM vs MVM vs RTM

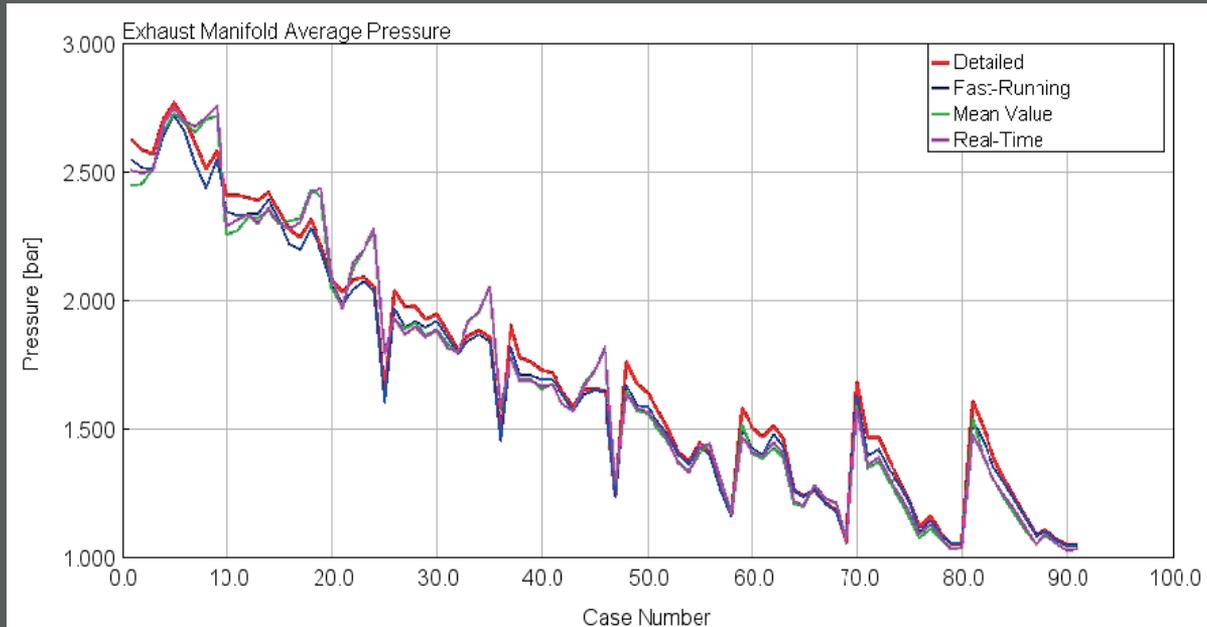


Figure 7.8 - Exhaust Manifold Pressure, DETM vs FRM vs MVM vs RTM

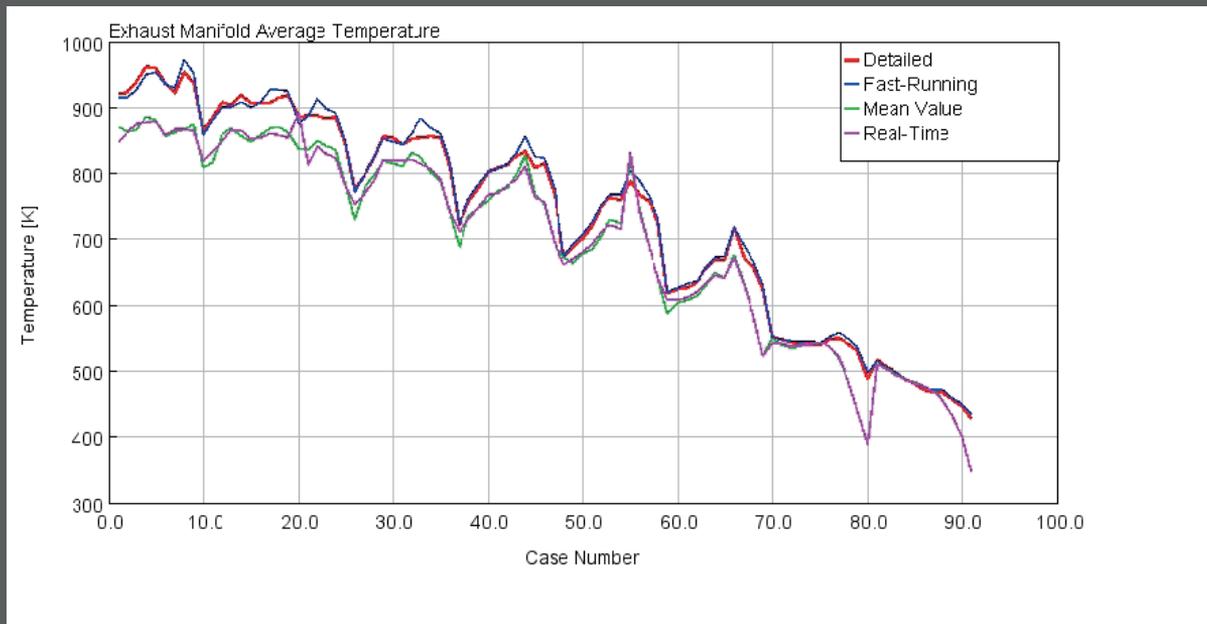


Figure 7.9 - Exhaust Manifold Temperature, DETM vs FRM vs MVM vs RTM

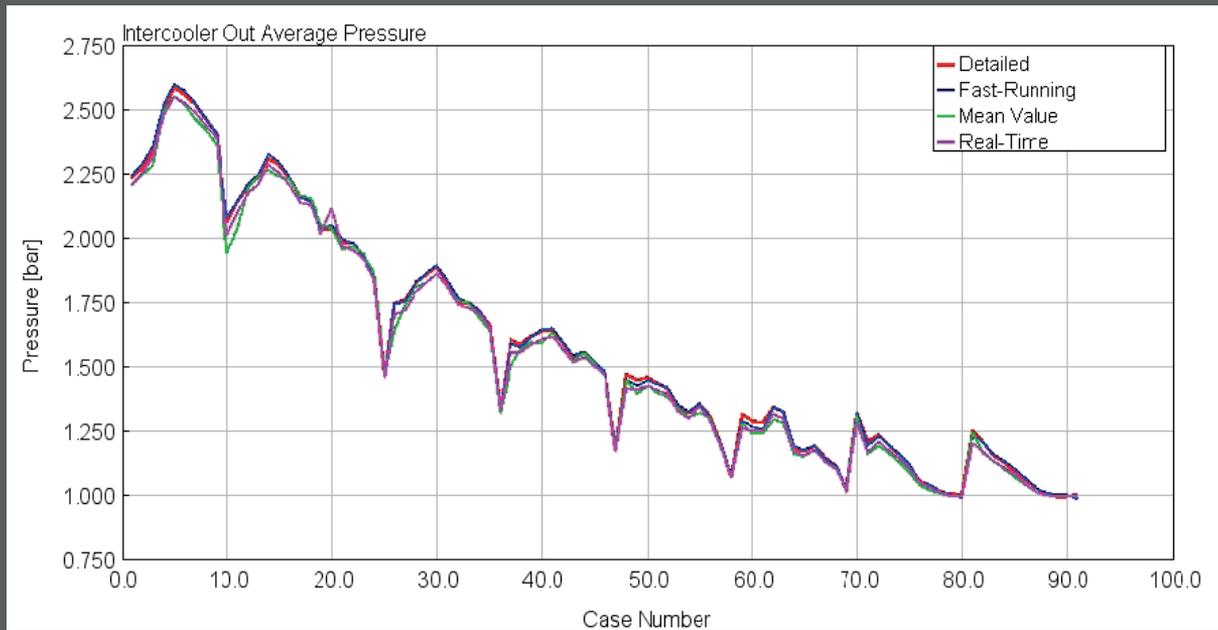


Figure 7.10 - Intercooler Out Pressure, DETM vs FRM vs MVM vs RTM

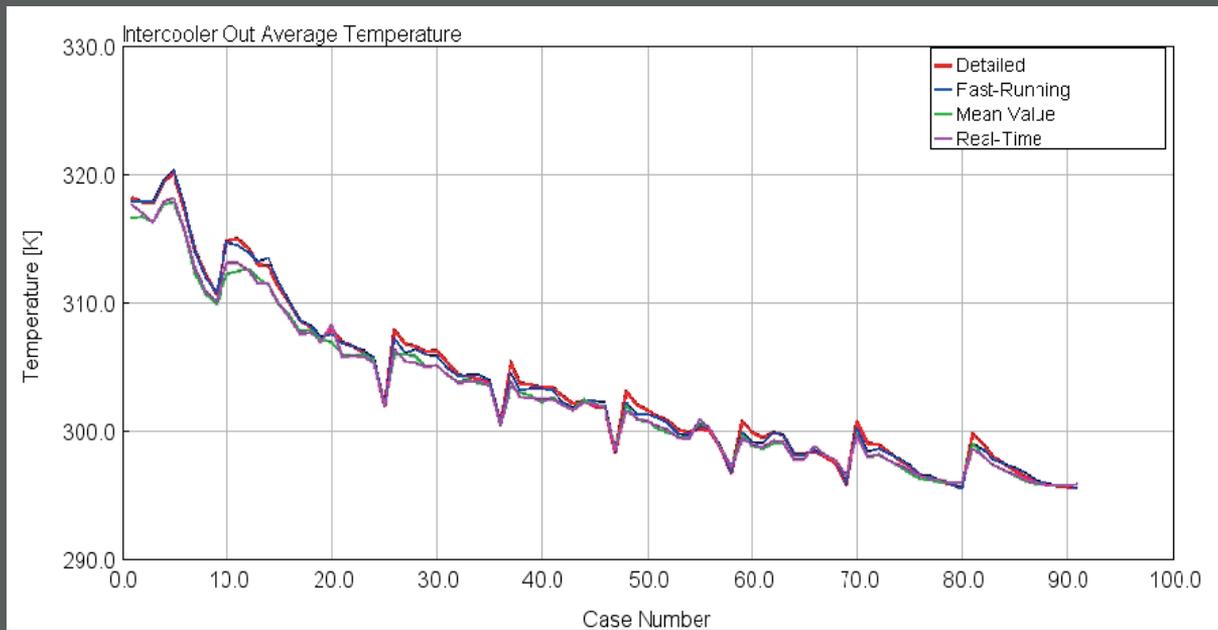


Figure 7.11 - Intercooler Out Temperature, DETM vs FRM vs MVM vs RTM

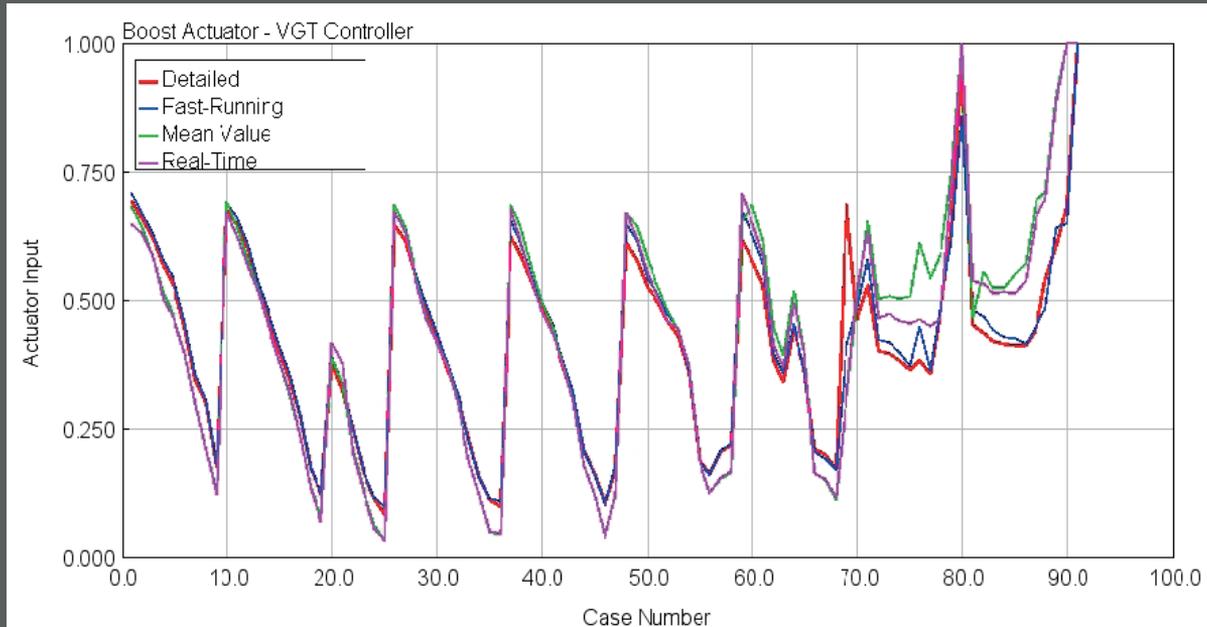


Figure 7.12 - VGT Controller, DETM vs FRM vs MVM vs RTM

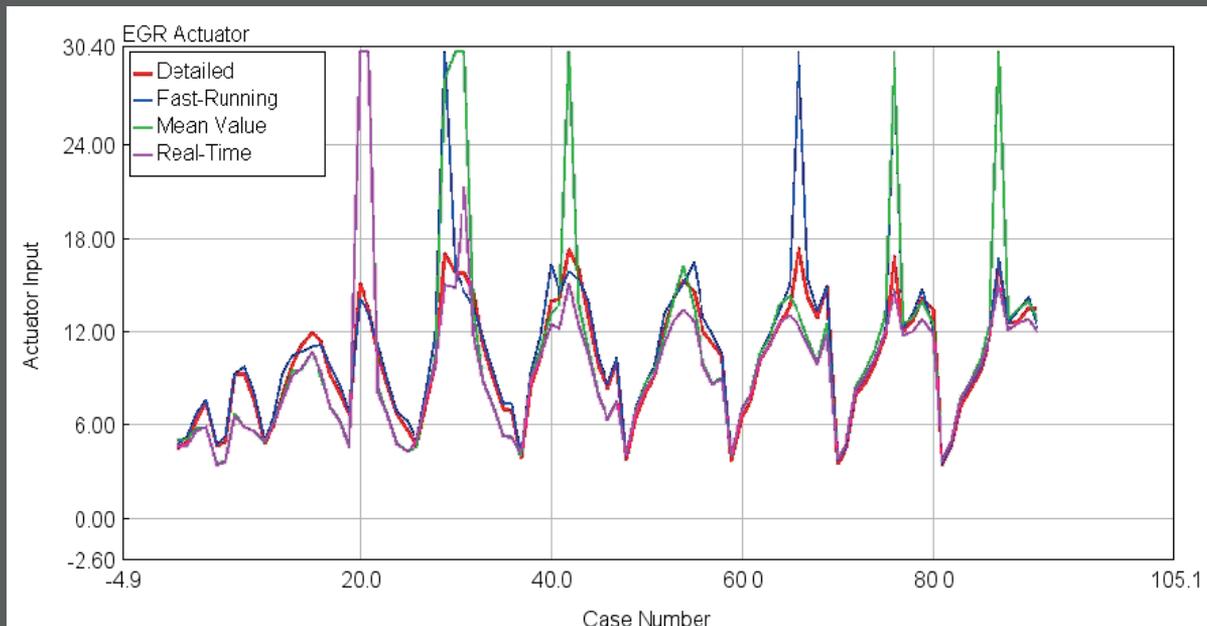


Figure 7.13 - EGR Actuator ww, DETM vs FRM vs MVM vs RTM

8

CONCLUSION

In this Master Thesis the core activity was to develop and validate the Real-Time model of the FPT FIA engine starting from the Detailed model provided by FPT. The switching and evolution from the Detailed to the Real-Time model has been obtained with GT-Power software by Gamma Technologies LLC. The purpose of the Real-Time model is the Hardware in the Loop testing phase to achieve the target mission together with the controllers in development.

First, the development of the Mean Value model has requested to manage the Detailed model, to understand its components and their working principles, as well as the requested input and the desired output. The Detailed model has been obtained from experimental measurements performed on the real engine in the test bed.

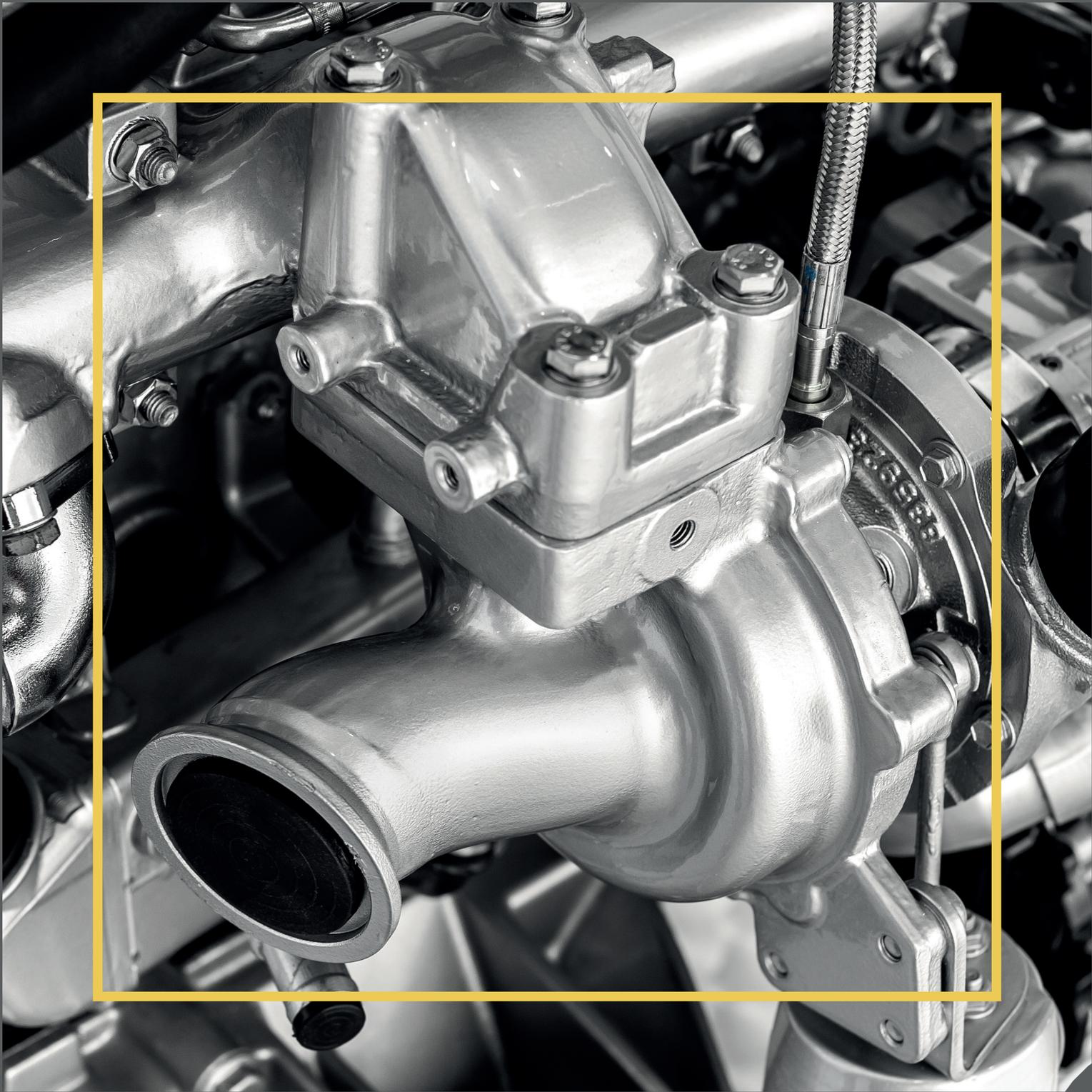
Afterwards, the Detailed model has been rearranged to comply with the Design of Experiment procedure. The target of the DoE is to create 2000 random cases to gather all the information concerning the engine parameters of interest. The input parameters for this phase are: engine speed, intake manifold pressure and temperature, exhaust manifold pressure, EGR fraction, fuel mass rate, and the start of combustion of the main injection. In each experiment generated by the DoE we have a random value of each parameter chosen from a user-defined range of values.

At this stage, the information are exploited to train the Neural Network of the key-parameter that will be supplied to the Mean Value Cylinder: volumetric efficiency, exhaust gas temperature and indicated mean effective pressure. The Neural Network creates a map per each parameter. In this way, the Mean Value Cylinder recalls the Neural Network map and so reducing the computational time in the successive simulations. In addition, a Neural Network for the Friction Mean Effective Pressure has been also trained.

Once the Neural Networks have been trained, again starting from the Fast-Running model (different from that used for the Design of Experiment), the Mean Value Model has been developed by including the Neural Network and by recalibrating the injection system. This because it is required that the Start of Injection of the main pulse is the only variable parameter; meaning that the start of injection of both the pilot and pre injection must refer to the start of injection of the main event. To ensure this to happen, the Dwell Time maps were reconstructed and supplied to GT-Power. In doing so, the start of injection of the main pulse is the independent parameter whereas the pilot and pre events are obtained by considering the fixed time step between the events given by the Dwell Time maps.

The validity of the Mean Value model passes through the comparison with the Detailed model, rather than the Fast-Running model. This because the Mean Value model must trace the trend of the experimental data directly measured on the real engine and virtually translated through the Detailed model. Eventually, the GT-Power license for Real-Time application provides to further evolve the Mean Value Model into the Real-Time model. The results of the Real-Time model show that it sticks to the Detailed model, but having a positive gain of -95,53% on the Factor of Real Time, compared to the Detailed Model. This fast response of the Real-Time model allow to use this latter in Hardware in the Loop testing phase.

Future works provide to implement the Real-Time model of the FPT FIA and the controllers in development into the NI PXI engine emulator. In addition, the Low Pressure of the engine model must be still verified.



-
- [1] A. Mittica, S. D'Ambrosio, *Combustion engines and their application*, Politecnico di Torino, 2017/2018
- [2] C. Novara, G. Belforte, *Automatic Controls*, Politecnico di Torino, 2017/2018
- [3] E. Spessa, *Design of Engine and Control System*, Politecnico di Torino, 2018/2019
- [4] F. Millo, *Engine Emission Control*, Politecnico di Torino, 2018/2019
- [5] Hannu Jääskeläinen, Magdi K. Khair, *Combustion in Diesel Engines*, 2010
- [6] J.E. Dec, *A Conceptual Model of DI Diesel Combustion Based on Laser-Sheet Imaging*, SAE Technical Paper 970873, 1997
- [7] C. K. Westbrook e J. E. Dec, *Diesel Combustion: An integrated view combining laser diagnostics, chemical kinetics, and empirical validation*, SAE technical paper, 1999.
- [8] Dec J., SAE 970873
- [9] Dec J. and Tree D., SAE Paper No. 2001-01-1295
- [10] Flynn et al., SAE Paper No. 1999-01-0509
- [11] W.L.Hardy, R.D.Reitz, *A study of the effects of High EGR, high equivalence ratio and mixing time on emission levels in a Heavy-Duty diesel engine for PCCI combustion*, SAE Technical Paper No. 0026, 2006.
- [12] J.B. Heywood, *Internal Combustion Engine Fundamentals*, McGraw Hill, 1988
- [13] T. Kamimoto and M. Bae, SAE Transactions – Journal of Engines (1988), SAE 880423
- [14] SAE 2005-01-0379
- [15] Horiuchi, M., K. Saito, S. Ichihara, *The Effects of Flow-through Type Oxidation Catalysts on the Particulate Reduction of 1990's Diesel Engines*, SAE Technical Paper 900600, 1990, doi:10.4271/900600.
- [16] Gamma Technologies, *GT-Power Training - Engine Performance Analysis*
- [17] GT-Suite, *Engine Performance Application Manual*, Gamma Technologies, 2017.
- [18] Finesso R., Marellò O., Misul D., Spessa E. et al., *Development and assessment of pressure-based and Model-based techniques for the MFB50 control of a Euro VI 3.0L Diesel Engine*, SAE Int.]engines 10(4:2017)
- [19] Marco Cavazzuti, *Optimization Methods: From Theory to Design Scientific and Technological Aspects in Mechanics*, Springer-Verlag Berlin Heidelberg, 2013.

SOU
RC
ES

SITO
GRA
PHY
