# Study of the Influence of Some Factors Regarding the Operation of the Internal Combustion Engine

Marius SINGUREANU and Ion COPAE

***Abstract*—The paper highlights the main possibilities of studying the influence of various parameters on the engine functionality. Mathematical algorithms are applied, which allows us to study these issues based on experimental data gathered throughout specific tests, carried out on certain vehicles. These algorithms are applied on vehicles with) electronic control for various automotive systems, thus the data was gathered from the vehicles built in CPUs. To this purpose we will show how to use sensitivity analysis, variance analysis, information theory and correlation analysis in automotive engineering.**

***Index Terms*—internal combustion engine, fuel consumption, correlation analysis, engine output power.**

## I. Introduction

The study of various factors over the vehicle's engine functionality is a continuous preoccupation for engineers. In the specific literature we can find quantitative and qualitative appreciations regarding the influence of functional, tuning, constructive and exploitation parameters over the power performance, fuel consumption and polluting emissions. We have to mention that in the field literature the influence of various factors is being analyzed following a very restrictive methodology like: to study the influence of just one factor while the others are considered to remain constant [2], which is obviously not in concordance with the reality. Throughout the paper this restriction is eliminated and functional interdependencies will always be emphasized/underlined especially in the case of vehicles that have ECU on board. Moreover, experimental research has confirmed that the parameters are not constant in time when dynamic regimes are present throughout exploitation.

## II. Simulation Configuration Steps

As a consequence the aim of the paper targets how some parameters (also called factorial parameters) affect engine performance. To this purpose experimental data are being used and mainly those parameters that are gathered from the vehicle's on board ECU equipped with gasoline injection systems; thus we target the measurable functional parameters like: engine speed $n$ and engine load (through the help of the throttle's position $\xi$ or intake air pressure $p_a$), ignition advance $\beta$ the quality of fuel – air mixture (through the help of air excess coefficient $\lambda$), injection duration $t_i$, etc.

The parameters over which the mentioned influencing parameters are analysed can be fuel consumption (through the help of hourly fuel consumption $C_h$, specific effective fuel consumption $c_e$, etc.), power performance (through engine power $P_e$, engine torque $M_e$, etc.); these encompass resulting parameters.

M. SINGUREANU is with the Military Technical Academy "Ferdinand I", Bucharest, Romania (e-mail: marius.singureanu@yahoo.com).
I. COPAE is with the Military Technical Academy "Ferdinand I", Bucharest, Romania (e-mail: copaeion@yahoo.com).

A first study procedure for the influence of certain parameters over the engine's running is *sensitivity analysis*. Sensitivity expresses the property of a parameter resulted by changing its value under the influence of certain factorial parameters. If we discuss only one factorial parameter we target simple sensitivity, otherwise we have to look at multiple sensitivity. Sensitivity encompasses a function that may vary (case where there is hetero-sensitivity) or may be constant (case where there is iso-sensitivity).

By definition, simple sensitivity is established by the following relation

$$S(y/x) = \frac{x}{y}\frac{\mathrm{d}y}{\mathrm{d}x} \tag{1}$$

where $x$ is the influencing factor (the factorial parameter), and $y$ is the resulting parameter.

From (1) we can see that the sensitivity is adimensional, thus $S$ is also known as sensitivity coefficient.

For example, if the intention is to establish the influence of the throttle $\mathrm{d}\xi$ and engine speed $n$ over the hourly fuel consumption $C_h$ (so efficiency is targeted), then we establish the following sensitivity functions:

$$S(C_h/\xi) = \frac{\xi}{C_h}\frac{\mathrm{d}C_h}{\mathrm{d}\xi};\; S(C_h/n) = \frac{n}{C_h}\frac{\mathrm{d}C_h}{\mathrm{d}n} \tag{2}$$

In the expressions from (2) all data are known from experimental tests or are calculated based on them (including the partial derivations); from (2) we can say that sensitivity function varies in time, thus we have a hetero-sensitivity, because all the parameters involved vary.

## III. Simulation Results

Fig. 1 presents the average values on each test for the sensitivity function in the case of 50 test runs carried out on a Logan Laureate vehicle; the resulted parameter is hourly fuel consumption, and the graphs prove the existence of different average values for different test runs.

Fig. 1 also shows that when looking at the big picture at all test runs, the most significant influence over the hourly fuel consumption is due to engine speed. The graphs also show that overall the engine speed influences twice as much as air-fuel mixture does (air excess coefficient) and almost 5 times as much as the throttle's position has (engine load).

Digital Object Identifier 10.32754/JMT.2019.2.05

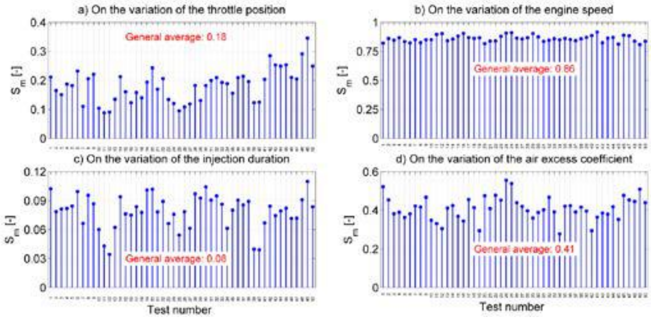


Figure 1. Average values for sensitivity function of hourly fuel consumption, 50 experimental test runs on Logan Laureate vehicle

The study on the influence of functional parameters can also call on dispersional analysis, better known under the name of *variance analysis* (ANOVA – **AN**alyse **O**f **Va**riance, MANOVA – **M**ultivariate **AN**alyse **O**f **VA**riance); dispersion, also called variance, has a special importance in the analysis on the influence of certain parameters onto the development of a certain dynamic process [1; 4].

The English mathematician and statistician Ronald Fisher, the creator of dispersional analysis, proved that by estimating the dispersion of a certain characteristic undergoing the influence of a parameter, and then eliminating its influence and comparing the two dispersions, we get quantitative information referring to this influence. As a result, dispersional analysis is all about comparing the two types of dispersions, factorial and residual. If the factorial dispersion is higher than the residual, then that specific factor has a sensitive influence on the analyzed process. Otherwise, if the factorial dispersion (individual or interacting with another factor) is lower than the residual one, then that specific factor has a negligible influence over the targeted process. Practically this comparison can be made by establishing the contribution of each factor in percentages and of the residual on the total dispersion.

Fig. 2 presents the results obtained by applying the generalized MANOVA algorithm the targeted parameters and the afferent interactions are being considered, and by studying the influence of engine speed $n$, throttle's position $\xi$, intake air pressure $p_a$ and air excess coefficient onto engine output power in the case of 50 experimental test runs carried out on a Logan Laureate vehicle. Afferent to this example, Fig. 3 presents the actual values (for each parameter in Fig. 3) and $D$ the dispersion images of functional parameters in the case of 50 experimental tests.

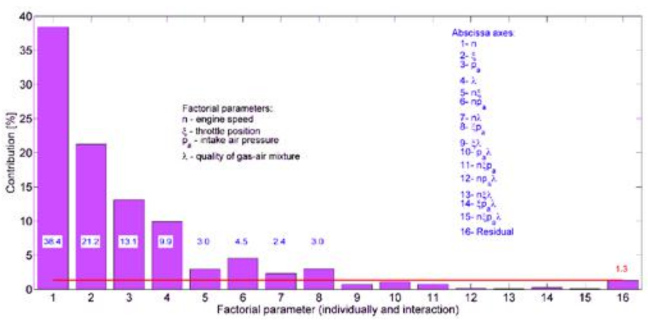


Figure 2. Study on the influence of certain factors over engine output power by applying generalized MANOVA algorithm, 50 experimental test runs, Logan Laureate vehicle.

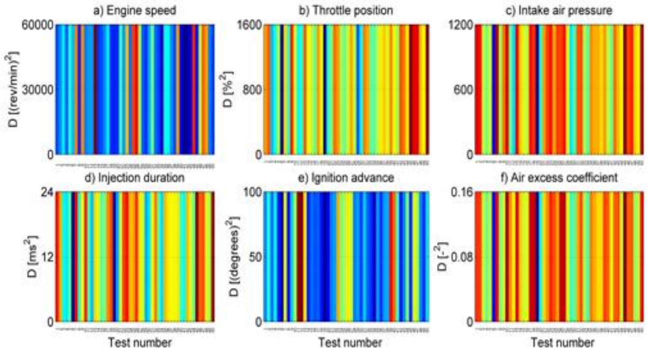


Figure 3. Values and dispersion images, 50 experimental test runs Logan Laureate vehicle

We can infer from Fig. 2 that the residual dispersion represents 1.3% from total dispersion; values higher than this have the dispersion afferent to engine speed (38.4%), throttle's position (21.2%), intake air pressure (13.1%), and the quality of air-fuel mixture by air excess coefficient (9.9%). Added to that, values higher than residual dispersions have the interactions engine speed - throttle's position (3.0%), engine speed-intake air pressure (4.5%), engine speed – excess air coefficient (2.4%) and throttle's position – intake air pressure (3.0%); the others have lower values than the residual dispersion, and therefore are not mentioned. So, engine speed and throttle's position have the most significant influences over the engine's output power, the first factor having an influence of about 1.4 times higher.

The graphs from Fig. 3 confirm the fact that various functional parameters have different influences on each test run and overall over engine output power.

The influence of certain factors over the engine's performance has an explicit interest, like in the presented examples, but has another interest also, that of prediction; to this purpose algorithms and concepts from *information theory* can be applied, thus calling on entropy and information [3, 5].

As it is already known, in order to characterize the uncertainty of a certain event, we use the entropy concept, as information represents the fundamental concept in predictions. The higher the entropy, the higher the uncertainty will be, consequently the prediction will also be lower.

Besides, mutual information constitutes a concept that offers a quantitative measure for uncertainty reduction, hence the increase of prediction degree. The higher the mutual information, the lower the uncertainties will be, therefore the more accurate the predictions. Mutual information is a basic concept when studying the evolution of processes and systems and it represents a measure of parameters interdependency. For this reason, when establishing mathematical models we have to choose those parameters that are characterized by the highest values of mutual information, because they ensure the best predictions; these parameters are called relevant parameters, attached to the concept of relevance.

For the reasons that were mentioned, it is considered that information theory constitutes a generalization of classic correlation, and mutual information represents a measure of relevance.

 Digital Object Identifier 10.32754/JMT.2019.2.05

Fig. 4 presents a graph where in its knots targeted parameters and their entropy values *H* are shown, and on the arches the values for the mutual information $I_{x,y}$. The deduced parameter is hourly fuel consumption in the case of 50 test runs carried out on a Logan Laureate vehicle. The mentioned parameter is presented in the superior part of the graph (so we target the vehicle's efficiency); the other 6 parameters constitute factorial parameters.

The graph from Fig. 4 shows that the pair hourly fuel consumption – engine speed has the highest value for mutual information (0.959 bits), followed by the pair hourly fuel consumption – air excess coefficient (0.363 bits); thus engine speed and air excess coefficient are the first top two relevant choices. Therefore if two mathematical models are established referring to efficiency, of type $C_h = f(n,\lambda)$, respectively $C_h = f(\xi,t_i)$, the first one will ensure a better prediction (a smaller modelling error) than the second for the values of hourly fuel consumption $C_h$.

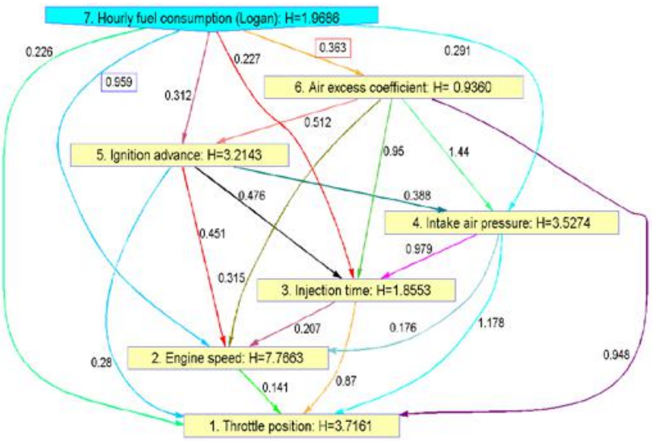


Figure 4. Graph that contains entropy and mutual information for 6 factorial information and the deduced parameter hourly fuel consumption, 50 test runs, Logan Laureate

The final aspect that was mentioned is confirmed by Fig. 5, where results are being presented in the case of some mathematical models on which the factorial parameters that were used are engine speed *n* and air excess coefficient $\lambda$, respectively throttle position $\xi$ and injection duration $t_i$ (Fig. 5b); on both models the resulting parameter is hourly fuel consumption $C_h$.

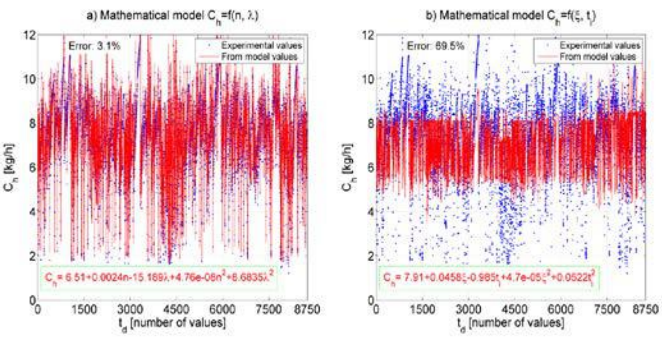


Figure 5. Establishing mathematical models based on information theory, for hourly fuel consumption 50 test runs of Logan Laureate vehicle

As we can see maximum prediction (acceptable modelling error of 3.1%) is being ensured by the mathematical model from Fig. 5a, on which the factorial parameters are the two relevant parameters that have the highest value for mutual information in Fig. 4 (0.959 bits and 0.363 bits). In exchange, the simulation error is higher in the case of Fig. 5b (69.5%, an unacceptable value), where the two variables with the smallest mutual information of Fig. 4 (0.226 bits and 0.227 bits) can be identified.

In the lower part of the graphs from Fig. 5 the expressions afferent to the two targeted mathematical expressions are presented; examination of Fig. 5a reveals that the generalized mathematical models (for 50 experimental test runs):

$$C_h = 6.51 + 0.0024n - 15.189\lambda + 4.76\cdot 10^{-8} n^2 + 8.6835\lambda^2 \quad (3),$$

which allows the calculus of hourly fuel consumption for the engine to depended on engine speed and its load latter through the intake air pressure).

Finally, in order to highlight the dependence character (linear or nonlinear) between the factorial and the deduced parameters, *correlation analysis* will be applied.

As we already know from classical statistics, simple correlation analysis targets the connection between a certain deducted parameter *y* and a factorial parameter x (influence factor). The index which is the most used to appreciate linear dependence between two variables is correlation coefficient (Pearson's coefficient), established with the expression [4, 5]:

$$\rho_{xy} = \frac{R_{xy}(0)}{\sqrt{R_{xx}(0)R_{yy}(0)}}, \quad (4)$$

with values $\rho \in [-1;1]$, a maximum possible inter-correlation (a perfect linear dependency) being for $\rho^2 = 1$. If $\rho = 1$ then we deal with a perfect direct linear dependency, and if $\rho = -1$ then we deal with a perfect indirect linear dependency; if $0 < \rho \le 1$ we deal with a direct dependency, and if $-1 \le \rho < 0$ there is an indirect dependency. So, as much as $\rho^2$ is further away from the value of 1 without (reaching the value of zero) the nonlinearity is highly accentuated.

In (4), at the abaci position we have the inter-correlation function in the origin of discrete time, meaning for $\tau = 0$, and under the square root are the self-correlation functions still for $\tau = 0$ (meaning its maximum values).

In the case of multiple correlation, the simultaneous influence of two or more factorial parameters (influence factors) over the deduced parameter. In this situation, the multiple correlation coefficient is used; it is calculated based on the simple correlation coefficient between the pair of parameters and taking into account the expressions for the correlation functions.

For example, the upper part of Fig. 6 presents the value for the simple correlation coefficients in the case of 50 experimental test-runs for Logan vehicle, and Fig. 6d the multiple correlation coefficients; the factorial parameters are the throttle's position $\xi$ (engine load) and engine speed *n*, and the deduced parameters are engine output power $P_e$ and hourly fuel consumption $C_h$.

Digital Object Identifier 10.32754/JMT.2019.2.05

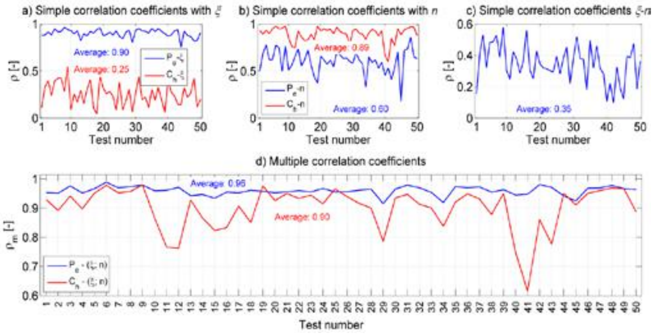


Figure 6. Values for simple and multiple correlation coefficients for engine output power and hourly fuel consumption, 50 experimental test runs, Logan Laureate Vehicle

## IV. Conclusion

From the graphs above we can identify a highly non-linear dependency between hourly fuel consumption and throttle's position (Fig. 6a), between engine output power and its speed (Fig. 6b), as well as between throttle's position and engine speed (Fig. 6c); these aspects have implications over the mathematical models established for the engine, which have to be mostly non-linear. The graphs from Fig. 6 also show another important aspect: multiple correlation coefficients have higher values than the afferent simple correlation coefficients; this aspect was to be expected, because the onboard computer oversees engine's operation based on the interdependence of several parameters.

It can be *concluded* that in order to study an engine operation it is necessary to analyze the concomitant influence of several factors, at the same time trying not to presuppose that some are constants, as the specific literature does. The study of various factors facilitates and supports the establishment of mathematical models for engine's operation mode.

## References

[1] Carey G., *Multivariate Analysis of Variance (MANOVA)*, Colorado State University, 1998

[2] Copae I., *Electronic Control of the Operation of Internal Combustion Engines*, Military Technical Academy Printing House, Bucharest, 2000.

[3] Gray R., *Entropy and Information Theory*, Stanford University, New York, 2007

[4] Murtagh F., *Multivariate Data Analysis*, Queen's University Belfast, 2000

[5] Watanabe S., *Information Theoretical Analysis of Multivariate Correlation*, IBM Journal,1990

 Digital Object Identifier 10.32754/JMT.2019.2.05