UNIVERSITY OF MODENA AND REGGIO EMILIA

METHODS OF SCIENCES AND ENGINEERING

THREE-YEAR DEGREE IN MANAGEMENT ENGINEERING

Development of a mathematical model for the two-legged walking: Fitting and study of a multivariate distribution that approximates the performance

Speaker: Graduating: Prof. Fabrizio Pancaldi Thomas Caleri Supervisor: Prof. Claudio Giberti

2017/2018 Academic Year

Index

1. Introduction

1.1 THE PEDESTRIAN WALKWAYS

In recent decades the pedestrian walkways have become increasingly important in contemporary architecture. With the continuous development of road networks are always looking for solutions to optimize the mobility of people. The pedestrian walkways allow crossing natural obstacles (such as rivers or gorges) and not natural (such as roads and railways). The realization of these structures also, has encourages the use of the bicycle and walking, thereby decreasing the environmental impact due to motor vehicles.

The continued spread of pedestrian walkways so, as well as closing the man's need to overcome natural obstacles (and not), it tends to push more and more people to adopt a lifestyle more "green" and to develop awareness of the importance of safeguarding the planet.Figure 1a shows the pedestrian walkway in Hong Kong while in Figure 2b shows the Millenium Bridge in London [1]

1.2 PROBLEMS RELATED TO THE DESIGN OF PEDESTRIAN WALKWAYS

The loads that these structures must support are generally low (400-500 kg / m2) for which, both for an aesthetic fact that design, are constructed using materials able to confer particular characteristics, such as slimness and lightness.

A case that catches the eye is the one concerning the "Millenium Bridge" in London. When it opened in 2000, thousands of pedestrians crossed it and after only two days were forced to close because of the resonance and vibration effects arising and which were not taken into account at the design stage. In addition, pedestrian walkways must be able to ensure the "Human comfort". The convenience of the pedestrian is therefore linked to possible vibrations and deformations that occur in the structure, hardly they pose a risk to safety, except for particular cases of resonance such as the one mentioned above, but can generate sensations of discomfort and fear to pedestrians in transit. For this we should try to keep the vibrations as far as possible from low frequencies perceptible by humans.

Over the years there have been several studies to search for a natural frequency threshold. For example, in NTC 2008 it has indicated that a naturala frequency higher than 5 Hz ensures a good level of comfort to the walker in the structures with rhythmic passage as the pedestrian walkways. The assessment of vibrations induced by humans is a much debated topic nowadays because there is no uniform method to define the inputs that the genres. [1][10].

1.3 AIM OF THE STUDY

The purpose of this thesis is to provide a statistical mathematical model that can be implemented in the study of stability of pedestrian walkways. In the first analysis we will introduce the theories and key concepts of the analyzed multivariate distributions. Then, using experimental data obtained thanks to a force floor, several statistical tests will be carried out through the use of the Matlab calculation software, to check whether there is a multivariate distribution that best describes the man walking process.

Figure 1a, pedestrian walkway in Hong Kong

Figure 1b, Millenium Bridge in London

2 The bipedal walking

First, it should point out that this argument uses as input the results obtained from the study done by Martina Fornaciari on human gait process (Development of a mathematical model for the two-legged walking: Applications in the study of stability of pedestrian walkways). This chapter shows some basics to introduce the aforementioned topic.

It will go to analyze the biped walking, or the process that allow people to move from one place to another, although it seems that a simple and automatic act to our eyes, so in reality is not, resulting in a rather complex mechanism. It is therefore essential, in the first instance, to understand what a complete walk cycle. It consists in the movement of the human body in the temporal window that runs from the beginning of a step with one foot up to the beginning of the next step with the same foot, as shown in Figure 2.

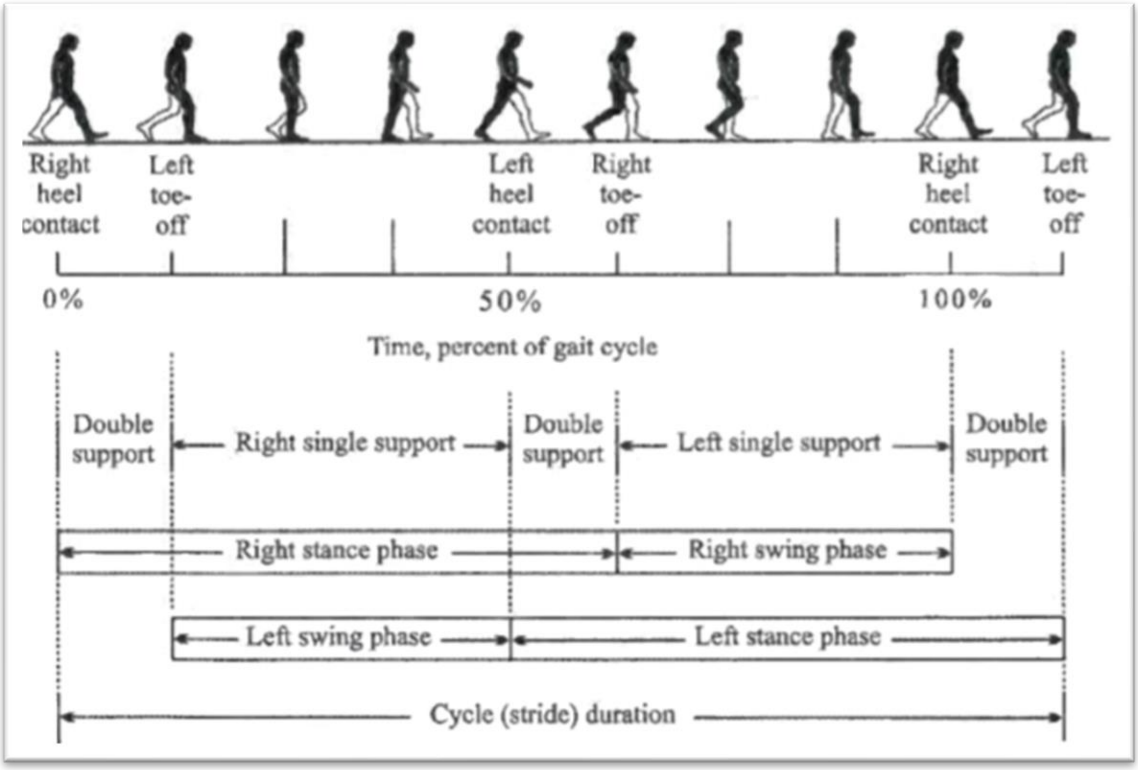

Figure 2, A single complete walk cycle (image courtesy of Racic et al.)

When you stop, both feet are in contact with the ground, so that the body remains in balance thanks to the double support ("double support"). When you start to walk instead, they highlight a number of consecutive events [2, 3, 4]:

1. In the first step the left foot rises from the ground ("toe-off left" in Fig. 2) initiating the movement, while the right foot acts as a single support body ("right single support" in Fig.2).
2. Subsequently, the body returns to a phase of dual standby because the left foot back again in contact with the ground ("left heel contact" in Fig. 2).
3. As a result, there is an opposite situation compared to that of the first phase, ie the right foot lifted from the ground ("right toe-off" in Fig. 2), while the left is still in the phase of contact so as to provide support individual claim ("left single support" in Fig.2).
4. There is then a further phase of dual standby when the right foot back again in contact with the floor ("right heel contact" in Fig. 2).
5. Finally, the cycle is complete when the left foot returns to rise from the ground ("toe-off left" in Fig. 2).

The foot then passes through two phases during each step:

1. The phase of oscillation ("swing phase"): with which reference is made to the time interval in which the foot is lifted from the ground.
2. The contact phase ("stance phase"): that is the period in which the foot is in contact with the plane, which starts at the moment when the heel touches the ground and ends with the complete support [5].

Even the body through two phases during the walking process:

1. double support phase("Double support" in Fig.2): accounts for one fifth of the walk cycle and occurs when both feet are in contact with the floor.
2. Single support phase: occurs when the left foot is in contact with the ground while the right is raised ("Left single support" in Fig.2) or vice versa ("Right heel contact" in Fig. 2) [6, 7].

To describe a person's walk is needed then use spatial and temporal parameters [8]. The spatial parameters of interest are:

- Stride length: distance between the heel of the right foot and that of the left foot during a step;
- Stride width: distance measured transversely between the two lines (respectively passing through the midpoints of the right and left heel), imaginary, that describe the trajectory followed by each foot respectively;
- Length of the walk cycle: Corresponds to the distance between the first and the second support of the same foot during a period of a [3] cycle.

While the time parameters are:

- walking speed: intensity of the horizontal speed in the gait direction;
- Cycle time: time window between two successive supports of the same foot to the floor [4];
- Step Frequency: number of steps at a time. [10]

2.1 THE FORCES INDUCED DURING THE WALK

In addition to the different parameters that characterize the man walk, we can ascertain the existence of two types of randomness: the variability inter and intra-subject variability [9].

1. inter-variable: it exists in relation to the fact that different people will have different key parameters related to the induced forces, the walking speed, etc.
2. intrasubjective Variable: It states that a person can not perform two consecutive identical steps. In practice, an individual generates different forces at every step.

The dynamic forces produced by humans during the walking process typically have components in directs vertical, horizontal-horizontal-parallel and transverse to the direction of movement. However, in this thesis we will consider only the vertical component because the others are not meaningful in the study of stability of pedestrian walkways. In addition, it has been shown that the vertical force has typically two peaks and a depression. As shown in Figure 2b, the first peak (F1) occurs at time T1, that is, when the heel touches the ground, as a result we are witnessing a depression (F2) when both heel that the tip of the foot are in contact with the ground . Finally, it has another peak (F3) to the time T3, when the heel is lifted from the ground,

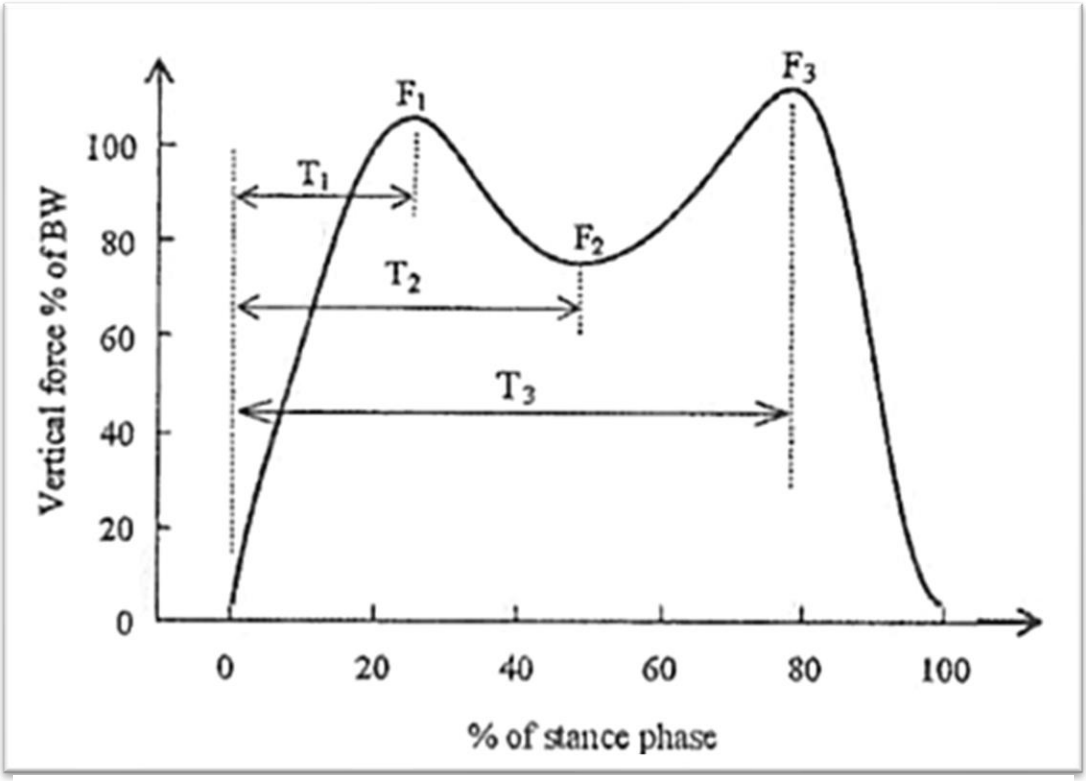

Figure 2b, the vertical force results during a single step (image courtesy of Racic et al.)

3 Data collection system

The data useful for performing this study were collected through a force floor (figure 3) and initially the variables taken into consideration were four, namely: the time between one step and the next, the force discharged to the ground during a step, the stride length during the walk and the angle of the step with respect to the direction parallel to the center line of the walkway. Thanks to a comprehensive study, performed by Martina Fornaciari, on the experimentally obtained data dependencies, three were considered relevant reports:

"No correlation between the length and angle of the step variables, negative correlation between the strength and the step time and negative correlation between time and stride length. In summary, time, stride length and strength they are linked to each other. Stride length and in fact take into account the time speed, and in the literature it is shown that the force depends on the speed. "[10]

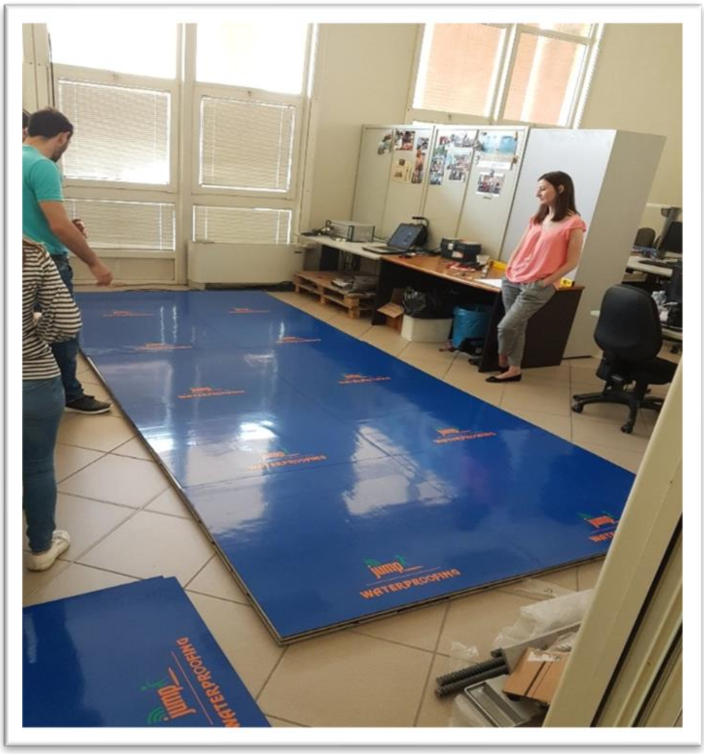

Figure 3, strengths floor used to make the measurements.

From these results it is clear that the mathematical model required to carry out this study will have to take in three random variables input ie the time between one step and the next, the force exerted at each step and the step length, for which you will have to do multivariate distributions.

In addition, it is important to emphasize that to obtain an effective correlation between the variables of interest would have to have a lot of data from so many people walking made one at a time, whereas, for reasons of time, we have only basic data relating to three available entities which differ in height and in proportion to the height, weight.

4 Data analysis

As mentioned several times, the aim of this study is to find a model that best describes the man walking process, and considering that the variables are more than you are forced to work with multivariate statistics .

4.1 RANDOM VARIABLES MULTIVARITE

The multivariate analysis field includes different statistical methods with which two or more random variables are considered related as a single entity and from which it is possible to extrapolate an overall result that takes account of existing relations between the [11] variables.

A random variable multivariate $\mathrm{X} = (\mathrm{X1}, \mathrm{X2}, \dots, \mathrm{Xn})$ It is therefore a function:

$$X: \omega \in \ \Omega \rightarrow X(\omega) \in \ R^n$$

For this reason, any point is chosen $(x_1, \dots, x_n) \in R^n$ it is possible to calculate the probability That the event:$P(X_1 \leq x_1, \dots, X_n \leq x_n)$

$$\{X_1 \leq x_1, \dots, X_n \leq x_n \} = \bigcap_{i=1}^{n} \{X_i \leq x_i\} \ \subseteq \ \Omega$$

From the conceptual point of view, the multivariate goes are not different from the univariate goes, but obviously, the mathematical treatment is more complex for the passage from R to R^n[12].

Before analyzing the various multivariate distributions studied is important to introduce some concepts and key definitions are useful for the study.

4.1.1 Media Campionaria

Given a particular sample and supposed not to know the true mean μ and the true variance σ^2, The primary goal is to try to estimate them reliably. Usually we try to estimateμ calculating the arithmetic average of the measured values:

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$$

To justify this, one can observe that $\bar{x}$ coincides perfectly with the value measured by the random variable defined as the arithmetic mean of the n random variables .For which may define a sample sample mean of the random variable:$X_i$$(\mathrm{X1}, \mathrm{X2}, \dots, \mathrm{Xn})\bar{X}$

$$\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_n$$

And it is considered an unbiased estimator of the true mean μ

The variance of the sample means it instead:

$$Var(\bar{X}) = \frac{1}{n}\, Var(X) = \frac{\sigma^2}{n}$$

It is none other than the true variance divided the n random variables.σ^2 [13]

4.1.2 Covariance sample

two samples and data$(\mathrm{X1}, \mathrm{X2}, \dots, \mathrm{Xn})(\mathrm{Y1}, \mathrm{Y2}, \dots, \mathrm{Yn})$ different but taken individually, it is often useful to understand if it exists between X and Y some form of linear dependence or correlation. For this reason it is necessary to estimate the covariance$\sigma_{X,Y}$ the joint variable (X, Y).

Defined sample covariance of the sample $((\mathrm{X}_i, Y_i), i = 1,2,..,n)$the random variable:

$$S_{X,Y} = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})(Y_i - \bar{Y})$$

In this way the true covariance is estimated:

$$\sigma_{X,Y} \approx \hat{\sigma}_{X,Y} = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y}) = \frac{1}{n-1}\left(\sum_{i=1}^{n} x_i y_i - \frac{1}{n}\left(\sum_{i=1}^{n} x_i\right)\left(\sum_{i=1}^{n} y_i\right)\right)$$

Where is it $\bar{x}$ and the arithmetic means of and. [13]$\bar{y} x_i y_i$

4.1.3 Joint distribution and marginal distributions

Having to deal with multivariate distributions, so it is often interested in studying probability problems related to the combined value of two or more random variables. To clarify this concept we define two random variables X and Y, it will follow that their joint distribution function is defined as follows:

$$F(a,b) = P\{X \leq a, Y \leq b\} \qquad -\infty < a, b < \infty$$

Once the joint distribution function can be traced to both the X distribution function:

$$F_x(a) = P\{X \leq a\}$$

$$= P\{X \leq a, Y < \infty\}$$

$$= P(\lim_{b\to\infty}\{X \leq a, Y \leq b\}$$

$$= \lim_{b\to\infty} \mathrm{P}\{X \leq a, Y \leq b\}$$

$$= \lim_{b\to\infty} \mathrm{F}(a, b)$$

$$\equiv F(a, \infty)$$

That the distribution function of Y:

$$F_y(b) = P\{Y \leq b\}$$

$$= \lim_{a\to\infty} \mathrm{F}(a, b)$$

$$\equiv F(\infty, b)$$

F_x e F_y are respectively the marginal distribution functions of X and Y. The latter will be very useful in our study, in fact, not knowing the trend of the joint distribution representing the walker, they are going to analyze the marginal, ie the vector of times between one step and the next, the vector of the forces discharged to the ground at every step and the vector of the lengths of each step while walking. [14]

4.1.4 Kolmogorv-Smirnov test

It is a nonparametric test that verifies the form of distributions, which compares the two data samples in the input and declares whether they come from the same continuous distribution. Obviously, the alternative hypothesis is that the two vectors derive from different continuous distributions. This test will be crucial in the course of studies carried out with the multivariate distributions, and is present in the toolbox of Matlab. Using the command "kstest2" and giving input two samples, the output result will be "1" if the test rejects the hypothesis, while appear "0" in the case of this latter being verified (standard significance level of 5%) . [15] [16] [17]

4.1.5 Copuling

Through the "Copuling" it is possible to obtain the joint construction of two or more random elements. It is essential for our study add concepts and definitions related to this theory.

Data (X_1, X_2, X_3) e $(\hat{X}_1, \hat{X}_2, \hat{X}_3)$ from two different families but which are copies in pairs (the variable does not necessarily take the same values but has the same distribution; for this is defined as a copy of), once obtained the two joint distributions, they are not necessarily distributed in same way. $\hat{X}_1 X_1 \hat{X}_1 X_1$

So, if you know all the marginal distributions of a model does not mean that we can derive the joint function uniquely. [18]

4.2 MULTIVARIATE NORMAL DISTRIBUTION

In statistics, the multivariate normal distribution is the generalization of the classic normal distribution to higher dimensions. It is useful, as in our case, when it has to do with more of a random variable.

Given the mean vector μ and the covariance matrix V, the joint density function of a multivariate normal distribution is:

$$f(x) = \frac{1}{\sqrt{(2\pi)^n|\mathbf{V}|}} \exp\left\{-\frac{1}{2}(x-\mu)\,\mathbf{V}^{-1}(\boldsymbol{x}-\boldsymbol{\mu})'\right\}$$

What you can do: $X \sim N_p(\mu, \mathbf{V})$

Although they note the mean vector and the covariance matrix, it has only a fraction of the information regarding the relationship between the variables. A multivariate normal distribution is in fact described by its joint distribution and its marginal distributions.$\mu\mathbf{V}$

Another important property is the following: given a vector of random variables, we can say that is distributed according to the multivariate normal if, for each, each linear combination is released under the ordinary. Therefore, X is distributed according to the multivariate normal if and only if every linear combination of the vectors has normal distribution. [19]$X = (X_1, X_2, \dots X_n) a \in \mathbb{R}^n X'_a = a_1 X_1 + a_2 X_2 + \cdots + a_n X_n X_i$

Since the goal is to find a model that best approximates the walkway process of man, it is essential to try to understand whether the multivariate normal distribution can be useful for this purpose.

Starting from the data collected through the strength of the floor, it will try to generate a random multivariate normal distribution function, which takes as input the mean vector and the covariance matrix obtained from the input sample. Once the generated Random distribution, expressed by the above definition, it will follow that each marginal generated will have normal distribution. For this reason, the latter, once obtained, will be compared with the marginal corresponding sample of the input, ie the vector representative of the time between one step and the next, that corresponding to the discharged force on the ground at every step and one representative of the length of each step. To make statistical comparison will be used to Kolmogorv-Smirnov test, so you will have to:

- If the test gives a positive result for all three pairs examined means that each of them has a normal distribution (level of significance of the test = 5%) and then you can deepen the study and assess whether the multivariate normal distribution approximates to best process of human walking.
- If the test result is negative even if only for a couple definitely means that the input sample is not approximated by a multivariate normal distribution.

These will be done through different tests Matlab to verify the above explained.

4.3 GAUSSIAN MIXTURE MODELS

A Gaussian mixture model (GMM) is a parametric probability density function represented as a weighted sum of the density of the Gaussian components.

The advantage of using a GMM is the ability to represent probability distributions in the presence of subpopulations, in fact points that the model produces is generated by a combination of n Gaussian distributions finished with unknown parameters. [20]

$$\boldsymbol{p(x|\lambda) = \sum_{i=1}^{M} w_i \, g(x|\mu_i, V_i)}$$

From the equation it is clear that a Gaussian mixture model is constituted by the weighted sum of the densities of each of the M components:

where is it:

- **x** is a vector of data in a continuous-dimensional value D
- $\boldsymbol{w_i}$, I = 1, ..., M are the proportional weights of the components of the mixture
- $\boldsymbol{g(x|\mu_i, V_i)}$, i = 1, ..., M are the Gaussian densities of the components of the mixture

Each of M component has a density of Gaussian function of the form:

$$\boldsymbol{g(x|\mu_i, V_i) = \frac{1}{(2\pi)^{D/2}|V_i|^{1/2}} \, exp\left\{-\frac{1}{2}(x-\mu)^T V_i^{-1}(x-\mu_i)\right\}}$$

where is it:

- $\boldsymbol{\mu_i}$ It is the mean vector
- $\boldsymbol{V_i}$ It is the vector of covariance
- The weight of the mixture satisfies the constraint that: $\sum_{i=1}^{M} \boldsymbol{w_i} = \mathbf{1}$

It can therefore be concluded that a complete model of the Gaussian mixtures is parameterized by the vector of the average values, from the covariance matrix and from the mixing weights of each component, which are collectively represented by the notation:

$\lambda = \{\boldsymbol{w_i},\ \boldsymbol{\mu_i}, \boldsymbol{V_i}\}$ i = 1, ..., M. [21]

It therefore will try to assess whether through a Gaussian mixture model it is possible to describe the man walking process.
Using the data obtained through the force floor, the first step will be to calculate the mean vector and the covariance matrix for each of the variables that will be used in input (variables: time between one step and the next, impressed force to each step, length between one step and the next) in order to subsequently generate the structure of Gaussian mixtures which describes the characteristics and showing the value of the respective mixing weights. Since this property is attempt to obtain, through the use of the toolbox of Matlab, an array of random variables distributed according to the structure of Gaussian mixtures. Will be carried out finally various statistical tests to determine whether the samples used in input,

4.4 COPULE

*the copule*They are functions that describe the dependencies between random variables, and provide a method to create distributions that process the correlated multivariate data. In this way, it is therefore possible to construct a multivariate distribution even when the marginal univariate not belong to the same family, but having a priori the degree of correlation between the variables. The copulas method is usable both with bivariate distributions that with higher dimensional distributions [22].

In this study we will use copulas to generate samples that simulate walking, Therefore the aim will be to produce the numerical backhoe that have a correlation structure according to measurable values and representing the variables of interest, ie the time between one step and the next, the pitch strength and the stride length, bypassing substantially the problem of knowledge of the joint distribution.

In other words we will use copulas as a generator of random samples of the three variables of interest, even without knowing the joint distribution.

To simulate dependent multivariate data through the use of copulas is necessary to specify the membership of the family, or the type of distribution (Gaussian, T Student etc ..), the degree of correlation between each variable and the respective marginal distributions.

It has the ability to run a parametric fitting of each of the input data taken separately and use their own estimates as marginal distributions. However, this method may not be as flexible, therefore it is convenient to use a non-parametric model, so as not to be forced to decide a priori the type of distribution. To have a significant discrete marginal distribution, the easiest method is to calculate the empirical cumulative distribution function (ecdf,) interpolating the trend between the midpoints of each step, so as to obtain a piecewise linear function. Finally, a method will be necessary to obtain the smoothing of the curve so as to be more significant.$F(x) = P(X \leq x)$

In the next chapter, in addition to the application of the method just explained, all the various tests carried out will be shown and illustrated through the Copule the purpose of generating of samples that simulate walking.

5 Statistical Tests

In this section we are going to analyze the results obtained by carrying out several tests through Matlab, an environment for the numerical calculation and statistical analysis. To carry out the tests have been used the data collected thanks to the strength of the floor.

Each test will consist of three phases:

1) **Data processing**

   It includes the loading phase of the experimental data in Matlab and the drafting of the code needed to carry out the "fitting" in accordance with the distribution that we intend to study.

2) **Statistical Test**

   Comparison between the experimental data and those obtained through the processing of the program. Using the "kstest2" command Matlab and giving input two samples is possible to evaluate, through the Kolmogorv-Smirnov test, if they belong or not to the same continuous distribution.

3) **Comparison charts**

   Analysis of the graphs that compare the trends of the data obtained experimentally and those generated in accordance with the distribution analyzed. Thanks to the function "difttool" Matlab is in fact possible to obtain the graphical representation of the vector input data in the form of histogram and then, using the "fit" function and specifying the type of distribution, a curve will be generated which approximates l ' trend. In case you do not know the distribution of the vector under consideration will suffice to set some parameters (bandwidth, domain) to have a non-parametric curve that approximates the better the performance of the carrier.

NOTE: During the tests it was found that the data collected from the subject of rather coarse A2 contain errors, especially those relating to the "TempoF" variable, so the results obtained from this sample will weigh differently than the others in the final conclusions. Will be marked with this symbol (*)

5.1 FITTING AND STATISTICAL TEST COMPARED TO A MULTIVARIATE NORMAL DISTRIBUTION

The following test has the purpose of verifying whether the man walking process can be described by a distributed model according to the multivariate normal.

*First, a sample is imported from Excel useful for testing. Subsequently, thanks to the function "**mvnrnd**" Matlab, which takes as input the mean vector "Mu" and the covariance matrix "sigma" of the test sample, is generated a matrix of random vectors chosen from the multivariate normal distribution. Finally, through theKolmogorov-Smirnov test vectors generated are compared with the corresponding marginal extracted from the input sample, to verify if they follow the same distribution.*

Below I developed the code commented:

```
x = xlsread ('Correlation Jessica11.xlsx');Data Matrix%
 % Imported from the first
 %sample

tempoF = x (:, 1); % Carrier that brings the measured time between
 % A step and the next

Force I = x (:, 2); % Vector which contains the Force be inscribed
 % Each step

LunghezzaP = x (:, 3);% Carrier that reports the length between
a
 And the next step%

meanF = mean (Force I); % Average carrier Forces
Meant = mean (tempoF); % Average carrier Times
meanL = mean (LunghezzaP);% Average vector lengths


mu = [Meant, meanF, meanL]; Vector% of average
Sigma = var (x); Covariance matrix%
```

```
R = mvnrnd (mu, Sigma, 415);% Returns a matrix of vectors
 % random chosen by the distribution
 % Multivariate normal with mean Mu and
 Covariance Sigma%


T = R (:, 1); Marginal% of R, representative TempoF

F = R (:, 2); Marginal% of R, representative Force I

L = R (:, 3);Marginal% of R, representative LunghezzaP

 % Kolmogorov-Smirnov test two samples

h1 = kstest2 (tempoF, T);
h2 = kstest2 (Force I, F);
h3 = kstest2 (LunghezzaP, L);


dfittool (T) % Fitting T
dfittool (F) % Fitting F
dfittool (L) % Fitting L
```

5.1.1 Statistical Test

The table below shows the results obtained by performing the Kolmogorv-Smirnov test between the experimental samples and those generated:

Test	Result (ls = 5%)
h1 = kstest2 (tempoF, T)	1
h2 = kstest2 (Force I, F)	1
h3 = kstest2 (LunghezzaP, L)	0

The h1 and h2 damage as a result test "1" this means that either reject the hypothesis that the two samples under test belong to the same continuous distribution

The h3 test instead, gives as a result "0" which means that the vector L, ie the marginal of the random matrix R generated, distributed according to the multivariate normal and the vector LunghezzaP, marginal of the input sample matrix, come from the same continuous distribution.

5.1.2 Comparison Charts

tempoF-T Chart

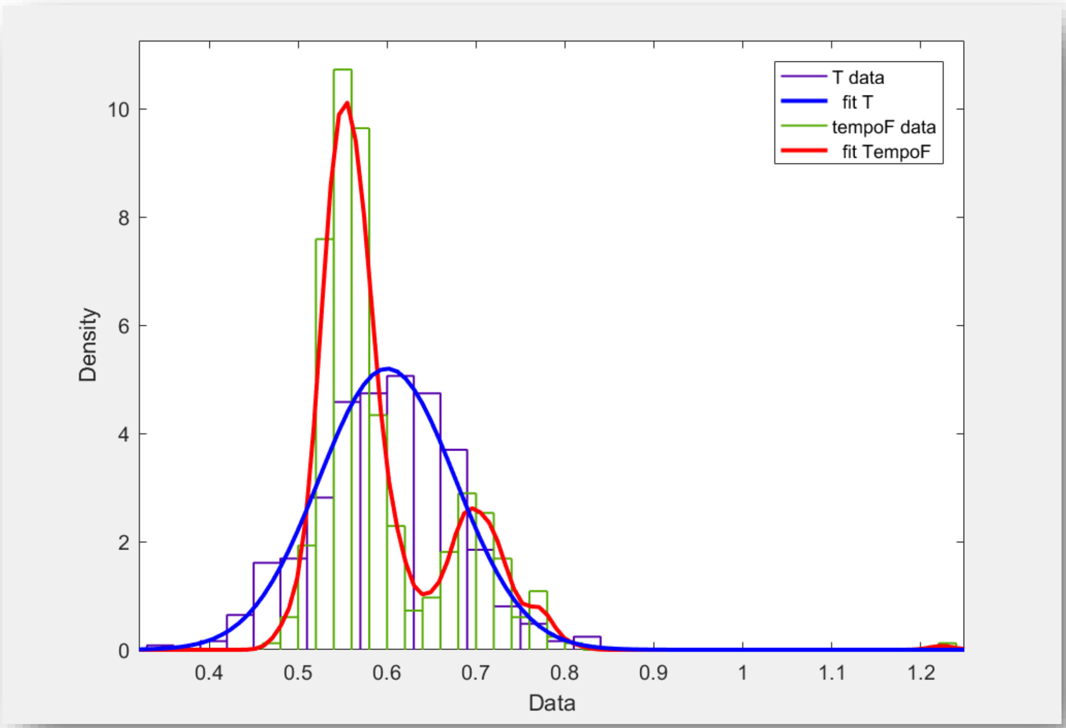

Figure 4, for histogram graph comparing the performance of tempoF and T

The graph in Figure 4, shows the trend of the input variable tempoF in green (curve which approximates the trend in red) and the generated variable T in purple (curve which approximates the trend in blue).

It can be stated that T approximates the normal, since it is a marginal amount of a multivariate normal distribution (ie R), while the same is not tempoF, which

shows a non-parametric trend and has two intervals where focus more values, the first between 0.5 and 0.6 sec, and the second between 0.65 and 0.75 sec.
The result obtained is in accordance with the Kolmogorv-Smirnov test, or that the two samples do not come from the same continuous distribution.

Chart Force I-F

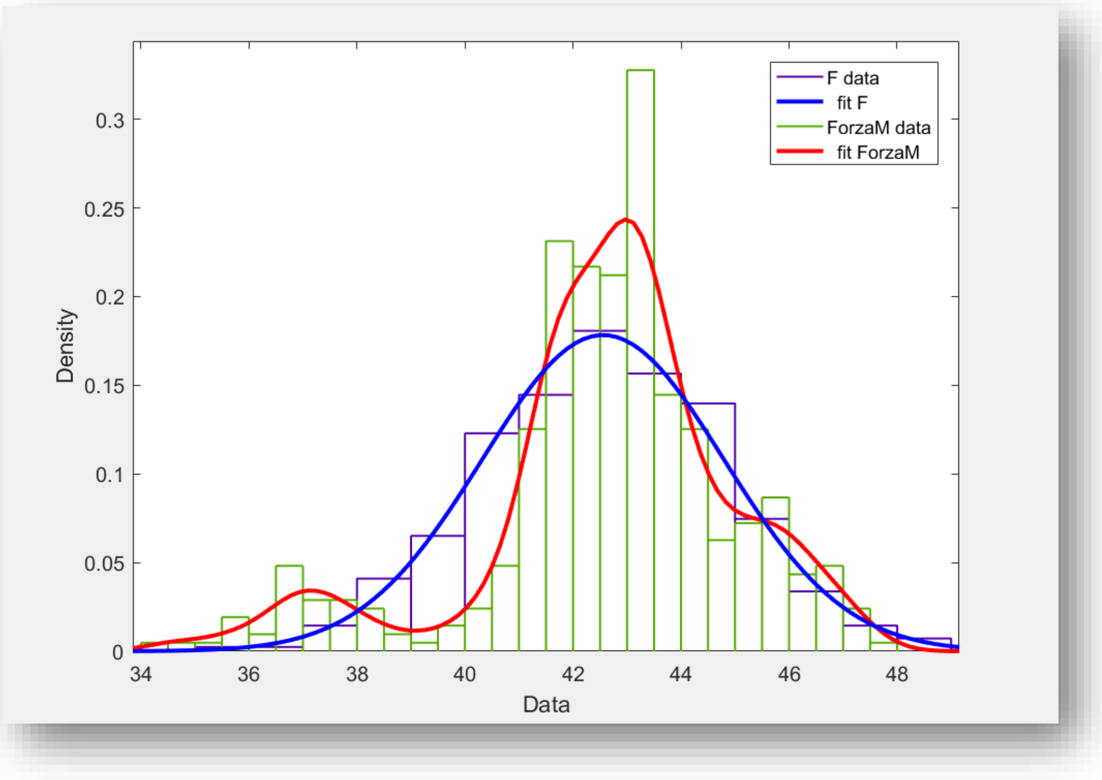

Figure 5, a histogram graph comparing the performance between Force I and F

The graph in Figure 5 shows the trend of the input variable Force I in green (curve which approximates the trend in red) and the carrier generated in purple F (curve which approximates the trend in blue).
Even in this case, as in the previous, F is distributed according to the normal, while Force I shows a non-parametric trend.
The result is in agreement with the Kolmogorov-Smirnov test, ie that the two input vectors do not belong to the same continuous distribution.

LunghezzaP Chart - L

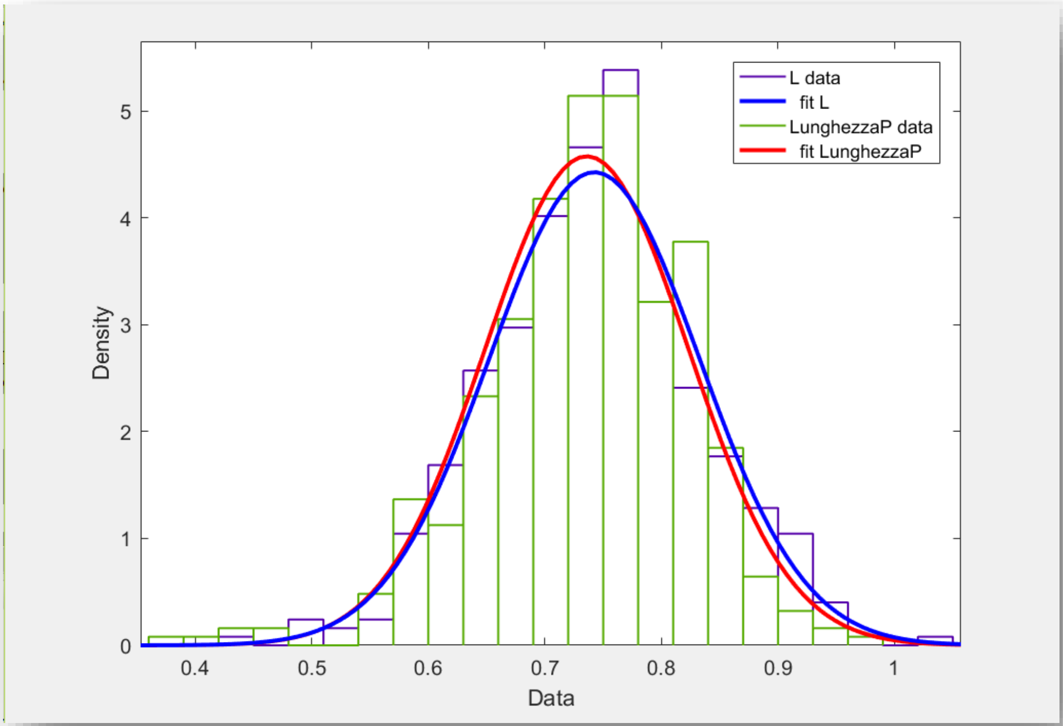

Figure 6, graph to Histogram comparing the performance between LunghezzaP and L

The figure shows the trend of the input variable LunghezzaP in green (curve which approximates the trend in red) and the carrier generated L in purple (curve which approximates the trend in blue).

In this graph, to differences of the previous cases, is known as the two curves are almost perfectly superimposed, this is once again in agreement with the result obtained by the Kolmogorov-Smirnov test, which stated that the two samples tested belonged to the same continuous distribution.

In addition, since it has been agreed that L is distributed according to the normal, marginal because of a multivariate normal distribution (ie R), it can be stated that even LunghezzaP is distributed according to the normal.

As a result, we will be reported the results obtained by the statistical and graphics tests, concerning the other samples collected through the strength of the floor.

5.1.3 Results obtained

Subject A2

Smirnov test Kolmogorv

Test	Result (ls = 5%)
h1 = kstest2 (tempoF, T)	1
h2 = kstest2 (Force I, F)	1
h3 = kstest2 (LunghezzaP, L)	1

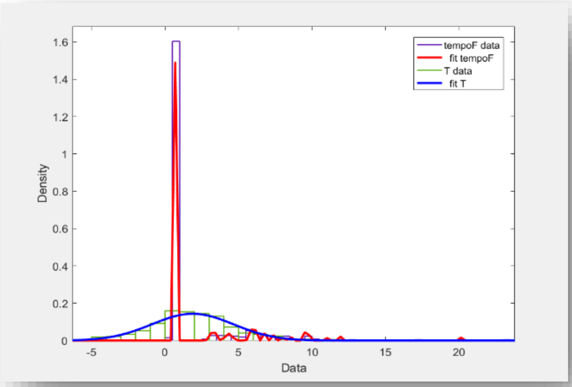

Figure 1, Histograms graph which compares the performance between tempoF and T (*)

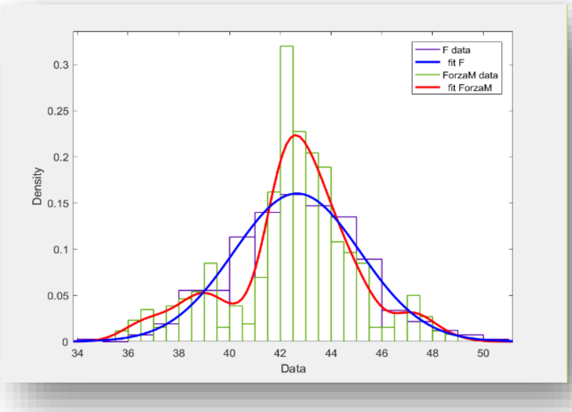

Figure 2, Histograms graph which compares the performance between Force I and F (*)

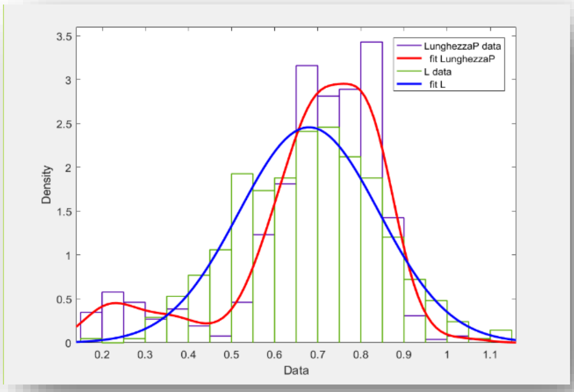

Figure 3, Histograms graph which compares the performance between LunghezzaP and L (*)

Subject B1

Smirnov test Kolmogorv

Test	Result (ls = 5%)
h1 = kstest2 (tempoF, T)	1
h2 = kstest2 (Force I, F)	1
h3 = kstest2 (LunghezzaP, L)	0

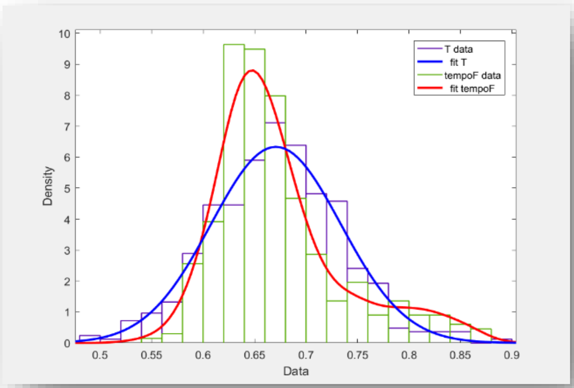

Figure 4, Histograms graph which compares the performance between tempoF and T

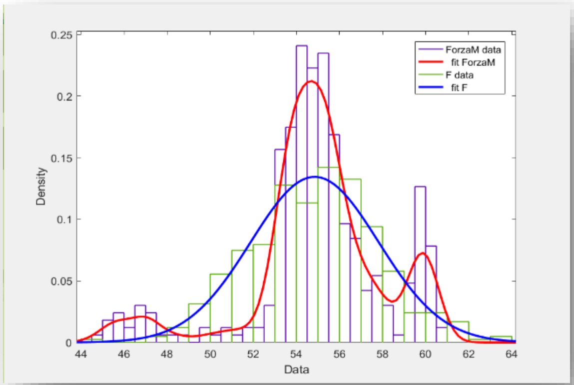

Figure 5, Histograms graph which compares the performance between Force I and F

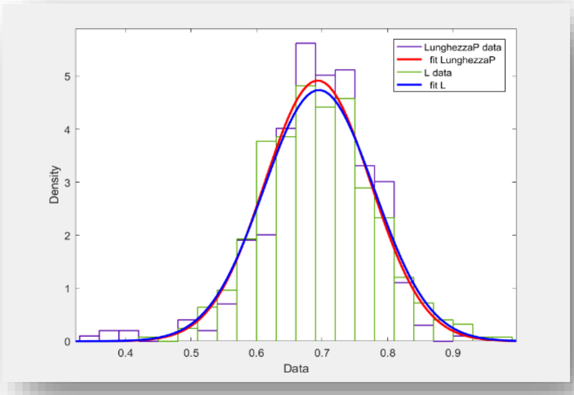

Figure 6, Histograms graph which compares the performance between LunghezzaP and L

Subject B2

Smirnov test Kolmogorv

Test	Result (ls = 5%)
h1 = kstest2 (tempoF, T)	1
h2 = kstest2 (Force I, F)	1
h3 = kstest2 (LunghezzaP, L)	0

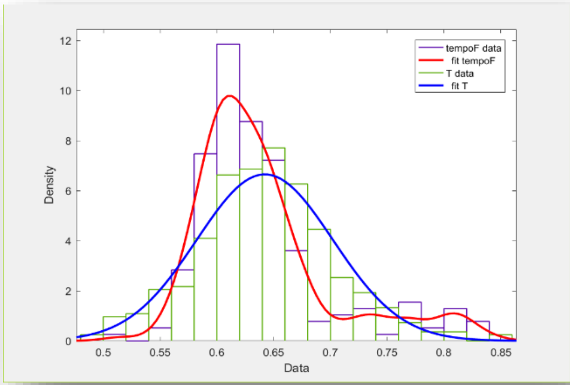

Figure 7, Histograms graph which compares the performance between tempoF and T

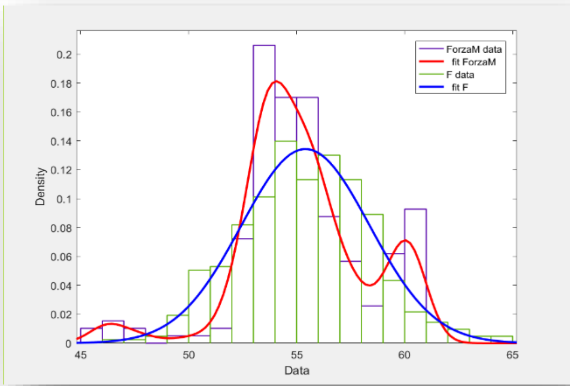

Figure 8, Histograms graph which compares the performance between Force I and F

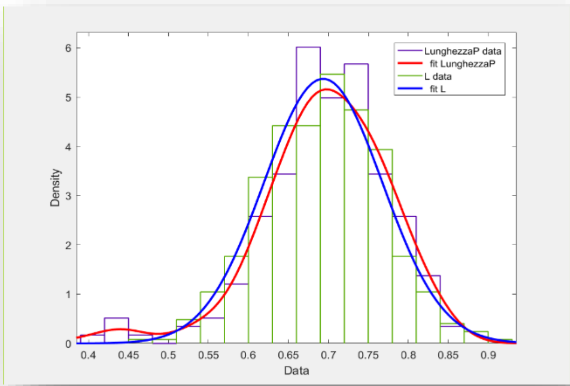

Figure 9, Histograms graph which compares the performance between LunghezzaP and L

Subject C1

Smirnov test Kolmogorv

Test	Result (ls = 5%)
h1 = kstest2 (tempoF, T)	1
h2 = kstest2 (Force I, F)	1
h3 = kstest2 (LunghezzaP, L)	0

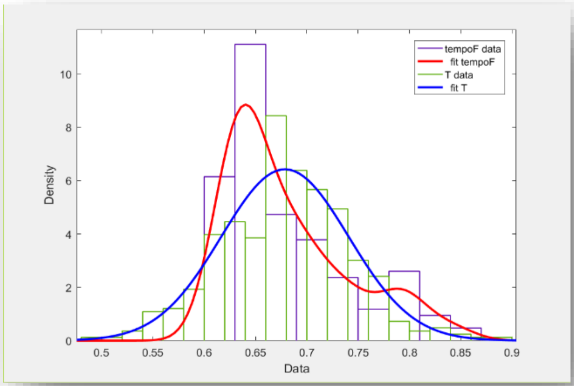

Figure 10, Histograms graph which compares the performance between tempoF and T

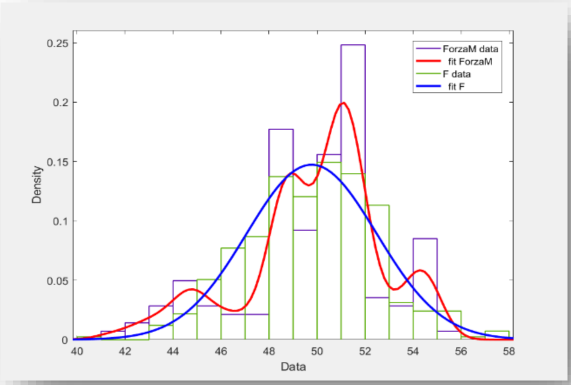

Figure 11, Histograms graph which compares the performance between Force I and F

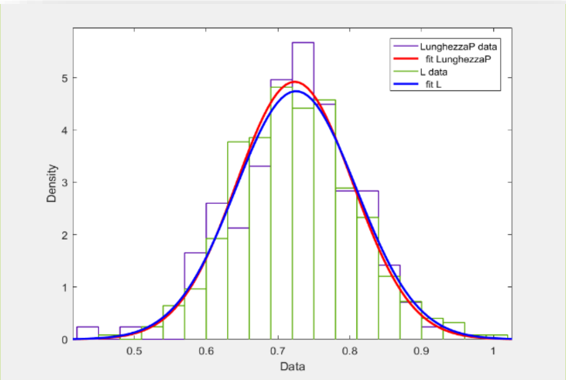

Figure 12, Histograms graph which compares the performance between LunghezzaP and L

Subject C2

Smirnov test Kolmogorv

Test	Result (ls = 5%)
h1 = kstest2 (tempoF, T)	1
h2 = kstest2 (Force I, F)	0
h3 = kstest2 (LunghezzaP, L)	0

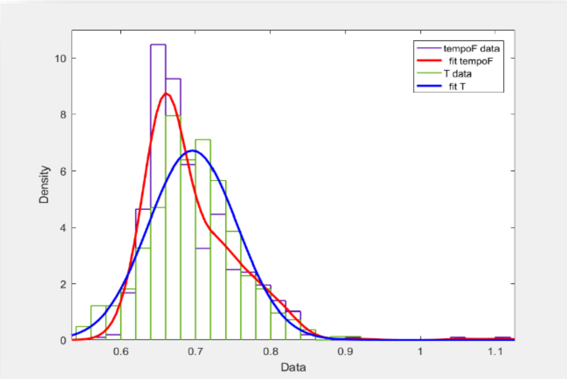

Figure 13, Histograms graph which compares the performance between tempoF and T

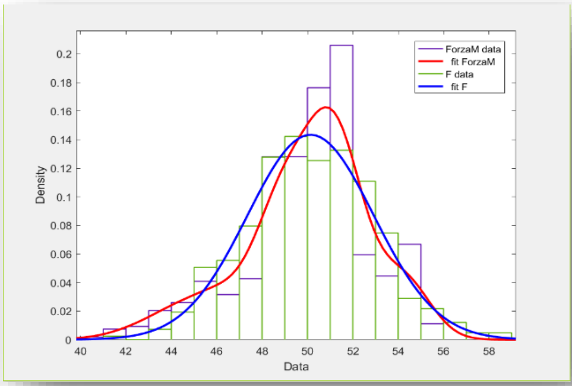

Figure 14, Histograms graph which compares the performance between Force I and F

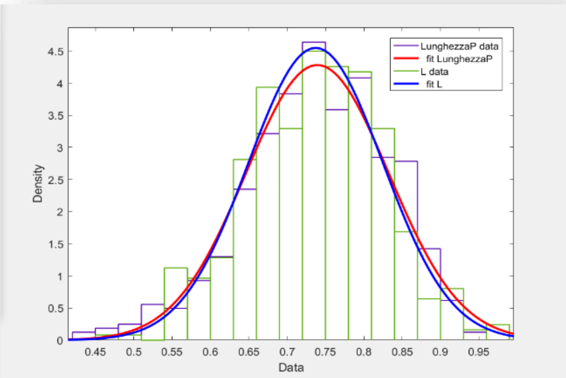

Figure 15, Histograms graph which compares the performance between LunghezzaP and L

5.2 FITTING AND STATISTICAL TEST OF A GAUSSIAN MIXTURE

The following test has the purpose of checking whether it is possible to describe the process of man walking through a Gaussian Mixture.

First you import from Excel a sample useful for testing. After imposed to 3 the number of degrees of freedom, using the "fitgmdist" function, which generates a structure that shows how it is made the mixture that best fits the data of the input sample. This structure will then be used through the "random" command to create n random variables distributed according to the structure of the generated Gaussian mixtures (in our case: n = 3). Finally, comparing vectors generated with the respective marginal of the input sample through the "kstest2" (of Kolmogorv-Smirnoff test).

Below is the code commented:

```
x = xlsread ('Correlation Jessica11.xlsx'); Data Matrix%
 % Imported the first
 %sample

TempoF = x (:, 1); % Carrier that brings the measured time
 % Between one step and the next

Force I = x (:, 2);% Carrier that shows the imprinted Strength
 % At each step

LunghezzaP = x (:, 3); % Vector which shows the distance between
 % A step and the next


meanF = mean (Force I); % Average carrier Forces
Meant = mean (TempoF); % Average carrier Times
meanL = mean (LunghezzaP); % Average vector lengths


mu = [Meant; meanF; meanL];Vector% of average
sigma = var (x); Covariance matrix%


df = 3;  Vinene set to 3% the number of
 % Of degrees of freedom
```

```
GMModel fitgmdist = (x, df) % This is the structure that tells how
 % Is made the mixture adapting to
 % Better than the data of the input sample


Y = random (GMModel, 415) % Return n random variables (n = 415).
 % Each column is a variable Y
 % Random extracted from the structure of
 % Gaussians mixtures generated GMModel.


T = Y (:, 1); Marginal% of Y corresponding to TempoF.
F = Y (:, 2);Marginal% of Y corresponding to Force I.
L = Y (:, 3); Marginal% of Y corresponding to LunghezzaP.


 % Plot the three marginal
dfittool (T)
dfittool (F)
dfittool (L)

 % Kolmogorov-Smirnov test two samples

h1 = kstest2 (TempoF, T)
h2 = kstest2 (Force I, F)
h3 = kstest2 (LunghezzaP, L)
```

5.2.1 Statistical Test

The table below shows the results obtained by performing the Kolmogorv-Smirnov test between the experimental samples and those generated:

Test Kolmogorov-Smirnoff	Result (ls = 5%)
h1 = kstest2 (tempoF, T)	0
h2 = kstest2 (Force I, F)	0
h3 = kstest2 (LunghezzaP, L)	0

As can be well noted in the table, in all three cases from the result of Kolmogorv-Smirnov test "0", this means that each of the marginal of the experimental matrix examined has the same distribution of marginal corresponding the Gaussian random mixtures matrix generated Y.

Generating more times the distribution Y and effecting the kstest2, it was found that not all test result occurred, but it can happen that the hypothesis that the two samples come from the same distribution is rejected, this is due to the process of generating random mix Gaussian Y.

NOTE: It has been noticed in fact that of about ten random generations of the mixture in structure, only one appears to have not all marginal distributed as the respective input variables. Also, please note that a statistical test with significance level of 5% is used.

5.2.1 Comparison Charts

TempoF-T Chart

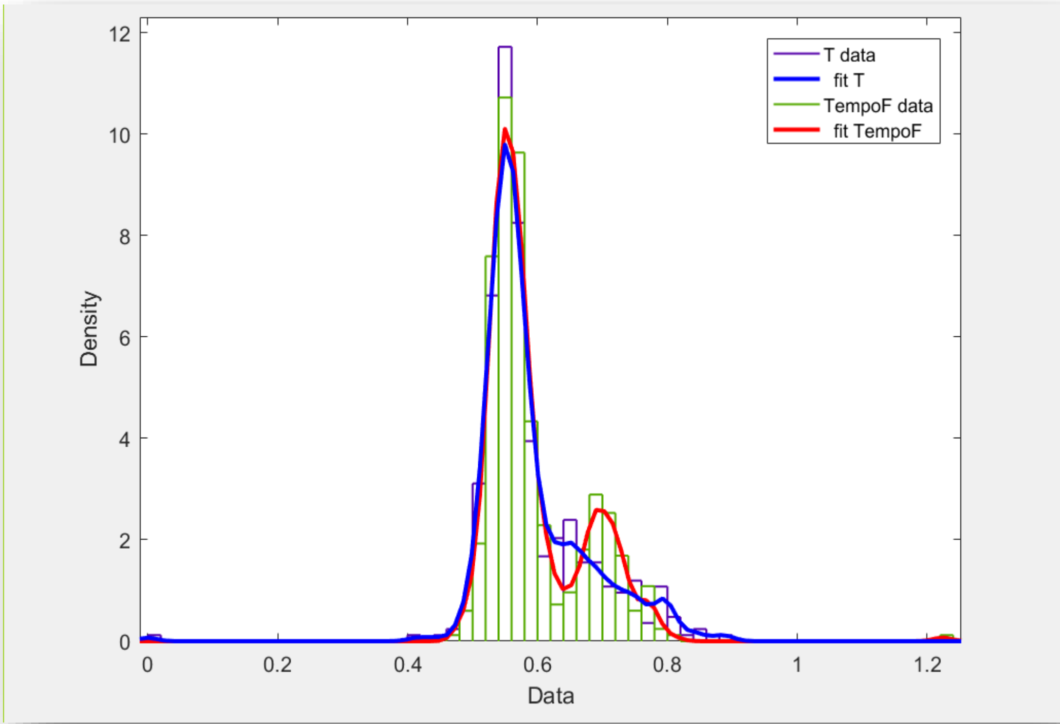

Figure 16, Histogram graph that compares the performance between TempoF and T

In the graph in Figure 24 shows the input TempoF in green variable (performance of red curve) and the carrier generated in purple T (trend in blue curve).

By analyzing the performance of the two curves can be seen which are almost perfectly superimposed in the range that goes from 0 to 0.6 seconds. Observing the two nell'intermezzo curves ranging from 0.6 to 0.9 seconds instead, we note a slightly different behavior, this may be due to non-accurate precision of the collected experimental data (TempoF curve), or to the approximation of the curve T, considering that has been randomly generated. If they are then neglected these small inaccuracies, one can conclude that the result obtained is in accordance with the Kolmogorv- Smirnov test, namely that the two test samples belong to the same distribution.

Chart Force I-F

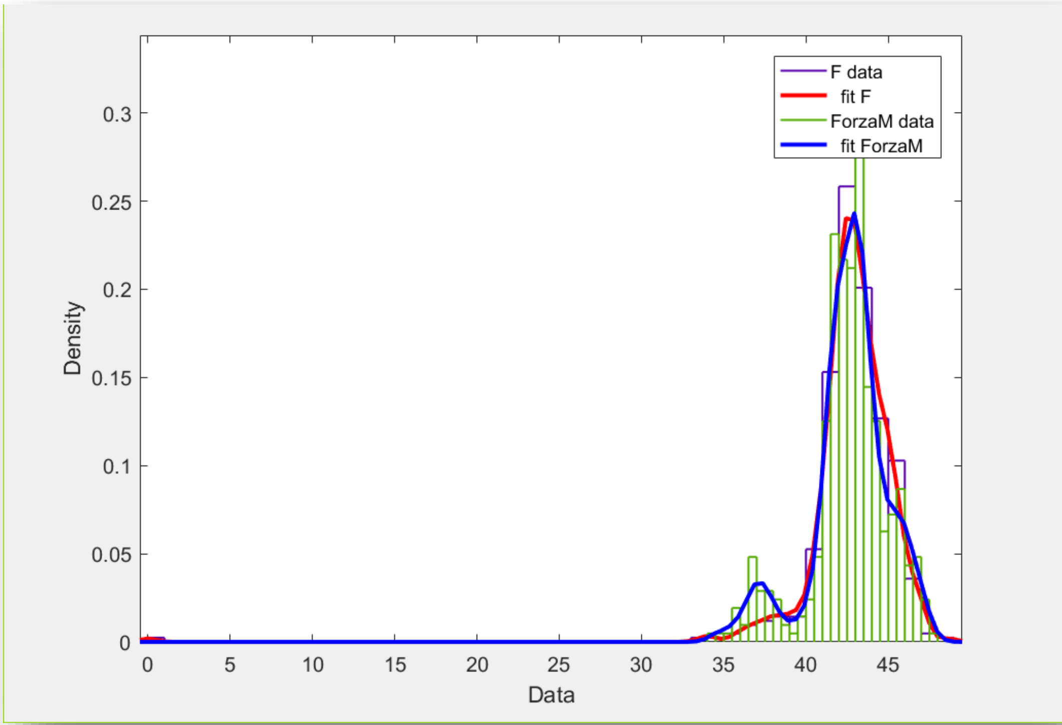

Figure 17Graph histogram that compares the performance between Force I and F

The graph in Figure 25 shows the trend of the vector F generated in purple (red curve) and the input variable Force I in green (blue curve).

Also in this case the two curves are almost totally superimposed in the range that goes from the go values comprised between 40 and 50 N, while it is noted that between 35 and 40 N Force I presents a slight peak unlike F.

Always considering a 5% significance level and assuming that these inconsistencies are caused by a not excellent collection of experimental data, or from an imperfection due to randomness in the generation of the Y matrix, one can confirm the result of the Kolmogorv-Smirnov test ie that the two variables come from the same continuous distribution.

LunghezzaP L-Graph

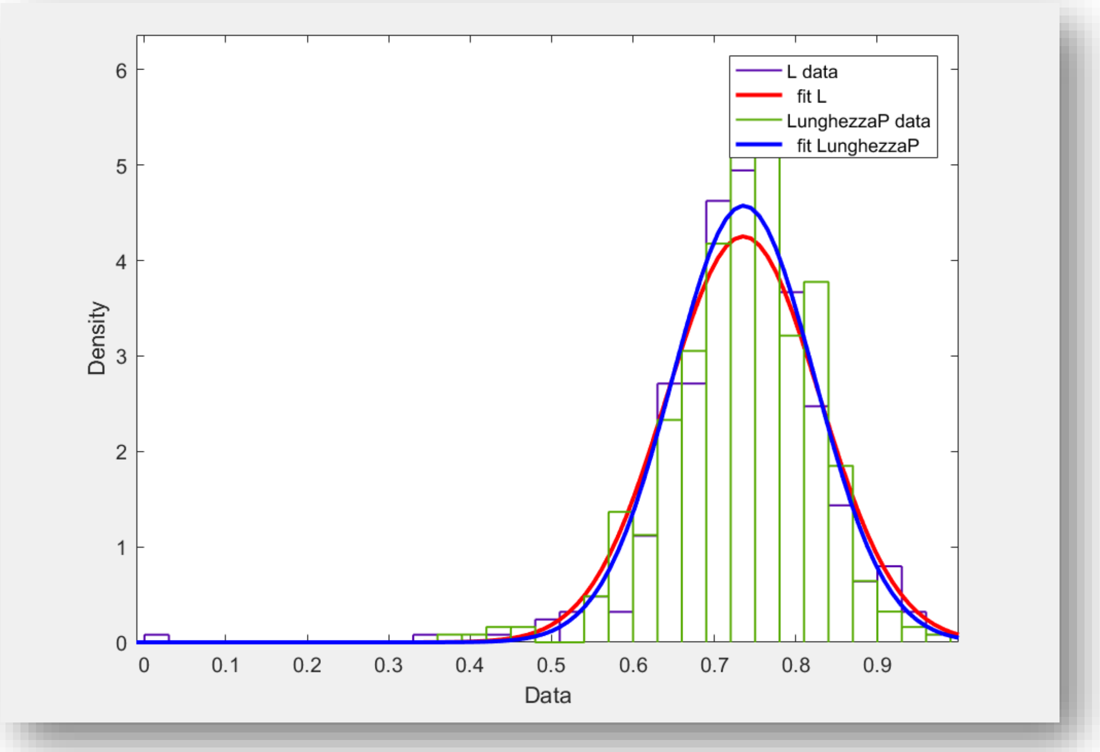

Figure 18, Histogram graph that compares the performance between LunghezzaP and L

This graph shows the trend of the vector L in purple (red curve) and the carrier LunghezzaP in green (blue curve).

In this last case we see how the curve L that LunghezzaP curve are almost perfectly superimposed and that both have a peak at about 0.75 m. This result is once again in agreement with that obtained from Kolmogrov-Smirnov test, namely that the two variables come from the same continuous distribution.

As a result, will be reported the results obtained by statistical tests and the respective graphs, the other samples tested.

5.2.3 Results obtained

Subject A2

Smirnov test Kolmogorv

Test	Result (ls = 5%)
h1 = kstest2 (tempoF, T)	1
h2 = kstest2 (Force I, F)	0
h3 = kstest2 (LunghezzaP, L)	0

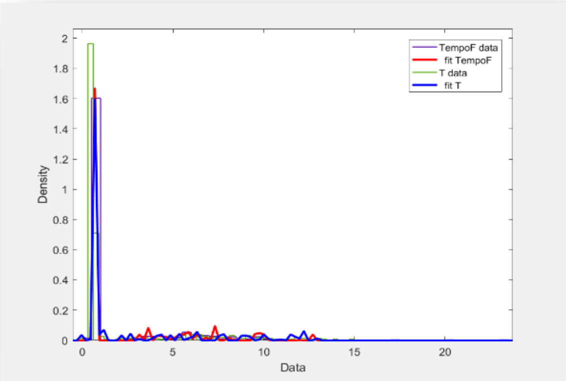

Figure 19, Histogram graph that compares the performance between TempoF and T (*)

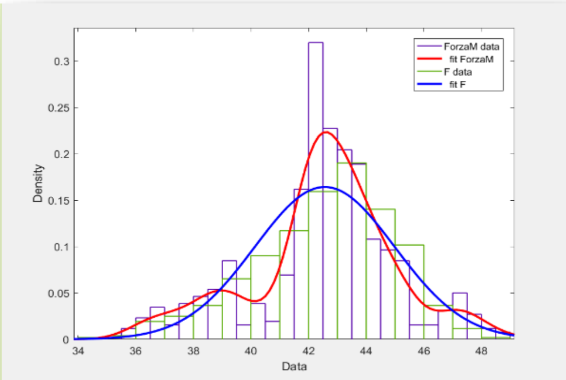

Figure 20Graph histogram that compares the performance between Force I and F

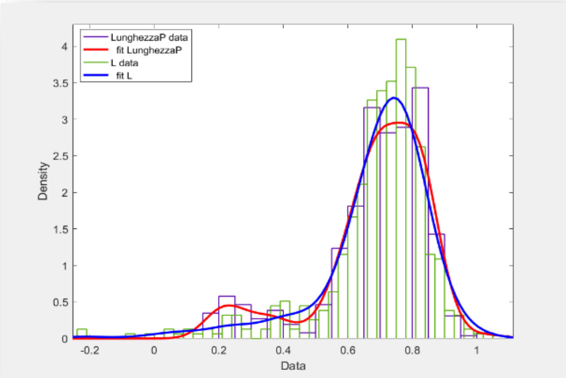

Figure 21, Histogram graph that compares the performance between LunghezzaP and L

Subject B1

Smirnov test Kolmogorv

Test	Result (ls = 5%)
h1 = kstest2 (tempoF, T)	0
h2 = kstest2 (Force I, F)	0
h3 = kstest2 (LunghezzaP, L)	0

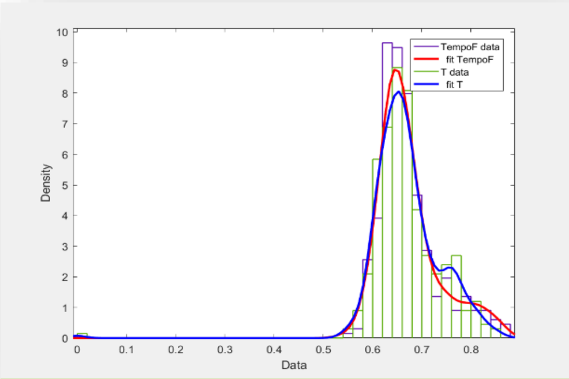

Figure 22, Histogram graph that compares the performance between TempoF and T

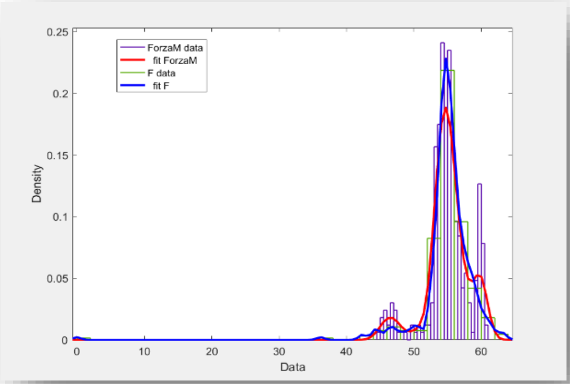

Figure 23Graph histogram that compares the performance between Force I and F

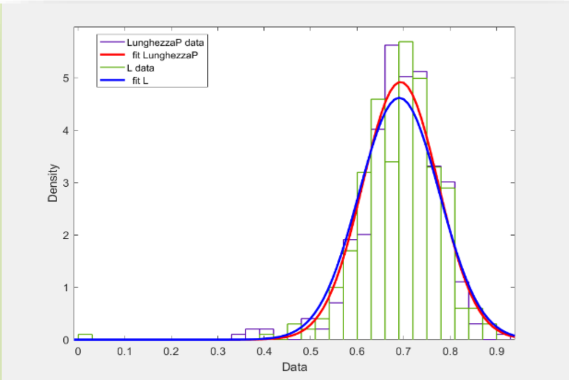

Figure 24, Histogram graph that compares the performance between LunghezzaP and L

Subject B2

Smirnov test Kolmogorv

Test	Result (ls = 5%)
h1 = kstest2 (tempoF, T)	0
h2 = kstest2 (Force I, F)	0
h3 = kstest2 (LunghezzaP, L)	0

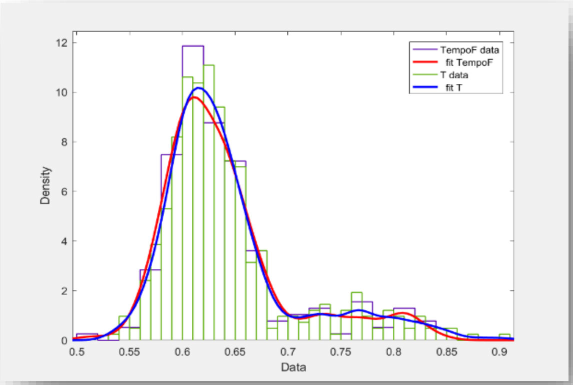

Figure 25, Histogram graph that compares the performance between TempoF and T

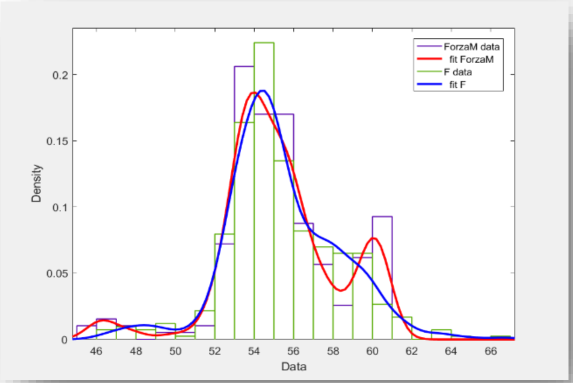

Figure 26Graph histogram that compares the performance between Force I-F

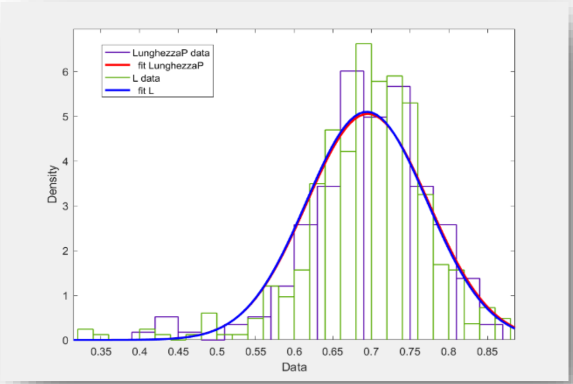

Figure 27, Histogram graph that compares the performance between LunghezzaP- L

Subject C1

Smirnov test Kolmogorv

Test	Result (ls = 5%)
h1 = kstest2 (tempoF, T)	0
h2 = kstest2 (Force I, F)	0
h3 = kstest2 (LunghezzaP, L)	0

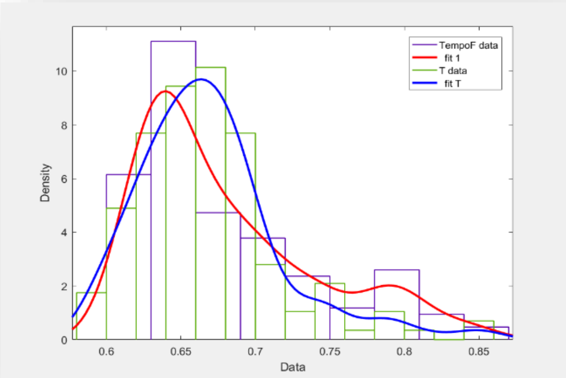

Figure 28, Histogram graph that compares the performance between TempoF and T

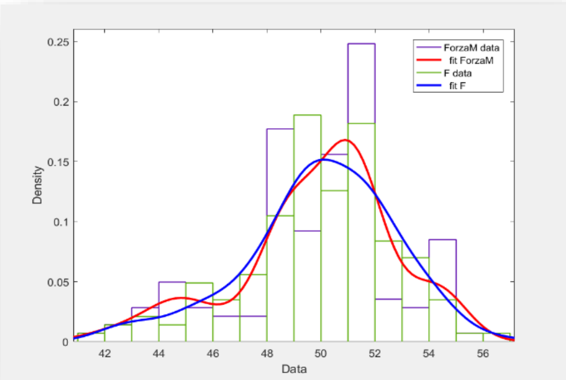

Figure 29Graph histogram that compares the performance between Force I-F

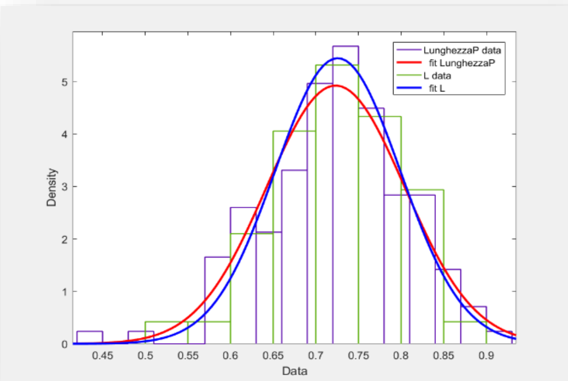

Figure 30, Histogram graph that compares the performance between LunghezzaP-L

Subject C2

Smirnov test Kolmogorv

Test	Result (ls = 5%)
h1 = kstest2 (tempoF, T)	0
h2 = kstest2 (Force I, F)	0
h3 = kstest2 (LunghezzaP, L)	0

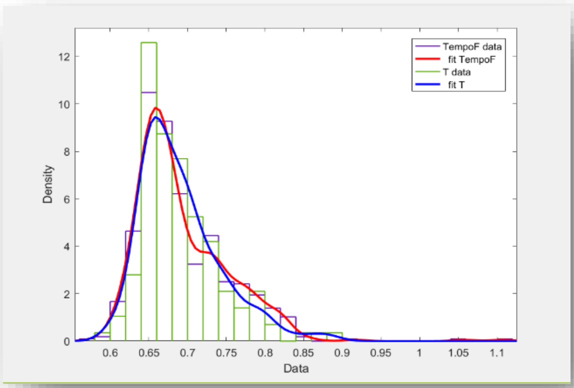

Figure 31, Histogram graph that compares the performance between tempoF and T

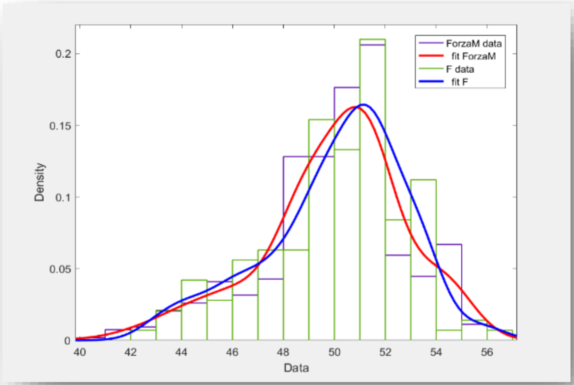

Figure 32Graph histogram that compares the performance between Force I-F

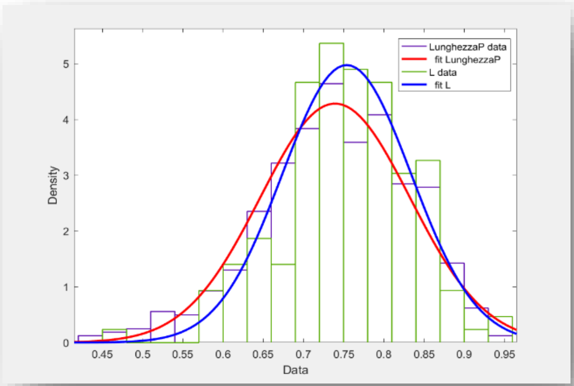

Figure 33, Histogram graph that compares the performance between LunghezzaP-L

5.3 COPULE STATISTICS

The goal is to verify if there is a method, through the copulas, to produce samples that simulate walking.

The test I will be divided into two parts, the first is to verify the method described in the previous chapter on the generation of discrete marginal distributions functions through a non-parametric approach.
While in the second part we will be precisely generated, through the copulas, of samples that simulate the behavior of a walker.
Matlab allows us to work to the maximum with bivariate, whereby the number of degrees will be given as input a pair of marginal obtained from an experimental sample (eg: TempoF, Force I) from which it will calculate the correlation matrix and after having chosen to freedom, through the "copularnd" command will be possible to extract the correlation matrix of the copula. Finally, thanks to the "ksdensity" functionit will be possible to make an estimate of the distribution of the kernel and to obtain the smoothed curve which interpolates the most of the performance from the points found in accordance with the correlation matrix of the copula in one step.
Next you will need to perform a statistical test to verify whether the two generated variables and those input dates are distributed in the same way, and then if the data obtained can be used to simulate a walker at best.

5.3.1 Part 1

Below is the code developed commented:

```
x = xlsread ('Correlation Jessica11.xlsx'); % Amount input data

TempoF = x (:, 1); % Carrier that brings the measured time between one
step and
 % the next one

Force I = x (:, 2); % Carrier that brings the force exerted at each
step

LunghezzaP = x (:, 3); % Vector which shows the distance between one
step and
 % the next one

meanF = mean (Force I); % Average carrier Forces
Meant = mean (TempoF); % Average carrier Times
meanL = mean (LunghezzaP); % Average vector lengths



mu = [Meant; meanF; meanL]; Vector% of average
sigma = var (x); Covariance matrix%

[Fi, xi] = ecdf (TempoF); % Returns the empirical function
 Fi% cumulative distribution evaluated in
 % Points xi, using the data carrier
 TempoF%.

% I get the graph of empirical cumulative distribution function

figures ()
stairs (xi, Fi,'B','LineWidth', 1)
hold on

% I get the smoothed curve of the graph through the ksdensity
function.

Fi_sm = ksdensity (TempoF, xi,'Function','Cdf','Bandwidth', 0.01);

% Comparison of the two curves obtained

plot (xi, Fi_sm,'R-','LineWidth', 1)

% Graph Legend

xlabel ('X1')
ylabel ('Cumulative Probability')
legend ('Empirical','Smoothed','Location','NW')
grid on
```

```
[Fi, xi] = ecdf (Force I); % Returns the empirical function
 Fi% cumulative distribution evaluated in
 % Points xi, using the data carrier
 Force I%.


% I get the graph of empirical cumulative distribution function
```

```
figures ()
stairs (xi, Fi,'B','LineWidth', 1)
hold on

% I get the smoothed curve of the graph through the ksdensity
function.

Fi_sm = ksdensity (Force I, xi,'Function','Cdf','Bandwidth', 0:15);

% Comparison of the two curves obtained

plot (xi, Fi_sm,'R-','LineWidth', 1)

% Graph Legend

xlabel ('X2')
ylabel ('Cumulative Probability')
legend ('Empirical','Smoothed','Location','NW')
grid on
```

```
[Fi, xi] = ecdf (LunghezzaP); % Returns the empirical function
 Fi% cumulative distribution evaluated in
 % Points xi, using the data carrier
 LunghezzaP%.


% I get the graph of empirical cumulative distribution function

figures ()
stairs (xi, Fi,'B','LineWidth', 1)
hold on

% I get the smoothed curve of the graph through the ksdensity
function.

Fi_sm = ksdensity (LunghezzaP, xi,'Function','Cdf','Bandwidth', 0.03);

% Comparison of the two curves obtained

plot (xi, Fi_sm,'R-','LineWidth', 1)

% Graph Legend

xlabel ('X3')
ylabel ('Cumulative Probability')
legend ('Empirical','Smoothed','Location','NW')
grid on
```

5.3.1.1 Comparison Charts

Chart Empirical-Smoothed discrete TempoF marginal distribution

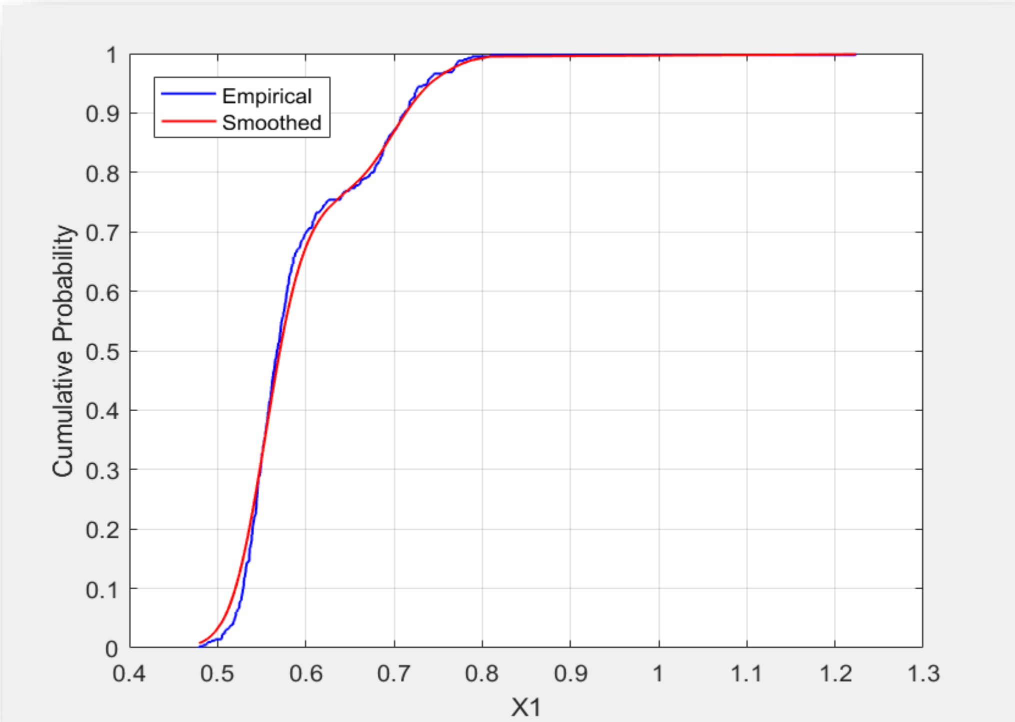

Figure 34, Shows the trend of the empirical marginal distribution (TempoF) and the respective "smoothing"

The graph in Figure 42 shows in blue the trend of empirical marginal distribution function of TempoF (ecdf), together with the "smoothed" curve obtained thanks to the "ksdensity function" in red.

It is noted that the blue curve is a piecewise linear function, or a broken obtained through the interpolation of points of the marginal TempoF input, while the red curve approximates the trend so as to obtain a continuous distribution.

Chart Empirical-Smoothed the discrete marginal distribution Force I

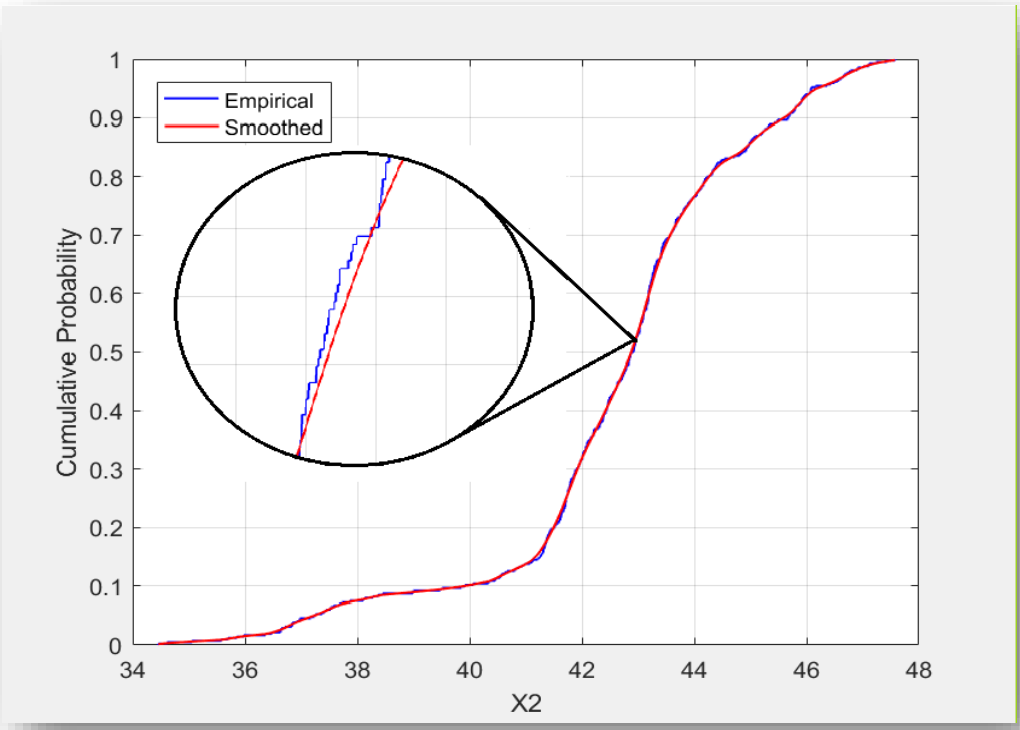

Figure 35, Shows the trend of the empirical marginal distribution (Force I) and the respective "smoothing"

The graph in Figure 43 shows in blue the trend of empirical marginal distribution function of Force I (ecdf), together with the "smoothed" curve obtained thanks to the "ksdensity function" in red.

In the illumination is visible the difference between the two patterns, namely the piecewise linear function obtained through the interpolation of points of the marginal grip Force I input (in blue) and the respective "smoothing" which outlines a continuous distribution (in red).

Chart Empirical-Smoothed discrete LunghezzaP marginal distribution

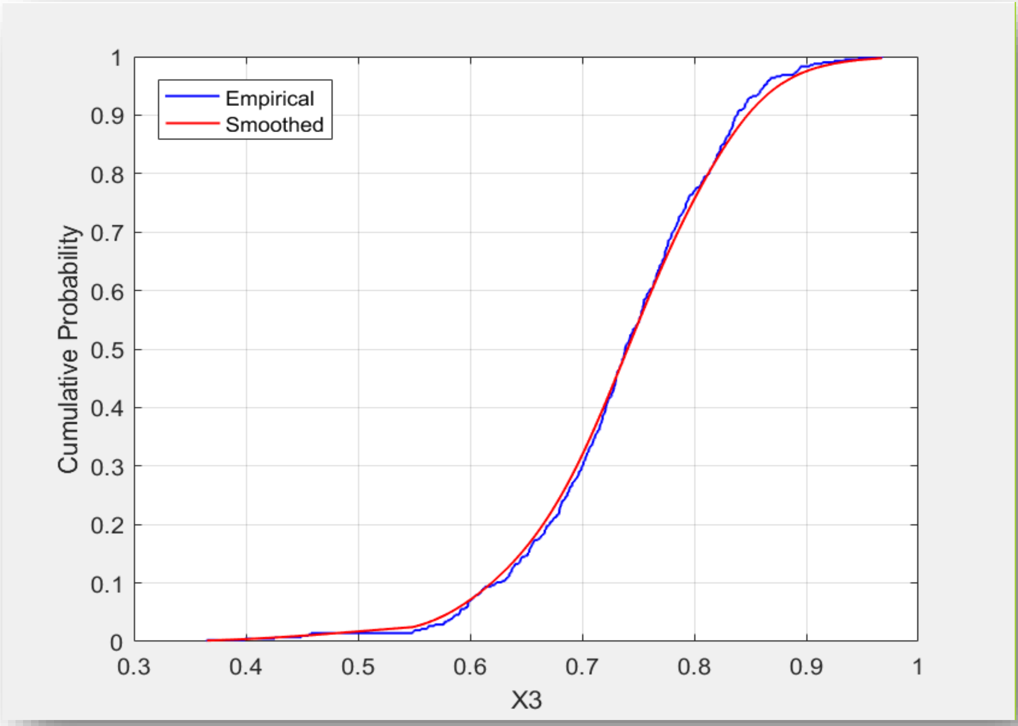

Figure 36, Shows the trend of the empirical marginal distribution (LunghezzaP) and the respective "smoothing

The graph in Figure 14 shows in blue the trend of the marginal distribution function of empirical LunghezzaP (ecdf), or the piecewise linear function obtained through the interpolation of points of the marginal grip Force I input, together with the "smoothed" curve obtained thanks to "ksdensity function" in red.

Using this method has therefore been possible to obtain the three marginal X1, X2 and X3 through a non-parametric approach

5.3.1.2 Results obtained

Subject A2

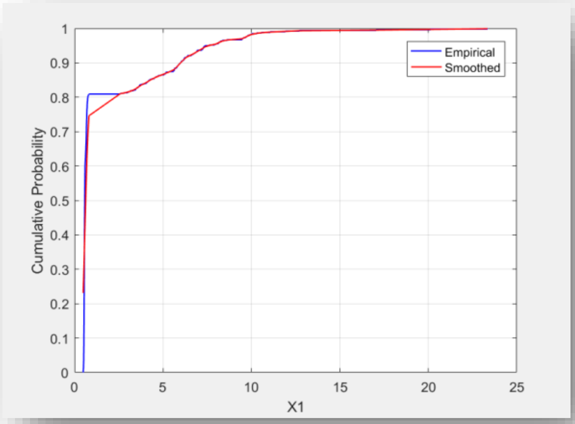

Figure 37, TempoF-marginal distribution obtained through a non-parametric approach (in blue) and "smoothing" (in red) (*)

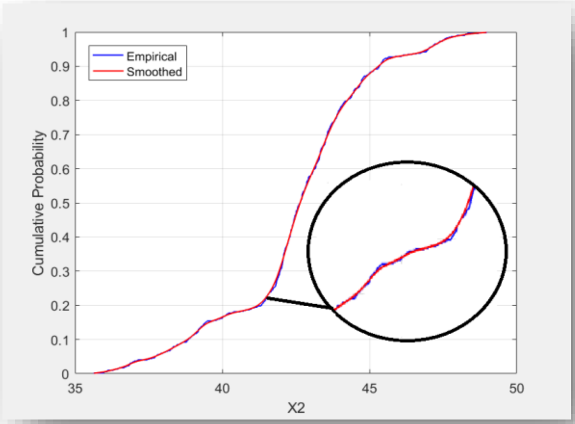

Figure 38, Force I-marginal distribution obtained through a non-parametric approach (in blue) and "smoothing" (in red)

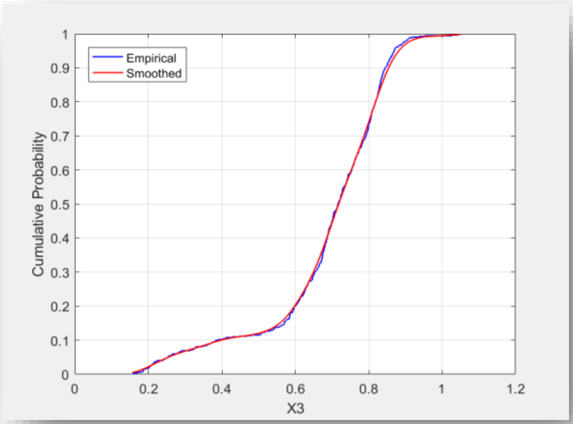

Figure 39, LunghezzaP-marginal distribution obtained through a non-parametric approach (in blue) and "smoothing" (in red)

Subject B1

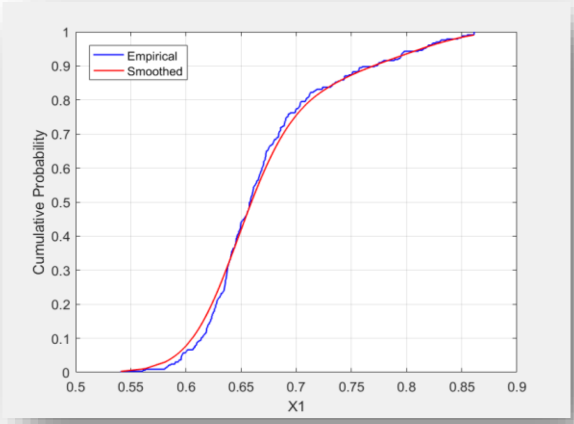

Figure 40, TempoF-marginal distribution obtained through a non-parametric approach (in blue) and "smoothing" (in red)

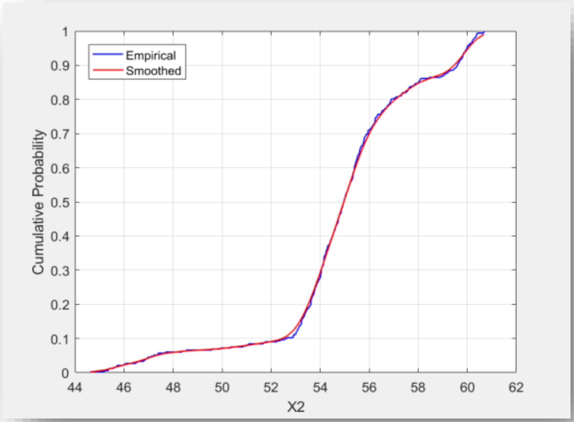

Figure 42, Force I-marginal distribution obtained through a non-parametric approach (in blue) and "smoothing" (in red)

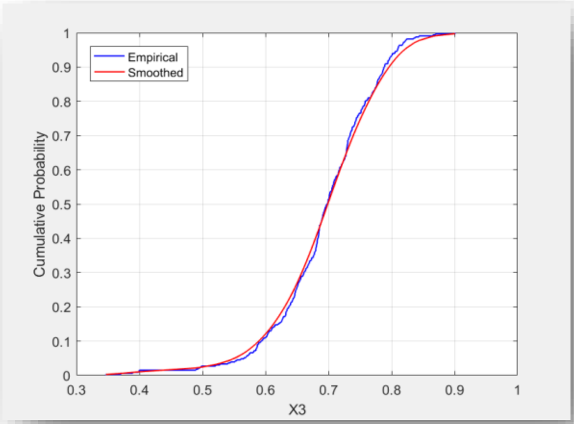

Figure 43, LunghezzaP-marginal distribution obtained through a non-parametric approach (in blue) and "smoothing" (in red)

Subject B2

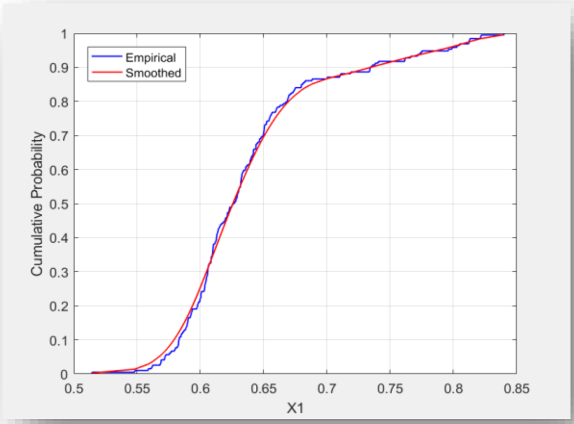

Figure 44, TempoF-marginal distribution obtained through a non-parametric approach (in blue) and "smoothing" (in red)

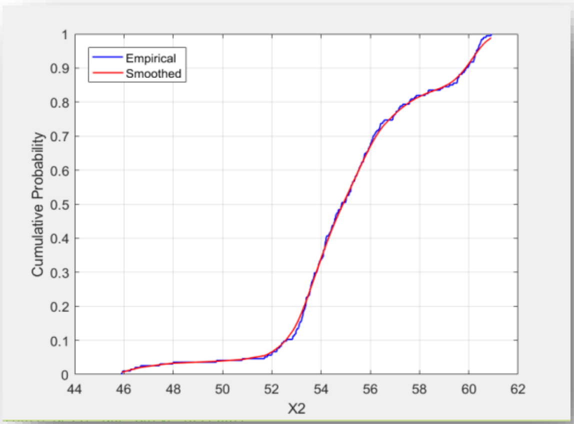

Figure 45, Force I-marginal distribution obtained through a non-parametric approach (in blue) and "smoothing" (in red)

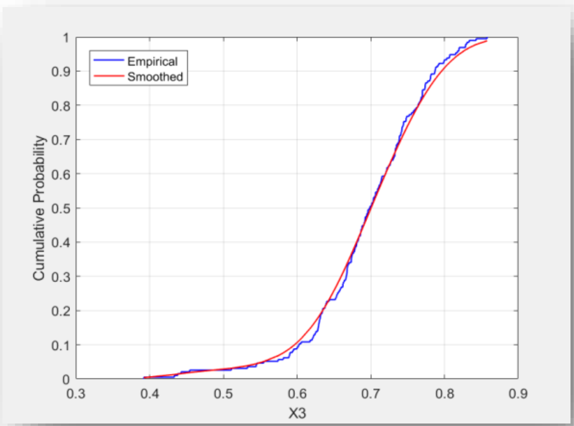

Figure 46, LunghezzaP-marginal distribution obtained through a non-parametric approach (in blue) and "smoothing" (in red)

Subject C1

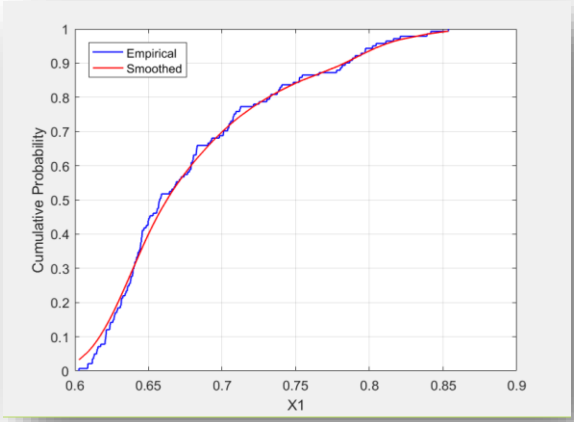

Figure 47, TempoF-marginal distribution obtained through a non-parametric approach (in blue) and "smoothing" (in red)

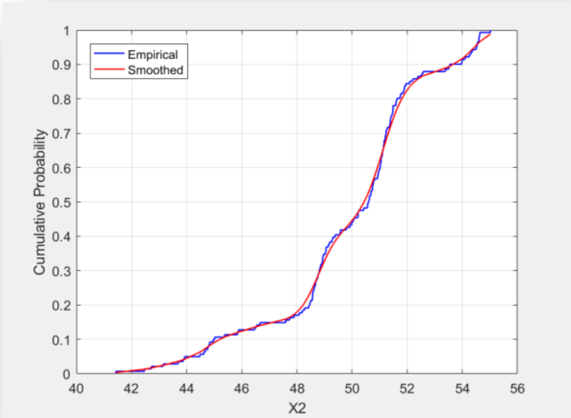

Figure 48, Force I-marginal distribution obtained through a non-parametric approach (in blue) and the "smoothing" (in red)

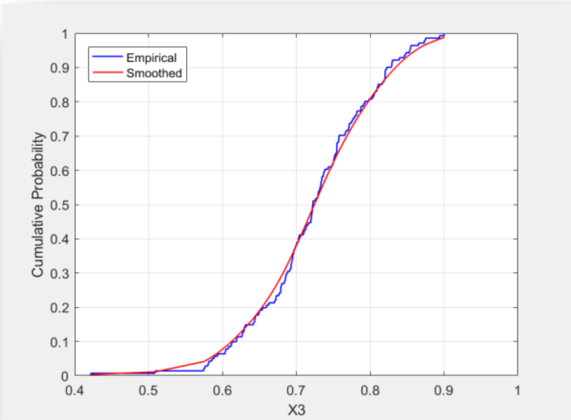

Figure 49, LunghezzaP-marginal distribution obtained through a non-parametric approach (in blue) and the "smoothing" (in red)

Subject C2

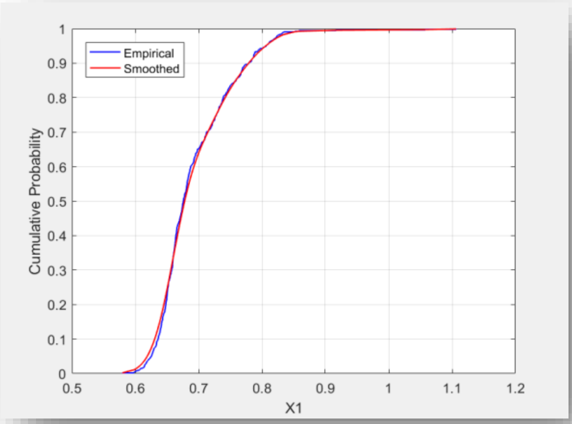

Figure 50, TempoF-marginal distribution obtained through a non-parametric approach (in blue) and "smoothing" (in red)

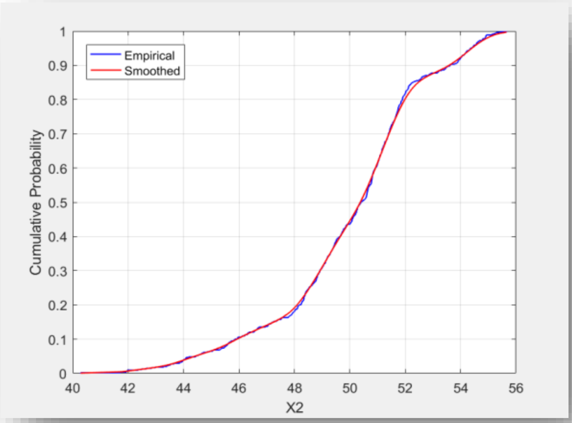

Figure 51, Force I-marginal distribution obtained through a non-parametric approach (in blue) and the "smoothing" (in red)

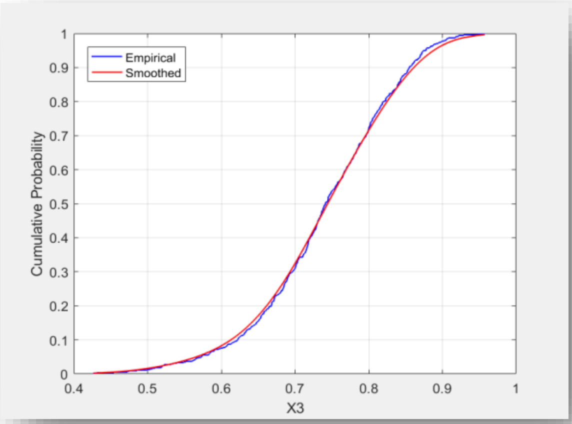

Figure 52, LunghezzaP-marginal distribution obtained through a non-parametric approach (in blue) and "smoothing" (in red)

5.3.2 Part 2

Later, it shows the code developed commented:

```
x = xlsread ('Correlation Jessica11.xlsx'); % Amount input data

TempoF = x (:, 1);% Carrier that brings the measured time between one
step and
 % the next one

Force I = x (:, 2);% Carrier that brings the force exerted at each step

LunghezzaP = x (:, 3); % Vector which shows the distance between one
step and
 % the next one

meanF = mean (Force I); % Average carrier Forces
Meant = mean (TempoF); % Average carrier Times
meanL = mean (LunghezzaP); % Average vector lengths


mu = [Meant; meanF; meanL]; Vector% of average
sigma = var (x); Covariance matrix%
```

```
nu = 3;% Number of degrees of freedom

tau = corr (x (:, 1), x (:, 2),'Type','Kendall');

% Returns the matrix of correlation coefficients between the two
marginal TIMEM% and Force I in the x input matrix.

rho = copulaparam ('T', Tau, nu, 'Type','Kendall');% Returns

% Of linear correlation parameters corresponding to the copula t.

n = 415;% Number of samples

U = copularnd ('T'[Rho 1; rho 1], nu, n);% N vectors generate random
from
 % T copula with parameters
 % Rho linear correlation.

Ksdensity X1 = (x (:, 1), U (:, 1),'Function','ICDF','Width', 0.02);
X2 = ksdensity (x (:, 2), U (:, 2),'Function','ICDF','Width', 0:15);

% The ksdensity function allows you to make an estimate of the marginal
distributions (smoothing of the trend) of the copula t obtained.


scatterhist (X1, X2,'Direction','Out')% Instogrammi the son of X1 and
X2


h1 = kstest2 (X1, TempoF)% Test Kolmogorov-Smirnoff between X1 and
TempoF
h2 = kstest2 (X2, Force I)% Test Kolmogorov-Smirnoff between X2 and
Force I

dfittool (X1)
dfittool (TempoF)% Plot of X1 and TempoF
```

```
dfittool (X2)
dfittool (Force I)% Plot of X2 and Force I
```

```
nu = 3; % Degrees of Freedom

tau = corr (x (:, 2), x (:, 3),'Type','Kendall'); %% Returns the matrix
of correlation coefficient between the two marginal Force I and
LunghezzaP of the x input data matrix.

rho = copulaparam ('T', Tau, nu, 'Type','Kendall');% Returns the linear
correlation parameters corresponding to the copula t.


% N generate random vectors from the t copula with linear correlation
parameters rho.

n = 415;
U = copularnd ('T'[Rho 1; rho 1], nu, n);% N vectors generate random
from
 % T copula with parameters
 % Rho linear correlation.

X3 = ksdensity (x (:, 2), U (:, 1),'Function','ICDF','Width', 0:15);
X4 = ksdensity (x (:, 3), U (:, 2),'Function','ICDF','Width', 0.05);

% The ksdensity function allows you to make an estimate of the marginal
distributions (smoothing of the trend) of the copula t obtained.


scatterhist (X3, X4,'Direction','Out')% Instogrammi the son of X3 and
X4

dfittool (X3)
dfittool (Force I)% Plot of X3 and Force I

dfittool (X4)
dfittool (LunghezzaP)% Plot of X4 and LunghezzaP

h3 = kstest2 (X3, Force I)% Test Kolmogorov-Smirnoff between X3 and
Force I
h4 = kstest2 (X4, LunghezzaP)% Test Kolmogorov-Smirnoff between X4 and
 % LunghezzaP
```

```
nu = 3; % Degrees of Freedom

tau = corr (x (:, 3), x (:, 1),'Type','Kendall'); %% Returns the matrix
of correlation coefficient between the two marginal LunghezzaP and TIMEM
of the x input data matrix.

rho = copulaparam ('T', Tau, nu, 'Type','Kendall');% Returns the linear
correlation parameters corresponding to the copula t.


% N generate random vectors from the t copula with linear correlation
parameters rho.

n = 415;
```

```
U = copularnd ('T'[Rho 1; rho 1], nu, n);% N vectors generate random
from
 % T copula with parameters
 % Rho linear correlation.

X5 = ksdensity (x (:, 3), U (:, 1),'Function','ICDF','Width', 0.05);
X6 = ksdensity (x (:, 1), U (:, 2),'Function','ICDF','Width', 0.02);

% The ksdensity function allows you to make an estimate of the marginal
distributions (smoothing of the trend) of the copula t obtained.


scatterhist (X5, X6,'Direction','Out')Son instogrammi% of the X5 and X6

dfittool (X5)
dfittool (LunghezzaP)% Plot the X5 and LunghezzaP

dfittool (X6)
dfittool (TempoF)% Plot X6 and TempoF

kstest2  (X5,  KunghezzaP)%  Test  Kolmogorov-Smirnoff  between  X5  and
LunghezzaP
kstest2 (X6, TempoF)% Test Kolmogorov-Smirnoff between X6 and TempoF
```

5.3.2.1 Statistical Test

The table below shows the results obtained by performing the Kolmogorv-Smirnov test for the subject A1.

Test Kolmogorov-Smirnoff	Result (ls = 5%)
h1 = kstest2 (X1, TempoF)	0
h2 = kstest2 (X2, Force I)	0
h3 = kstest2 (X3, Force I)	0
h4 = kstest2 (X4, LunghezzaP)	0
h5 = kstest2 (X5, LunghezzaP)	0
h6 = kstest2 (X6, TempoF)	0

According to the results obtained in the table, each of the marginal distribution estimates of the copula t obtained, has the same distribution of the corresponding marginal used as input to generate it.

5.3.2.2 Comparison Charts

Histogram graph of X1 and X2

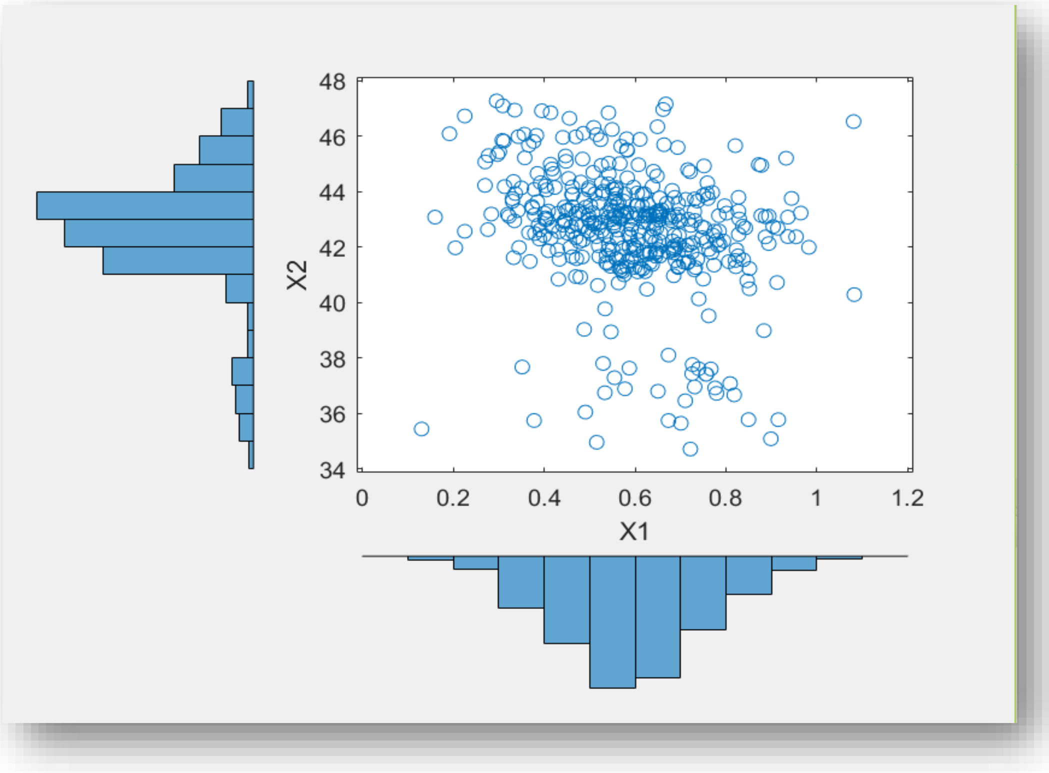

Figure 53Graph showing on the abscissa of the histogram X1 and in ordinate of the histogram X2

In the figure they are shown X1 and X2, namely the two marginal distributions extracted from copula t generated and estimated thanks to the "ksdensity" function.

The variables are represented through two histograms, the axis of abscissas corresponds to X1 to X2 while that on the ordinate. In the middle shows the points touched by the two distributions, it is noted that the highest concentration area is identified by the following coordinates: 0,4- 0,8 sec and 41.0 to 45.0 N.

Graph of histogram X3 and X4

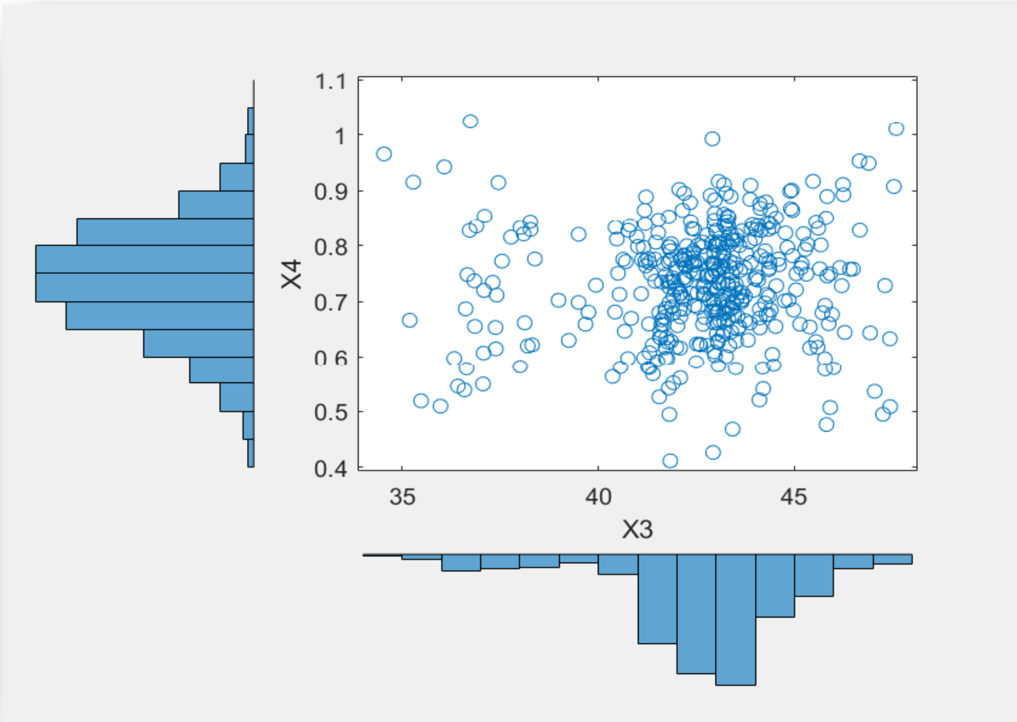

Figure 54Graph showing on the abscissa of the histogram X3 and X4 in the ordinate of the histogram

In the figure they are shown X3 and X4, namely the two marginal distributions extracted from copula t generated and estimated thanks to the "ksdensity" function.

The vectors are represented through two histograms, the axis of abscissas corresponds to X3 while that on the ordinate to X4. In the middle I come represented the points touched by the two distributions that mainly accumulate in the intervals comprised between 40- 45 N and 0.6 - 0.9 seconds.

Graph of histogram X5 and X6

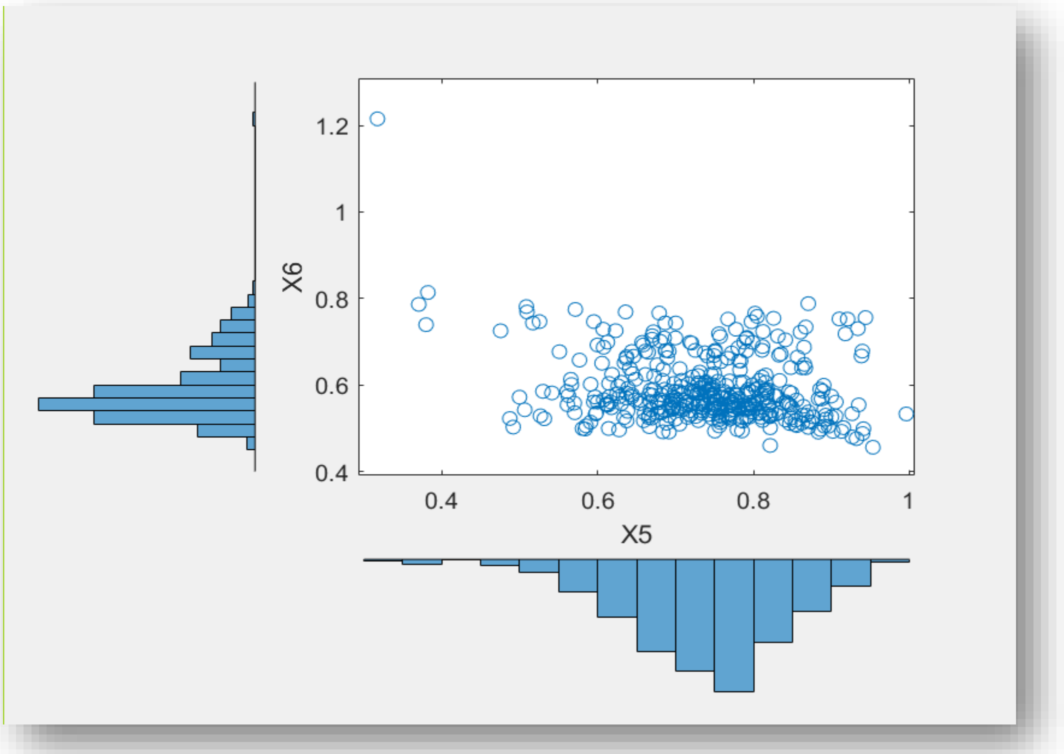

Figure 55Graph showing on the abscissa of the histogram X5 and X6 on the ordinate of the histogram

In the figure they are shown X5 and X6, ie the two marginal distributions extracted from copula t generated and estimated thanks to the "ksdensity" function.

The vectors are represented through two histograms, the axis of abscissas corresponds to X5 while that on the ordinate to X6. In the middle I come represented the points touched by the two distributions. One can note that most of the values is thickened in the intervals comprised between 0.6 and 0.085 me 0.45 and 0.8 seconds.

Comparison X1-TempF

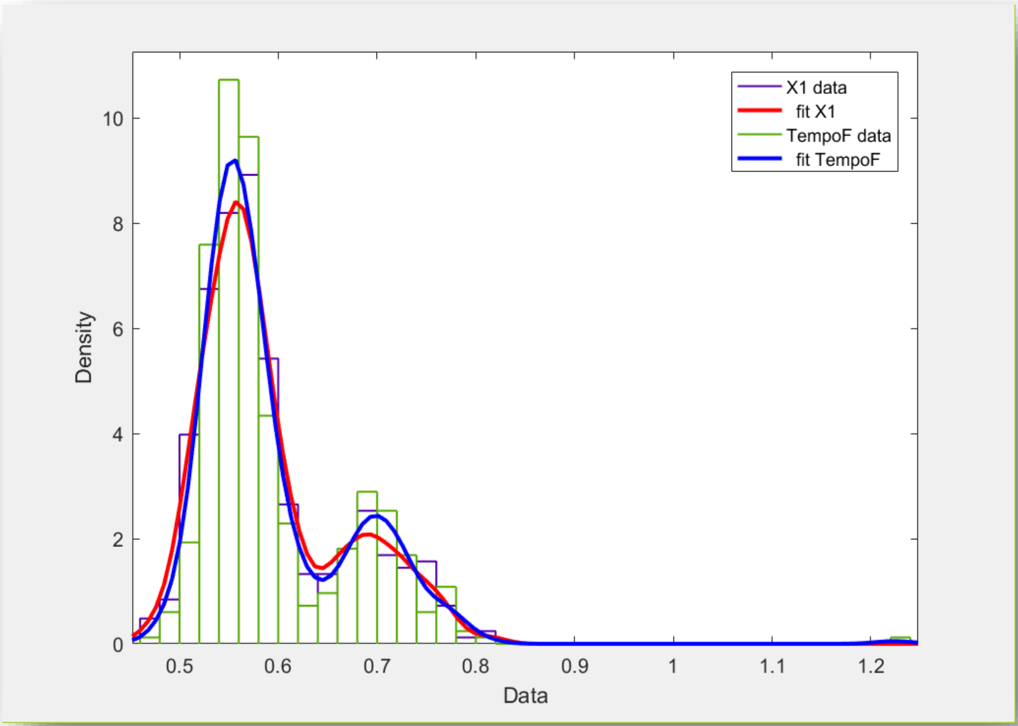

Figure 56, Graph histograms that shows the trend of the marginal TempoF and the carrier generated X1

The figure above shows the trend of X1 in purple (red curve) and TempoF in green (blue curve).

It is noted that the graph reflects the result obtained by kstest2, in fact, the two curves have very similar trends, ie a peak, where it has a high concentration of values, in the range between 0.5 and 0.6 and a second lower peak in 'range from 0.65 to 0.8 seconds. The approximation of the generated curve is defined by the parameter "width" of the "ksdensity" function with which it is possible to modify the bandwidth and thereby obtain a more or less precise performance, in this case has been set at 0.02. It can be said that the graph obtained confirming the outcome of the test Kolmogorv-Smirnov, namely that X1 and TempoF come from the same continuous distribution.

Compare X2-Force I

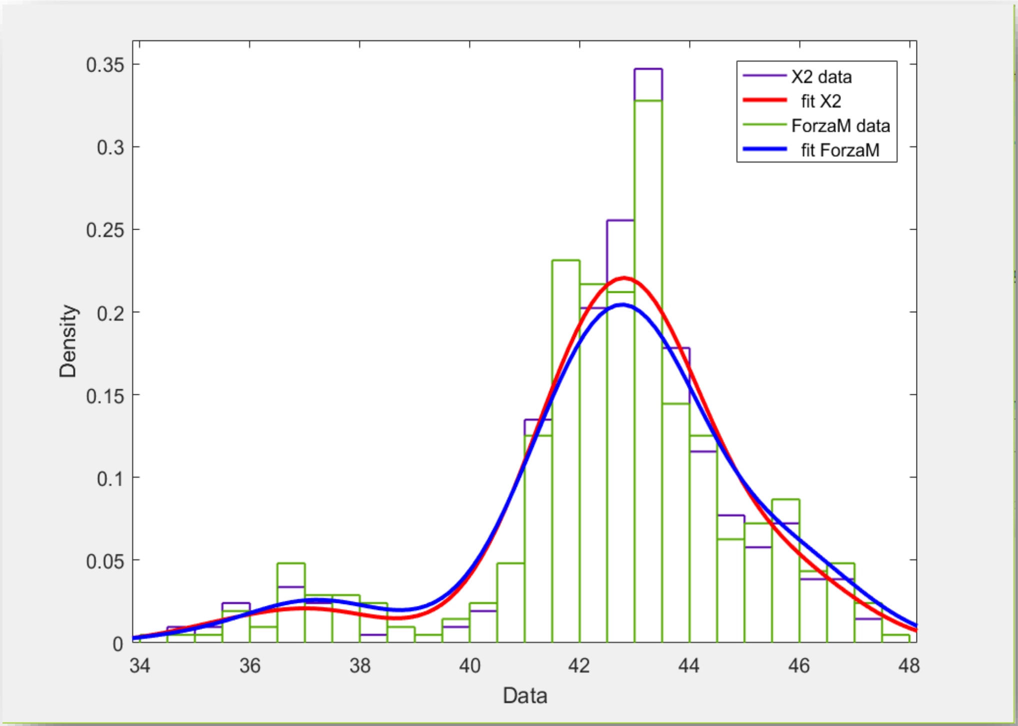

Figure 57, Graph histograms that shows the trend of the marginal Force I and the vector generated X2

The figure shows the trend of generated variable X2 in purple (red curve) and in green that of the input variable Force I (curve in blue).

The trend of the two curves shows a strong peak values in correspondence interval that goes from 42 to 44 N. Also in this case one can note how well the two curves prove almost entirely overlapping, this is also confirmed by kstest2 that , being the result 0, certifies that they come from the same continuous distribution. For this case, the "width" parameter that determines the approximation of the curve and has been set at 0.15.

Compare X3-Force I

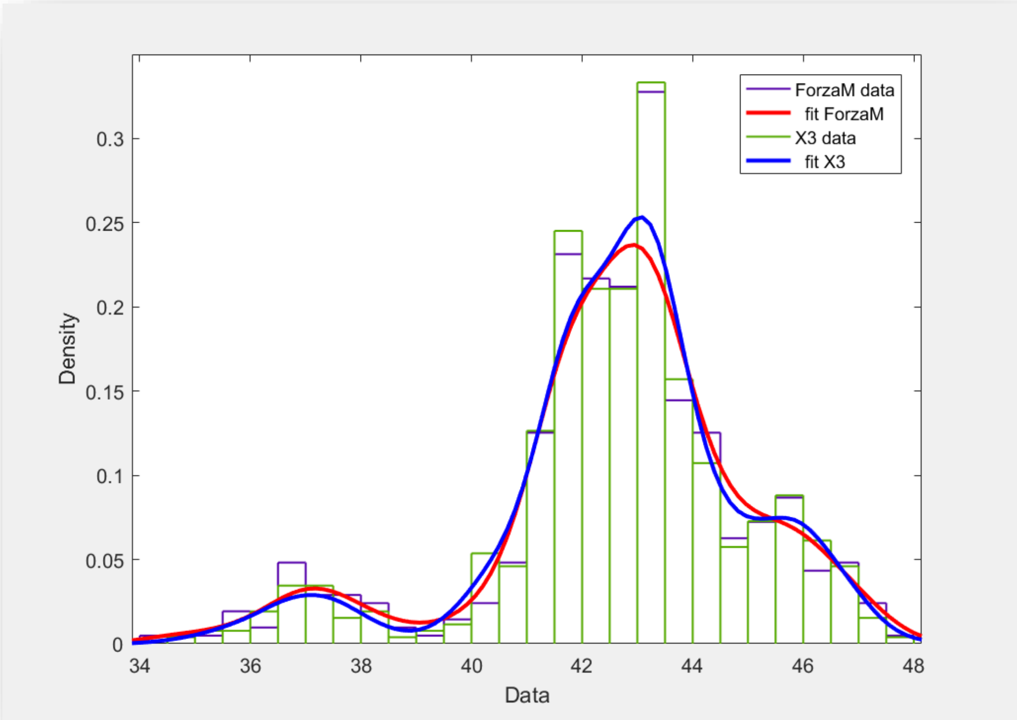

Figure 58, Graph histograms that shows the trend of the marginal Force I and the generated vector X3

The figure shows the trend of Force I variable in purple (red curve) and in green that the generated variable X3 (blue curve).

It can be seen as the trends of the two curves are very similar, with a peak in the range of values comprised between 41 and 44 N, this is also confirmed by the Kolmogorv-Smirnov test, which certifies that the two variables follow l ' trend of the same continuous distribution. Again, the "width" parameter was set to 0.15.

Compare X4-LunghezzaP

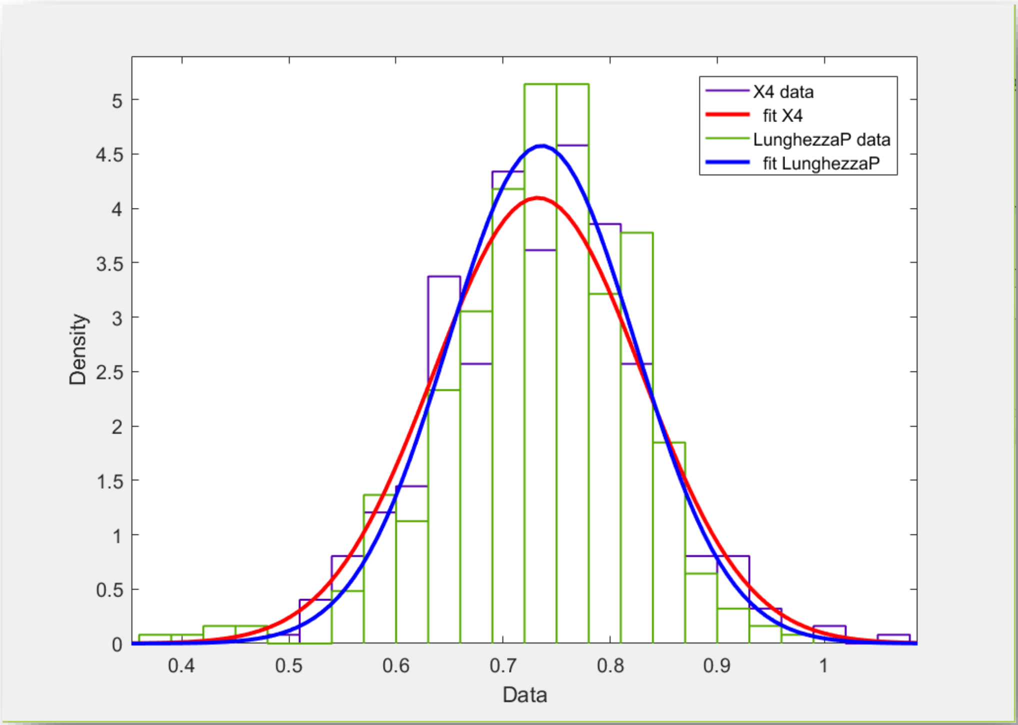

Figure 59, Graph histograms that shows the trend of the marginal LunghezzaP and the carrier generated X4

The figure shows the trend of the input variable LunghezzaP in green (blue curve) and in purple what the generated variable X4 (red curve).

Also in this case the two curves have very similar trends with a peak of values comprised in 'interval ranging from 0.6 to 0.85 seconds. Moreover, even the Kolmogorv-Smirnov test is in accordance with the graphic result, because it confirms that the two samples come from the same continuous distribution. The "width" parameter was set to 0.05 in this case.

Compare X5-LunghezzaP

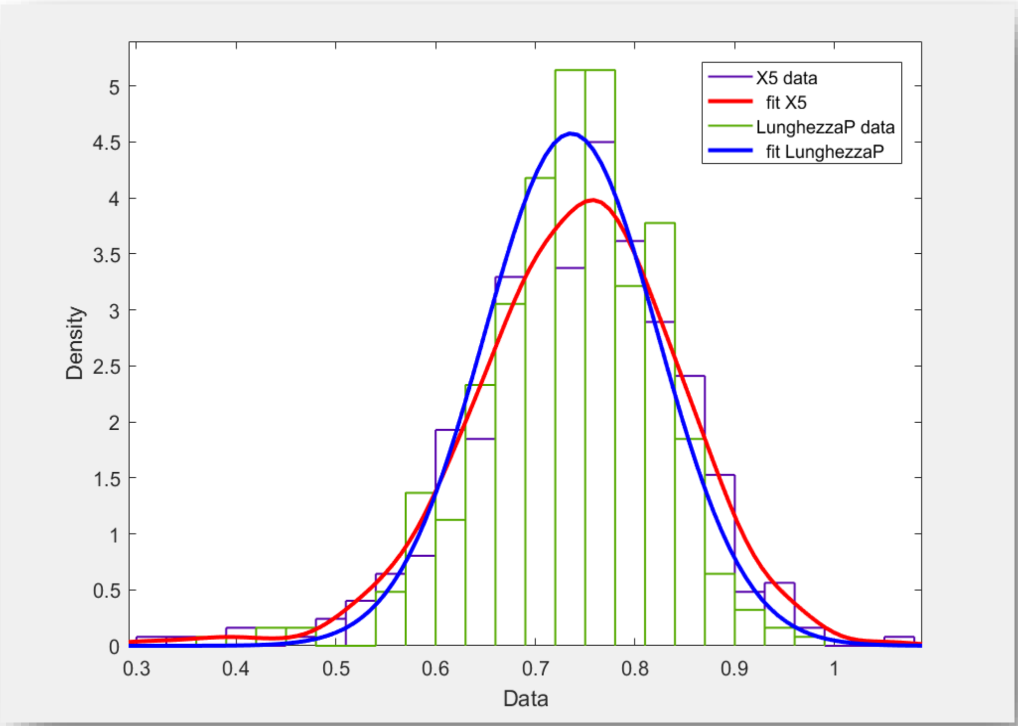

Figure 60, Graph histograms that shows the trend of the marginal LunghezzaP and the generated vector X5

The figure shows the trend of LunghezzaP in green (blue curve) and in purple to X5 (red curve).

The results obtained are very similar to those of the previous case, for which, also considering the outcome of the Kolmogorv-Smirnov test, it can be concluded that even X5 and LunghezzaP are distributed in the same way. The "width" parameter that determines the approximation of the generated curve X5, has been set to 0.05 in this case.

Compare X6-TempoF

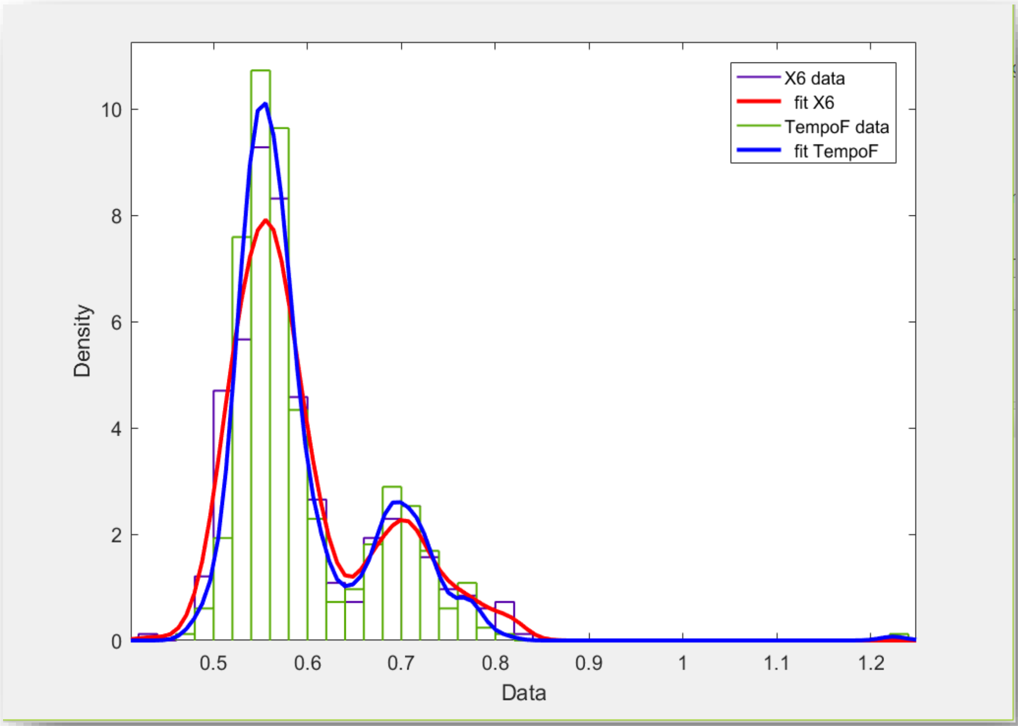

Figure 61, Graph histograms that shows the trend of the marginal TempoF and the generated vector X6

The figure shows the trend of the input variable TempoF in green (blue curve) and in that the generated variable X6 purple (red curve).

The two curves have similar trends, with values mostly concentrated in the range of 0.75 and 0.8 seconds. Even in this case, in agreement with the result obtained from Kolmogorv-Smirnoff test, it can be asserted that the LunghezzaP variables and X6 follow the trend of the same continuous distribution.

Through this method it is therefore possible to obtain the variables backhoe (for example:(X_1, X_2, X_4) e $(X_6, X_{3,} X_5)$) That they represent different walkers sample that can be used as input for a simulation.

As a result, we will be reported the results obtained using the other subjects as input.

5.3.2.3 Results obtained

Subject A2

The table below shows the results obtained by performing the Kolmogorv-Smirnov test for the subject A1.

Test Kolmogorov-Smirnoff	Result (ls = 5%)
h1 = kstest2 (X1, TempoF)	1
h2 = kstest2 (X2, Force I)	0
h3 = kstest2 (X3, Force I)	0
h4 = kstest2 (X4, LunghezzaP)	0
h5 = kstest2 (X5, LunghezzaP)	0
h6 = kstest2 (X6, TempoF)	1

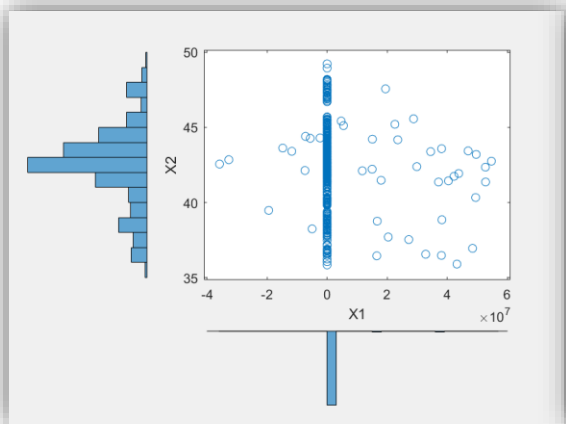

Figure 69 Graph showing on the abscissa of the histogram X1 and in ordinate of the histogram X2 (*)

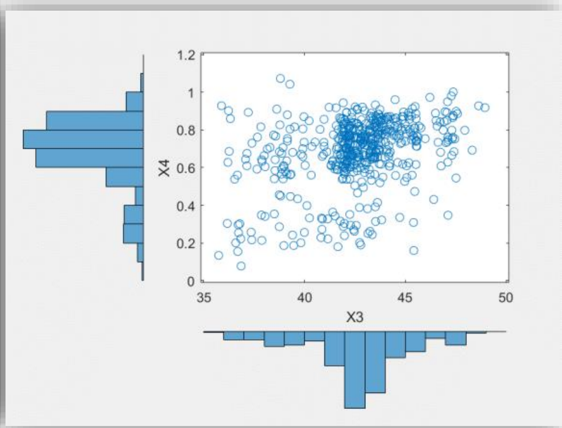

Figure 70, Graph showing on the abscissa of the histogram X3 and in ordinate of the histogram X4

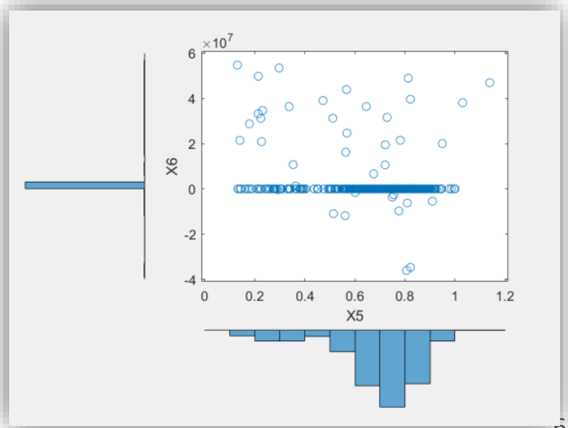

Figure 62Graph showing on the abscissa of the histogram X5 and X6 on the ordinate of the histogram (*)

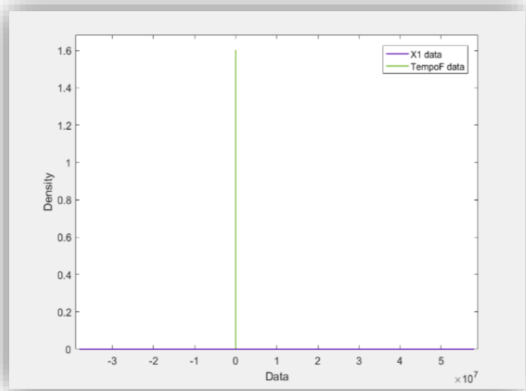

Figure 63, Graph histograms that shows the trend of the marginal TempoF and the carrier generated X1 (*)

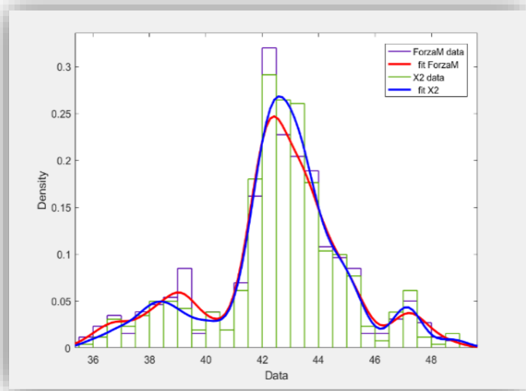

Figure 64, Graph histograms that shows the trend of the marginal Force I and the vector generated X2

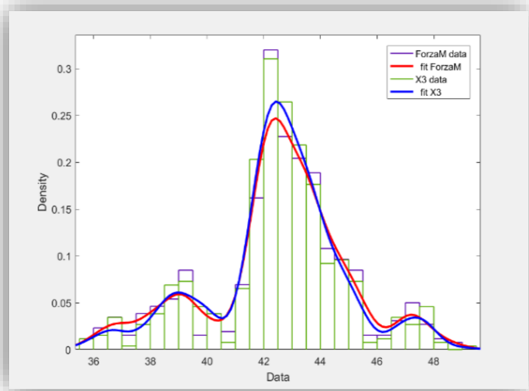

Figure 74, graph histograms that shows the trend of the marginal Force I and the generated vector X3

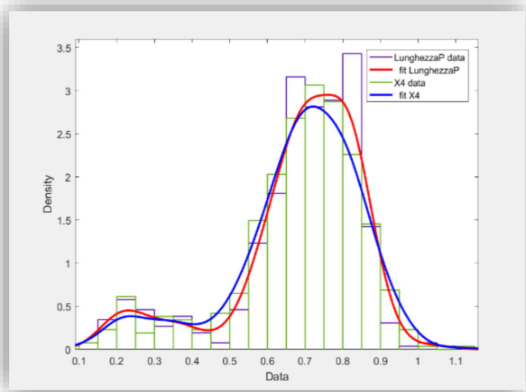

Figure 75, graph histograms that shows the trend of the marginal LunghezzaP and the carrier generated X4

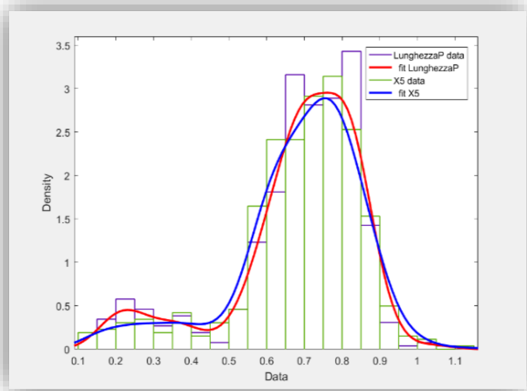

Figure 76, graph histograms that shows the trend of the marginal LunghezzaP and the generated vector X5

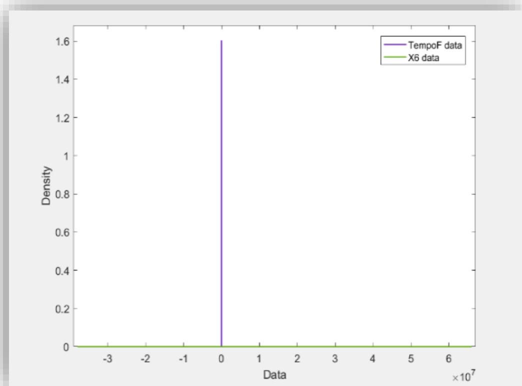

Figure 77, graph histograms that shows the trend of the marginal TempoF and the generated vector X6 (*)

Subject B1

The table below shows the results obtained by performing the Kolmogorv-Smirnov test for the subject A1.

Test Kolmogorov-Smirnoff	Result (ls = 5%)
h1 = kstest2 (X1, TempoF)	0
h2 = kstest2 (X2, Force I)	0
h3 = kstest2 (X3, Force I)	0
h4 = kstest2 (X4, LunghezzaP)	0
h5 = kstest2 (X5, LunghezzaP)	0
h6 = kstest2 (X6, TempoF)	0

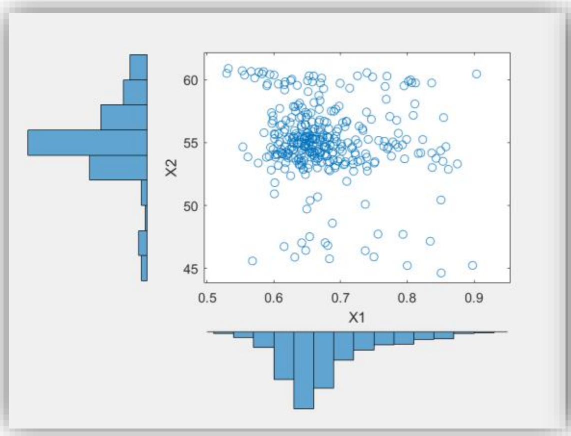

Figure 66Graph showing on the abscissa of the histogram X1 and in ordinate of the histogram X2

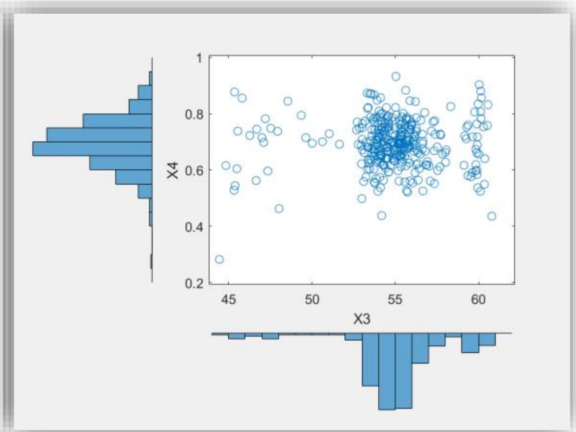

Figure 65Graph showing on the abscissa of the histogram X3 and X4 in the ordinate of the histogram

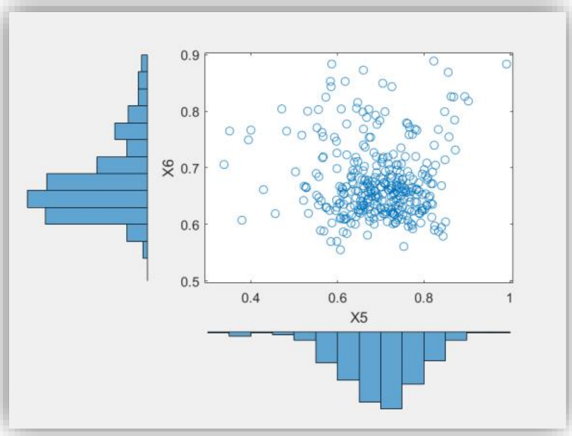

Figure 67Graph showing on the abscissa of the histogram X5 and X6 on the ordinate of the histogram

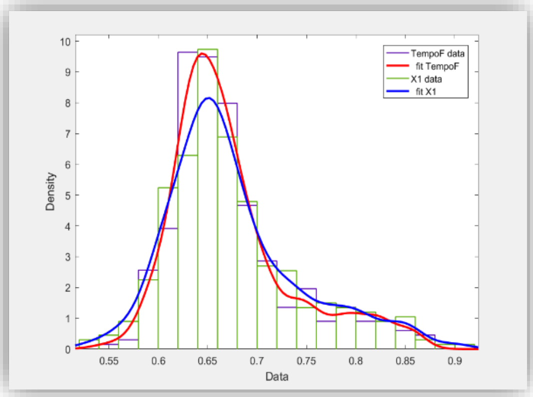

Figure 69, Graph histograms that shows the trend of the marginal TempoF and the carrier generated X1

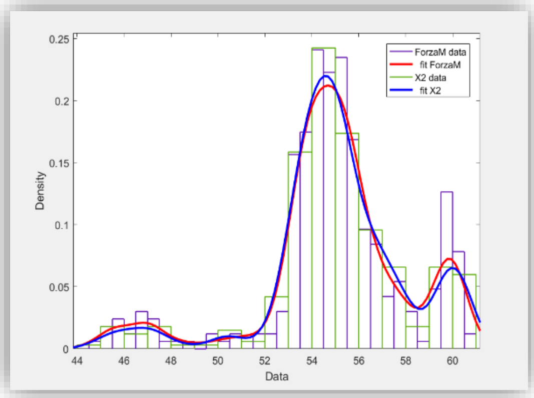

Figure 68, Graph histograms that shows the trend of the marginal Force I and the vector generated X2

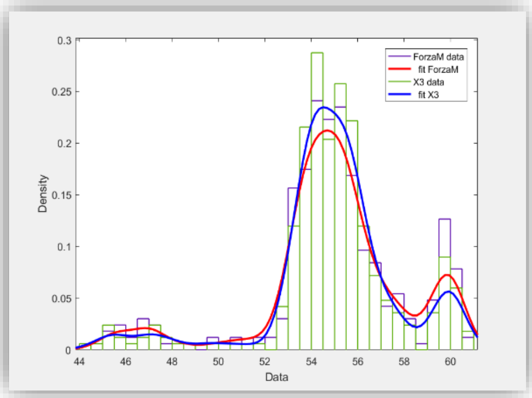

Figure 71, Graph histograms that shows the trend of the marginal Force I and the generated vector X3

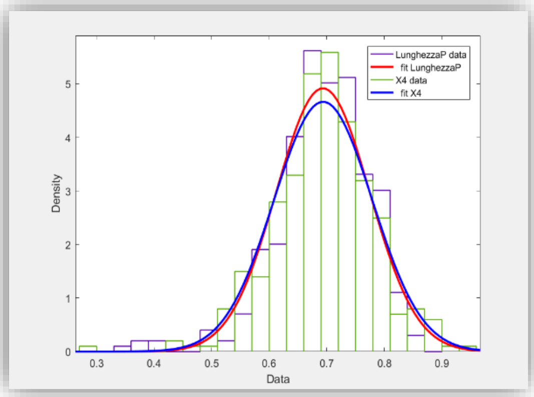

Figure 70,, Graph histograms that shows the trend of the marginal LunghezzaP and the carrier generated X4

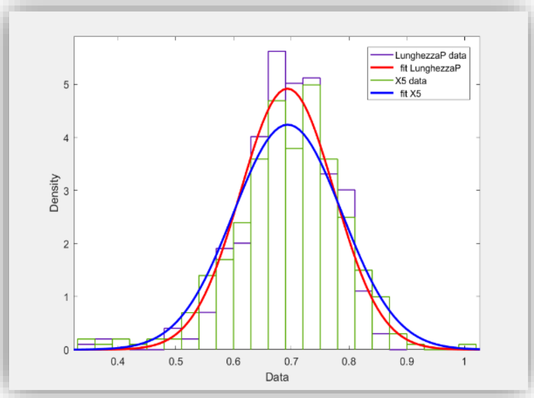

Figure 73, Graph histograms that shows the trend of the marginal LunghezzaP and the generated vector X5

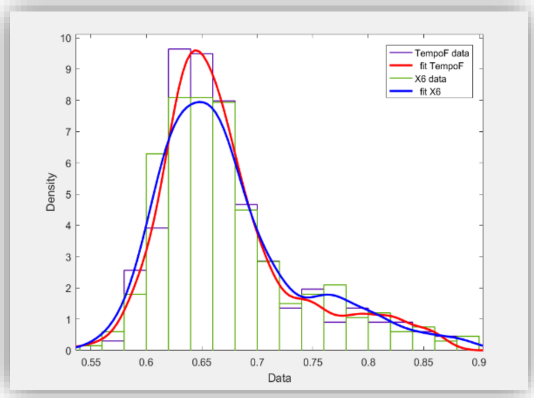

Figure 72, Graph histograms that shows the trend of the marginal TempoF and the generated vector X6

Subject B2

The table below shows the results obtained by performing the Kolmogorv-Smirnov test for the subject A1.

Test Kolmogorov-Smirnoff	Result (ls = 5%)
h1 = kstest2 (X1, TempoF)	0
h2 = kstest2 (X2, Force I)	0
h3 = kstest2 (X3, Force I)	0
h4 = kstest2 (X4, LunghezzaP)	0
h5 = kstest2 (X5, LunghezzaP)	0
h6 = kstest2 (X6, TempoF)	0

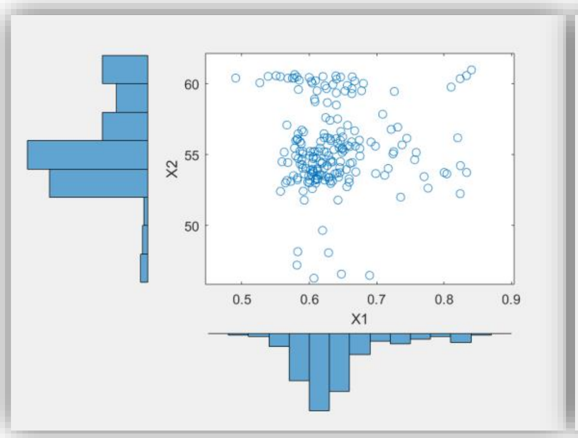

Figure 75Graph showing on the abscissa of the histogram X1 and in ordinate of the histogram X2

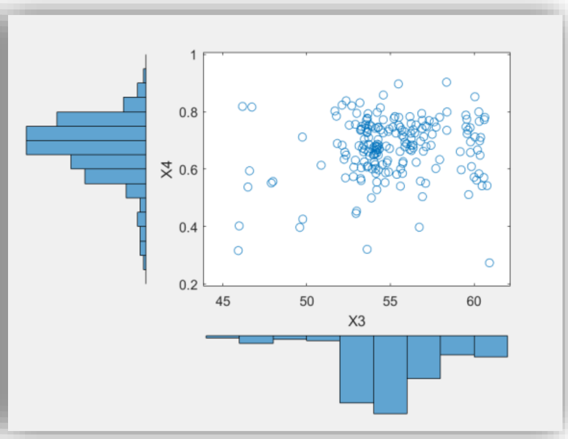

Figure 74Graph showing on the abscissa of the histogram X3 and X4 in the ordinate of the histogram

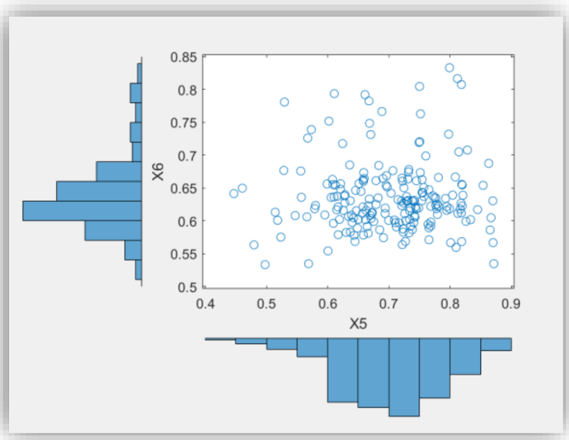

Figure 76Graph showing on the abscissa of the histogram X5 and X6 on the ordinate of the histogram

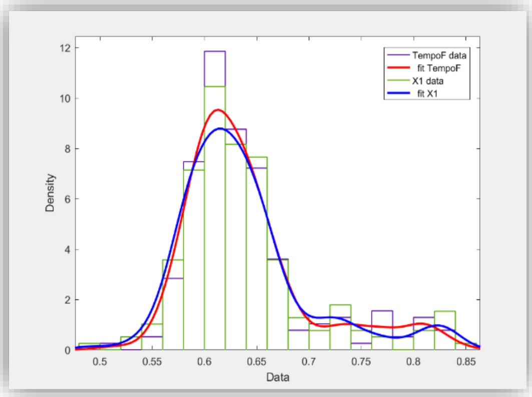

Figure 78, Graph histograms that shows the trend of the marginal TempoF and the carrier generated X1

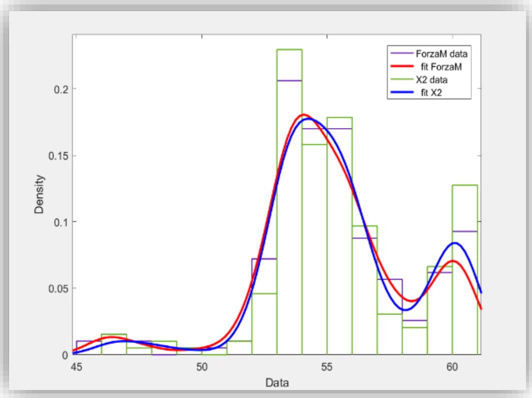

Figure 77, Graph histograms that shows the trend of the marginal Force I and the vector generated X2

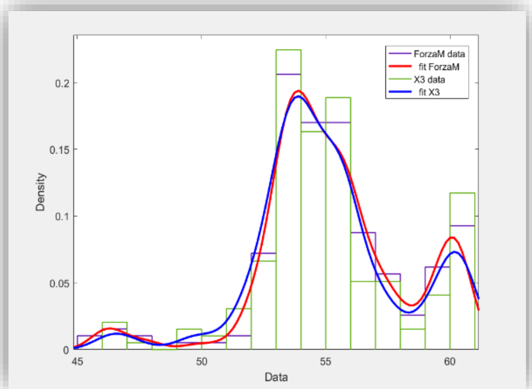

Figure 80, Graph histograms that shows the trend of the marginal Force I and the generated vector X3

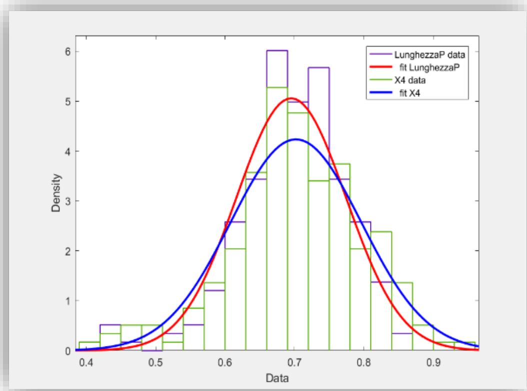

Figure 79, Graph histograms that shows the trend of the marginal LunghezzaP and the carrier generated X4

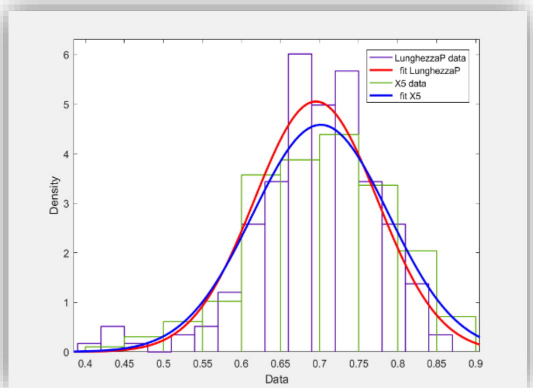

Figure 82, Graph histograms that shows the trend of the marginal LunghezzaP and the generated vector X5

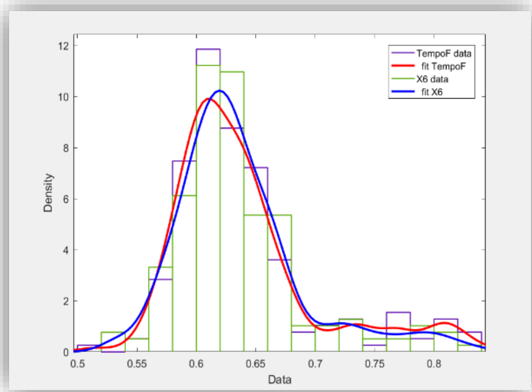

Figure 81, Graph histograms that shows the trend of the marginal TempoF and the generated vector X6

Subject C1

The table below shows the results obtained by performing the Kolmogorv-Smirnov test for the subject A1.

Test Kolmogorov-Smirnoff	Result (ls = 5%)
h1 = kstest2 (X1, TempoF)	0
h2 = kstest2 (X2, Force I)	0
h3 = kstest2 (X3, Force I)	0
h4 = kstest2 (X4, LunghezzaP)	0
h5 = kstest2 (X5, LunghezzaP)	0
h6 = kstest2 (X6, TempoF)	0

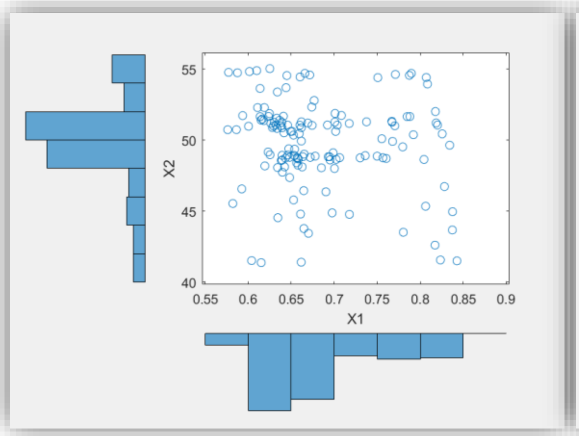

Figure 96: Graph showing in abscissa the X1 histogram and the ordinate of the histogram X2

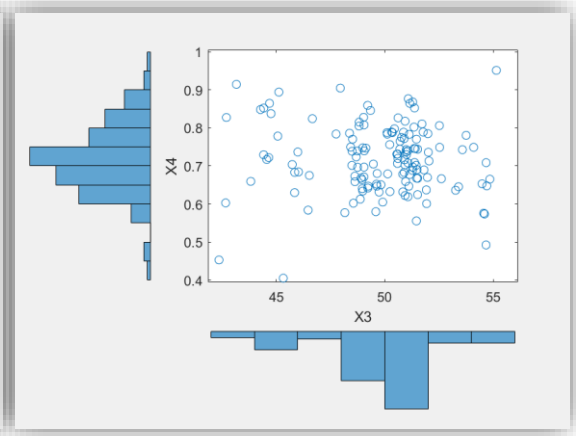

Figure 97, Graph showing on the abscissa of the histogram X3 and X4 in the ordinate of the histogram

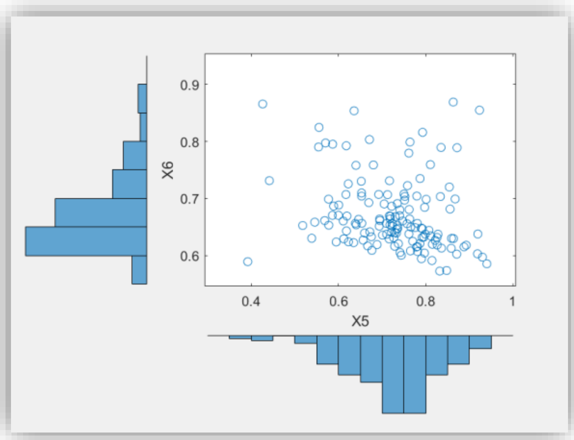

Figure 83Graph showing on the abscissa of the histogram X5 and X6 on the ordinate of the histogram

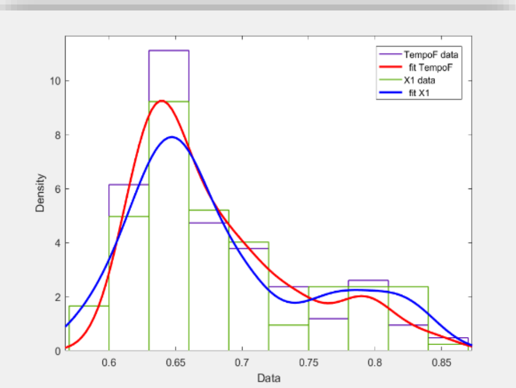

Figure 85, Graph histograms that shows the trend of the marginal TempoF and the carrier generated X1

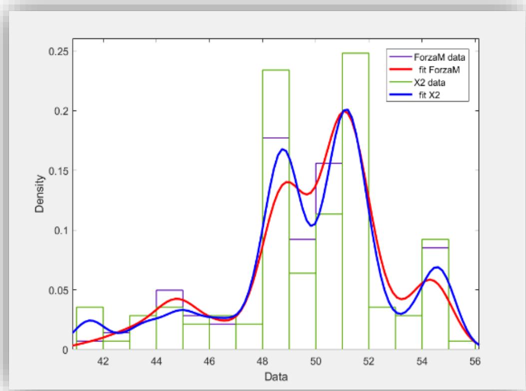

Figure 84, Graph histograms that shows the trend of the marginal Force I and the vector generated X2

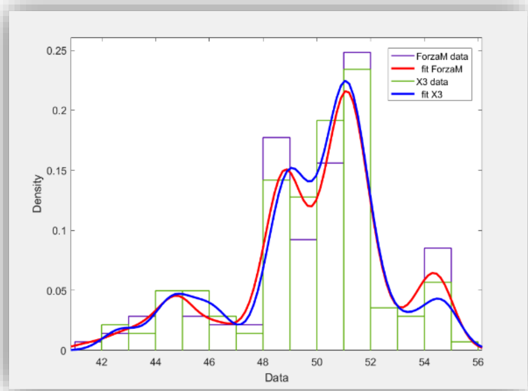

Figure 87, Graph histograms that shows the trend of the marginal Force I and the generated vector X3

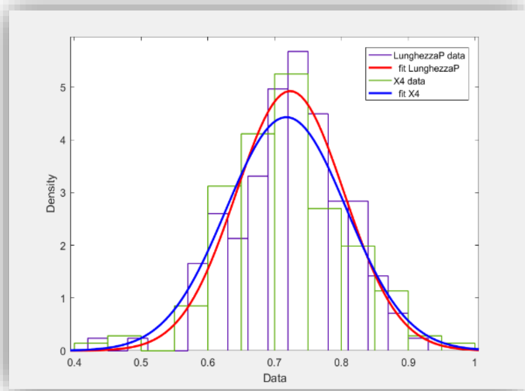

Figure 86, Graph histograms that shows the trend of the marginal LunghezzaP and the carrier generated X4

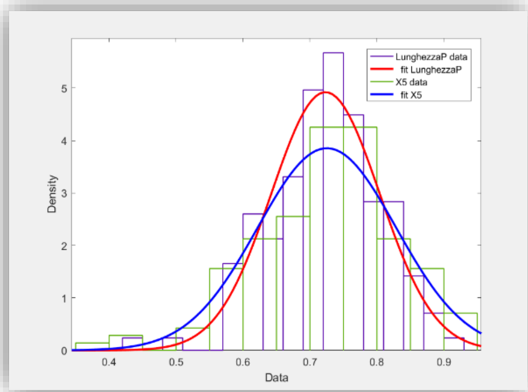

Figure 89, Graph histograms that shows the trend of the marginal LunghezzaP and the generated vector X5

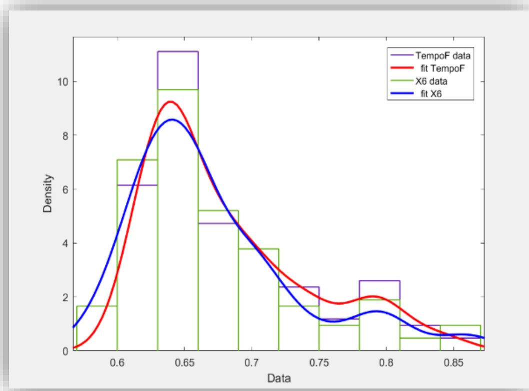

Figure 88, Graph histograms that shows the trend of the marginal TempoF and the generated vector X6

Subject C2

The table below shows the results obtained by performing the Kolmogorv-Smirnov test for the subject A1.

Test Kolmogorov-Smirnoff	Result (ls = 5%)
h1 = kstest2 (X1, TempoF)	0
h2 = kstest2 (X2, Force I)	0
h3 = kstest2 (X3, Force I)	0
h4 = kstest2 (X4, LunghezzaP)	0
h5 = kstest2 (X5, LunghezzaP)	0
h6 = kstest2 (X6, TempoF)	0

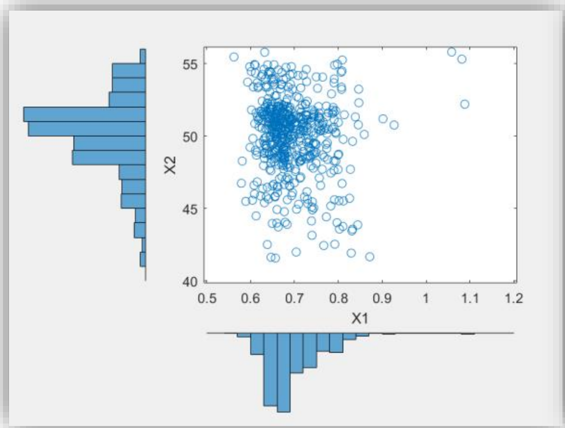

Figure 90Graph showing on the abscissa of the histogram X1 and in ordinate of the histogram X2

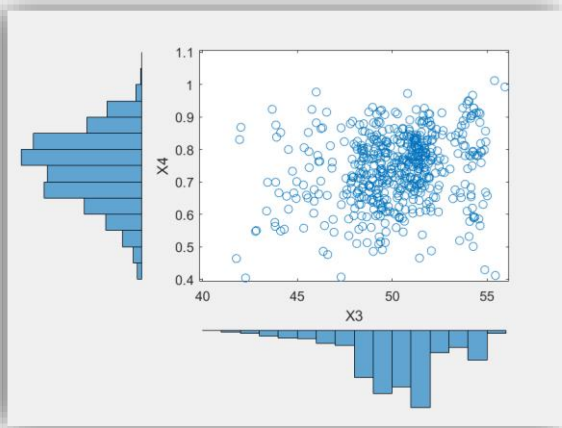

Figure 91, A graph that shows on the abscissa of the histogram X3 and X4 in the ordinate of the histogram

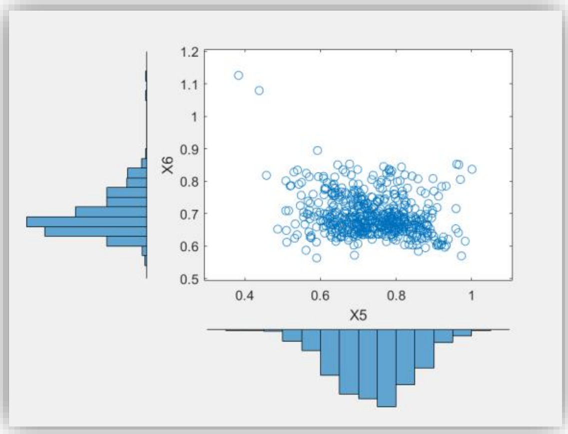

Figure 92, A graph that shows on the abscissa of the histogram X5 and X6 on the ordinate of the histogram

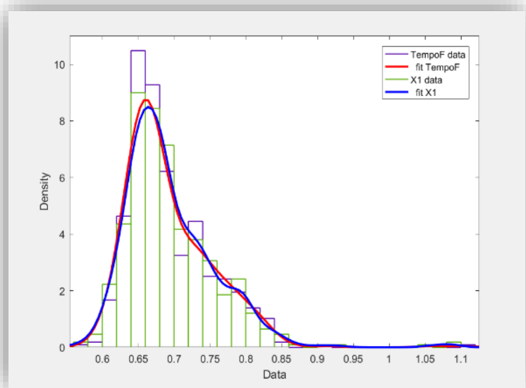

Figure 100, graph histograms that shows the trend of the marginal TempoF and the carrier generated X1

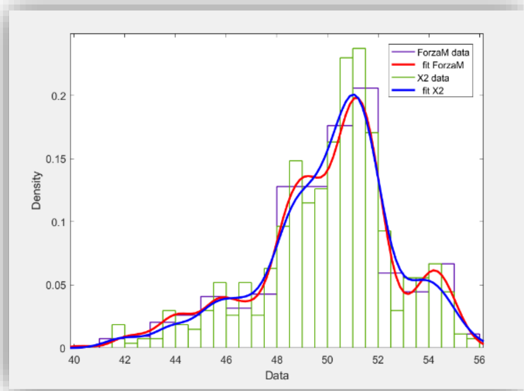

Figure 101, graph histograms that shows the trend of the marginal Force I and the vector generated X2

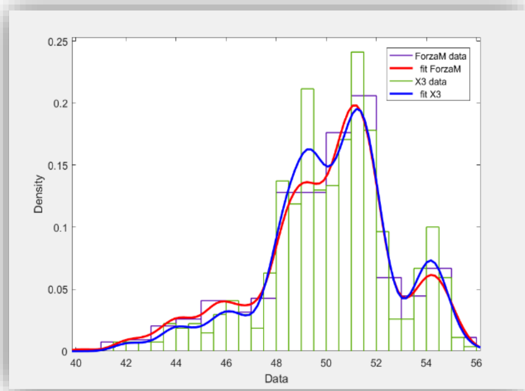

Figure 103, graph histograms that shows the trend of the marginal Force I and the generated vector X3

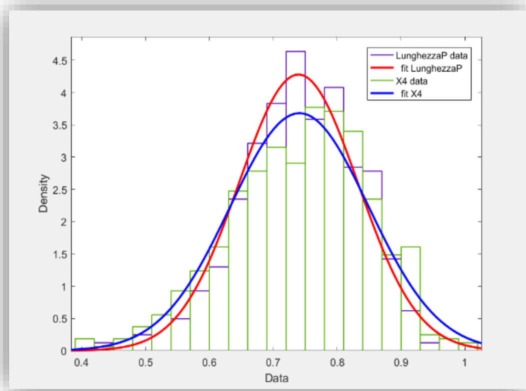

Figure 104, graph histograms that shows the trend of the marginal LunghezzaP and the carrier generated X4

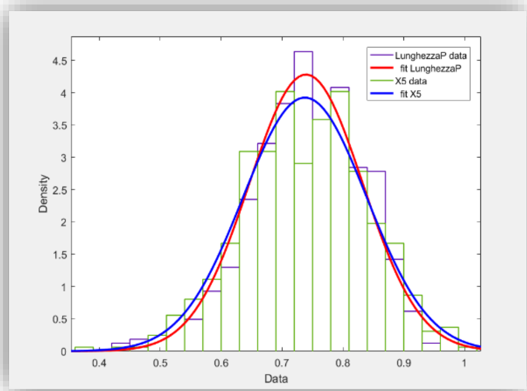

Figure 105, graph histograms that shows the trend of the marginal LunghezzaP and the generated vector X5

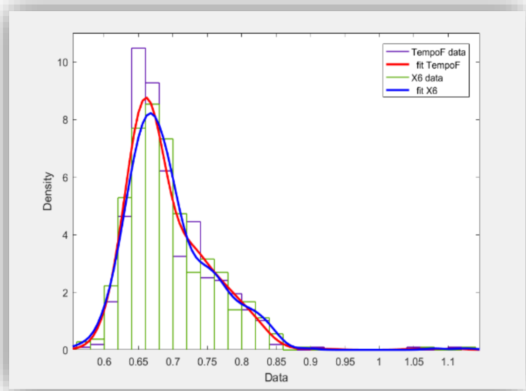

Figure 106, graph histograms that shows the trend of the marginal TempoF and the generated vector X6

6 Conclusions

Before moving to conclusions, I would like to point out that the study was influenced by the "goodness" of the sample data used in the study. As mentioned several times, the purpose of this analysis is to find a model that best describes the man walking process to be used in the design phase and improve the comfort of pedestrian walkways.

- The first type of tests carried out showed that it is not possible to describe the phenomenon of interest through a multivariate normal distribution, since each sample tested did not pass the Kolmogorv-Smirnov test. It is noted, however, that at each test (not considering the Subject A2) the variable that corresponds to the step length and the corresponding marginal extracted from the multivariate normal distribution generated, are distributed in the same way. This means that the "step length" variable may be distributed according to the normal.

- The test results relating to the Gaussian mixture models have been more than satisfactory, in fact, excluding the subject A2, each sample generated by the "mixture" of the three components of interest, it has exceeded the Kolmogorv-Smirnov test. One can therefore conclude that the system formed by the three variables of interest (Time, strength and length) could be described by a Gaussian Mixture Model.

- Finally, if you want to perform a simulation through which is generated a finite walker, ideally the input data should reflect what is known about the dependencies of the quantity that is being analyzed. Through the use of copulas is possible to do so, then create a walkers generator that takes into account the dependencies between the input variables can be used as input for a simulation.

6.1 FUTURE DEVELOPMENTS

There would be several points to be analyzed for future research. First, it should be emphasized that the results obtained are influenced by several limitations, already discussed previously. Therefore, in the future it would be desirable to overcome these problems, for example, by using, for the measurements, pedestrian structures of greater dimensions in order to allow subjects tested to perform multiple steps to every stride. It would also be sensible to perform simulations to test the response of the structure of a pedestrian walkway excitation induced by a walker. To this purpose it could be used copulas of the method developed for generating synthetic samples to be used as input.

Finally, this study takes into account only a pedestrian who moves at a constant speed, it would therefore be interesting to consider also the behavior of people in transit groups on the same walkway pedestrian, since there still exists a group-force model that is generally accepted.

Bibliography

[1] CLCD MRM Antinori, pedestrian walkways steel, Palermo: Dario Flaccovio Publisher, 2017

[2] VTRHJ a. TF lnman, Human Walking, Philadelphia, PA: Williams & Wilkins, 1981

[3] M. Perc, The Dynamics of Human Gait, Eur. J. Phys., 2005, p 525-534.

[4] CLDBL a. JCO Vaughan, Dynamics of Human Gait, Cape Town, South Africa: Kiboho Publishers, 1992

[5] E. Ayyappa, Normal Human Locomotion, Part I: Basic Concepts and Terminology, American Academy of orthotists & Prosthetists, 1997.

[6] No Messenger, Moving the Human Machine: Understanding the Mechanical Characteristics of Normal Human Walking, 1994, p. 352-357

[7] MW Whittle, Gait Analysis: An lntroduction, Oxford, UK: Butterworth-Heinemann, 2014

[8] JGJG a. AJM Rose, Human Walking, Philadelphia, PA: Lippincott Williams & Wilkins, 2006.

[9] KMSAB Newell, "The Nature of Movement Variability," in Motor Behavior and Human Skill, Champaign, Human Kinetics Publishers, 1998, pp. 143-160.

[10] Martina Fornaciari, Development of a mathematical model for the two-legged walking: Application in the study of stability of pedestrian walkways, University of Modena and Reggio, 2018

[11] Alessandro Lubisco, Analysis Multivariate Statistics, University of Bologna, 2009

[12] Daniela de Canditiis, MULTIVARIATE RANDOM VARIABLE, pp. 1-2, 2013.

[13] W. Navidi, Probability and statistics for engineering and the sciences, McGraw-Hill, 2006.

[14] Sheldon M. Ross, Probability calculation, Apogee, Milan, pp 233-234, 2004

[15] Massey, FJ "The Kolmogorov-Smirnov Test for Goodness of Fit". Journal of the American Statistical Association. Vol. 46, No. 253, 1951, pp. 68-78.

[16] Miller, LH "Table of percentage points of the Kolmogorov statistics." Journal of the American Statistical Association. Vol. 51, No. 273, 1956, pp. 111-121.

[17] Marsaglia, G., J. W. Tsang and Wang. "Evaluation of the distribution of Kolmogorov." Journal of Statistical Software. Vol. 8, No. 18, 2003.

[18] Luigi Accardi, Yun-Gang Lu, Igor Volovich, Probability towords 2000 Springer-Verlag New York, Inc., 1998, pp 319-321

[19] Geoffrey R. Grimmett & David R. Stirzaker, Probability and Random Processes, Third_edition, Oxford University, pp 115-117, 2001

[20] McLachlan, G., and D. Peel. Finite Mixture Models. Hoboken, NJ: John Wiley & Sons, Inc., 2000.

[21] Douglas A. Reynolds, Gaussian Mixture Models. In: Li SZ, Jain, A. (eds) Encyclopedia of Biometrics. Springer, Boston, MA, 2009

[22] Xuexing Zeng, Jinchang Ren, Zheng Wang, Stephen Marshall, Tariq Durrani, Signal Processing, Elsevier, pp 691-702, 2014

Acknowledgments

I want to say a few words of thanks to the people who have helped me during this study. First of all, I would keep to thank Fabrizio Pancaldi and Claudio Giberti professors for their valuable advice and support received. Then, a special thanks to my friends and classmates, who have always supported me and heard in these three years. Finally, a special thanks are due to my family, that gave me the opportunity to continue their studies and has always believed in me, it is mainly to them that today I managed to reach this milestone.