Lessons learned using machine learning to link third body particles morphology to interface rheology

Rabii Jaza, Guilhem Mollon, Sylvie Descartes, Amandine Paquet, Yves Berthier

Version of Record: https://www.sciencedirect.com/science/article/pii/S0301679X20304588
Manuscript_3b682df9448c59e3358c4d698855d31d

Lessons learned using Machine Learning to link third body particles morphology to interface rheology

Rabii Jaza[1], Guilhem Mollon[1*], Sylvie Descartes[1], Amandine Paquet[1], Yves Berthier[1]

[1] Université de Lyon, LaMCoS, INSA-Lyon, CNRS UMR5259, F-69621, France

[*]Corresponding author: guilhem.mollon@insa-lyon.fr

SUMMARY

This paper reports a preliminary investigation on the ability of Machine Learning algorithms to relate the morphology of third body particles to the rheology of the contact interface that created them. A testing campaign is performed on a pin-on-disc tribometer, followed by a comprehensive observation of the worn surfaces. Several Machine Learning algorithms are then used to establish and quantify the logical relations between the rheological and a morphological databases built from this campaign. Success rates and thorough analysis of their predictions are used to validate the general approach and to propose possible improvements. It appears that Machine Learning presents an interesting potential in quantitative tribological analysis if the morphological and rheological databases are properly enriched.

Keywords: Third Body, Morphology, Rheology, Machine Learning

I. INTRODUCTION

The tribological triplet [1-4] is a conceptual framework that was proposed as an attempt to unify and structure the science of dry mechanical contacts by considering as a whole the different scales where the relevant physics take place. The scale of the third body (i.e. the layer of solid matter trapped in the interface and separating the contacting bodies) is particularly challenging because it is both pivotal in the understanding of friction and wear and hard to model. Indeed, if the third body is defined as the interfacial layer within a dry contact interface, it means that any piece of matter ejected from the contact is a wear debris; as such it is no more considered as being part of the third body in the contact. Likewise, when opening a contact to observe it, what is to be seen on the surfaces may be quite different from the third body that initially existed during the contact, because of a strong modification of the environment caused by the contact opening (dramatic changes in normal and tangential stress, release of internal stresses, no more strain rate, temperature cooling, exposition to the gaseous environment, etc.). The only relevant place to characterize the mechanical behaviour of the third body is thus inside a contact, and this is precisely where the third body is the less accessible. Moreover, any quantity measured in a classical tribometry experiment will be collected quite far from the contact itself. Hence, no tribometer can claim to be directly measuring the rheology of the third body (its flow capacity as a function of shear strain and rate coupled with adhesion on first bodies surface) without any bias.

Presented this way, this approach looks like an experimental dead-end. Apart from improving local probing of dry third body during contacts, a possible solution to circumvent this limitation is to employ numerical modelling of the interface. This does not come without difficulties, since the micrometric scale is notoriously challenging for numerical simulations: it is too large to use first principles simulations, and the kinematics of third body flow are too complicated and

poorly constrained to use fluid-like continuum approaches. The best candidates for such a task have been identified to be Discrete Element Modelling (DEM) [5] and its extensions and improvements [6-10]. Among them, the Multibody Meshfree Approach was recently proposed [11-12]. In this approach, the main principle is to consider the third body layer as an assembly of a large number of deformable grains. Each grain has its own properties in terms of size, shape, mechanical constitutive behaviour, and interaction laws with its neighbouring grains. Recent numerical results have shown that such simulations are able to describe a large variety of flow regimes which experimentally make sense, and to provide the corresponding frictional response, as shown for example in [13]. However, reproducing reality is not predicting it, and many open questions remain regarding the realistic character of such simulations. Yet, they provide both local and dynamic data, while experiments provide data that are either dynamic but global (e.g. friction coefficient provided by a tribometer) or local but static (e.g. post-mortem observation of surfaces). As such, they have the potential to complement nicely classical tribometry if they are experimentally validated. It requires to verify that the predicted flow regimes are in agreement with the morphologies observed *post-mortem* in the contacts.

From an experimental viewpoint at the local scale, the examination of third body morphology started as individual characterization of wear debris. Several studies tried to relate every shape they found to a possible source. Seven shapes have been proposed, only based on visual characterization, with their possible origins [14-16]. Other studies focused on giving metric dimensions to the feature describers when relating them to the wear type taking place in the contact [17-21], highlighting a wide range of size. For example, abrasive wear may lead to particles that have a thickness from 2 µm to 5 µm and a length from 25 µm to 100 µm. In fatigue wear, particles are found as spherical grains with a radius from 3 µm to 10 µm, as a group of small particles with a width ranging from 10 µm to 100 µm or as laminar particles with a length

starting from 20 µm to 50 µm [22]. The work of Kowandy et al. [23] shares some similarities with the present study, such as studying the morphological properties of third body particles created during a tribological experiment and then relating those descriptors to the wear regime taking place in the contact. In their work, the chosen descriptors were the length, the perimeter, the area and the Spike parameter (an indicator for the angularity of the particle [24]).

The next step for relating wear particles to tribological scenarios was to use Machine Learning as an automatized expert system. Based on first attempts to extract data from wear particle images [25-26], several automated statistical approaches were proposed for wear particles classification such as Artificial Neural Networks (ANN) [27-29], Support Vector Machines (SVM) [30-31, or Classification and Regression Trees (CART) [32]. More recently, a coupled approach using both conventional ANN and Deep Learning was applied to the same problem with promising results [33]. For all these studies, wear particles were first extracted from the contact and isolated by techniques such as ferrography, in order to acquire the microscopic images and to build relevant databases. This is a limitation, because (i) this operation involves a large amount of human work and expertise and (ii) the displacement of the particles may have caused the modification of their morphology due to their fragility. For these reasons, the number of particles analysed was typically pretty low. However, this research proved the efficiency of using wear debris morphology to understand the wear that takes place in the contact.

As a continuation of these pioneering works, a preliminary methodological study is proposed in this paper, which aims to further explore the experimental gap between two classes of experimental results regarding third body: Rheology and Morphology. Both terms deserve to be well-defined in that context. For the present framework rheology means any quantitative information that describes the flow of the third body. Since no such direct measurement exist,

it is necessary to rely on indirect measurements such as the friction coefficient measured by a tribometer. Hence, the term rheology in the context of this study does not follow the conventional practice (where “rheology” often means “rheological law”) but is to be understood in a much broader sense. It could also include any evolution or transformation of the third body during the sliding contact, may it be from micromechanical or physico-chemical points of view, for example (although, as detailed in the next section, the reported rheological data is currently purely mechanical, in the form of a friction coefficient). Morphology, on the other hand, will describe any information regarding the shape and the aspect of the third body as observed after opening the contact and observing it at the micrometric scale. Following the classification proposed in the previous paragraph, rheological data is “global and dynamic” while morphological data is “local and static”. To build this link, the following methodology is employed: a tribological rig (described in Section II) is used to run an experimental campaign while systematically varying several key parameters, and a structured rheological database is built from these results (Section III). The worn surfaces of the first bodies are then systematically submitted to microscopic observation in order to extract morphological properties of the created third bodies, and to build a morphological database (Section IV). Machine learning techniques are then applied to determine to what extent these two databases can be logically related (Section V). Finally, the limitations and the potential of this methodology are discussed and the lessons learned when applying it for the first time are detailed (Section VI).

II. EXPERIMENTS

This study does not seek to recreate a specific industrial contact but to investigate a concept. To do so, a dedicated tribological test bench is set up (a pin-on-disk tribometer shown in Fig. 1) in order to have control on the majority of the parameters. The tribometer consists in an arm holding the pin in contact with the disk surface, and a motor to keep the disk in a continuous rotational motion. The counterweights are used to oppose to the weight of the pin, of its supporting arms and of the different components of the arm itself (force sensor, fixing screws, etc.). When applying 0 N, the system is in a state of equilibrium and the normal force detected at the tip of the pin is very close to zero. Once the pin is secured on its supporting arm, this arm keeps two rotational degrees of freedom. The axis system in Fig. 1 is used as reference. The pin's supporting arm can rotate around the *Y-axis* so that the pin can go up and down. It is only stopped by the disk's surface. In addition, the arm has approximately ±2° of rotational freedom angle to turn around the *Z-axis*. This freedom is used by a force sensor mounted on the supporting arm to follow the tangential force. The sensor used is the SIKA FTCA50. It is a S-type sensor that can measure compression and tension. Its measuring range goes from 1 to 50 N. The sensor sends its signal through a conditioner to an acquisition system (OROS35) to analyse it. The sensor needs to be calibrated because its response varies depending on the way it is fixed on the tribometer arm. This process consists in applying a set of known static loads and analysing the response of the sensor. It is performed just before and just after each test in order to ensure its linear response and the absence of drift during the test.

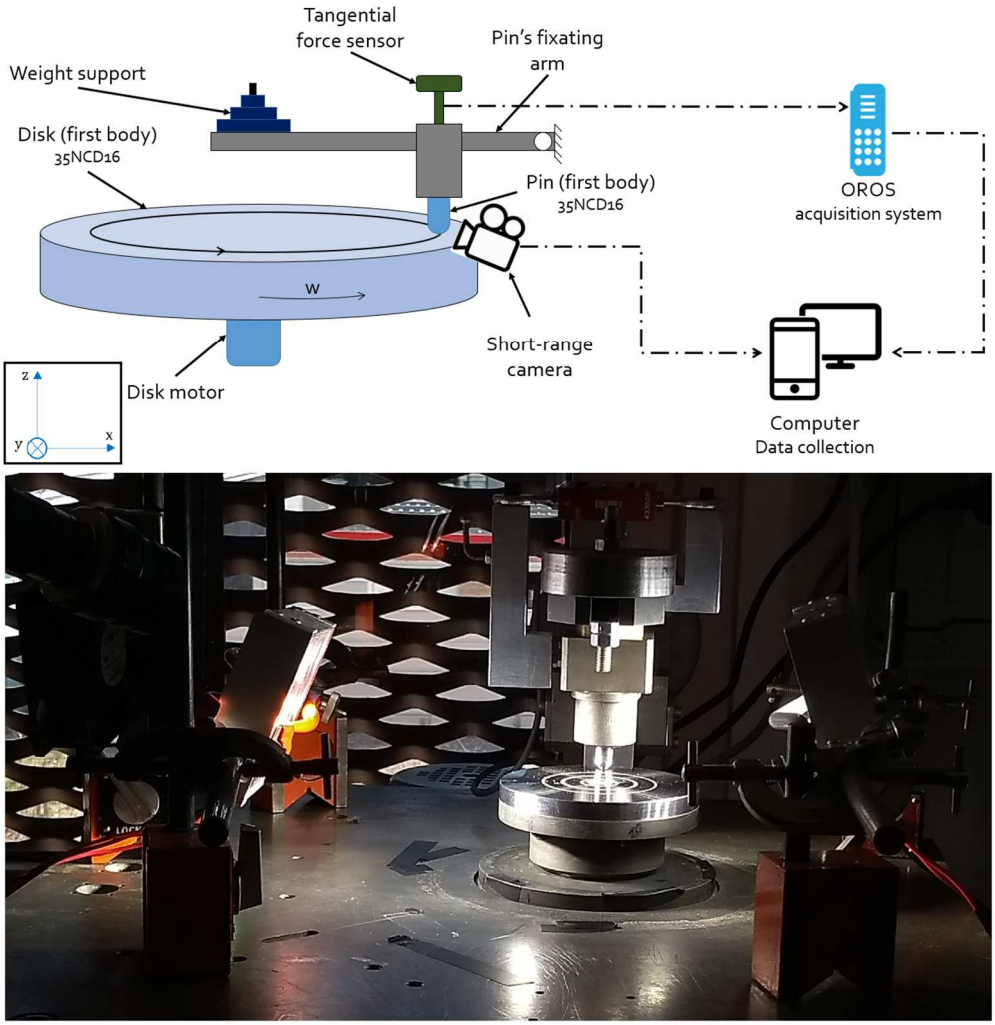

Fig. 1. Experimental apparatus

For the pin and the disc, a low-carbon alloy steel (NF $35NCD16$ steel, or $35NiCrMo16$ steel according to the DIN) is employed. Its hardness is close to 50 HRC. It is highly used in the manufacturing of highly stressed parts regardless of their dimensions across different industrial fields, thanks to its high hardenability and important impact strength level. In the present case, it was chosen for its ability to generate a sufficient amount of wear debris. The dimensions of both the pin and the disk are imposed by the tribometer. The disk has a 109 mm diameter and a 10 mm height, allowing it to be scanned in the Scanning Electron Microscope (SEM Thermofischer Quanta 600) without cutting it. Whereas the pin is a cylinder that has a 5 mm radius and a 25 mm height, with a hemisphere-like tip that measures 5 mm of radius. Both the

pin and the disk have a surface roughness equal to R_a=0.2 µm. For each test, a new pin and a new track on a given disc are used. Before the test, the surfaces of both the pin and disk are cleaned (brushing, water bath, drying, acetone ultrasonic bath, drying, ethanol ultrasonic bath, drying) in order to remove any oil-like residue or particles resulting from the manufacturing process or the handling that may have an effect on the sliding test.

The utility of using *in situ* real time visualization of the tribological test has been proven in the observation and the quantification of the wear debris [34-36], especially with cameras evolving and their specifications getting better. Primarily, those recordings can help deduce the rheological properties of the third body particles, such as the ductility and the cohesion of those particles [37]. In addition, the *in situ* recordings can help explain the micro events that may appear in the friction signals since they follow the contact in real time. The test bench is therefore equipped with two high-speed cameras (GO 5000) with CMOS sensors, a 2560*2048 resolution, and a frame rate up to 62 fps. The cameras are equipped with a 35mm lens and a variety of extension tubes to enhance the focal distance and to get the best recordings possible. Additional light sources are important in order to reduce the exposure time and to pick up sharp details of the object. The lightings mounted are two XL2150 from Spectrum Illumination. They are ultra-high power LEDs producing 559 lm per LED of pure white light. The lights can be moved and their angle can be changed to ensure that the coverage area is the same as the contact area. The cameras can be placed to follow different spots with the help of their supporting fixing arms. The inlet of the contact, the outlet, and the sliding track (away from the contact) can be followed, as well as the contact perpendicularly when using a transparent first body. In the present study, it was decided to supervise only the outlet of the contact between the pin and the disk in order to follow the behaviour of the third body particles.

The applied force is kept constant equal to 10 N, which means applying a static weight of 600 g, in order not to reach the tangential force sensor limits. The shape of the first bodies and the normal load ensure a point contact in the initial state of the experiment with an initial Hertzian maximum pressure equal to 977 MPa at the contact point. The initial average Hertzian pressure in the contact area is equal to 651 MPa. Rotational speed values are in the range 5-6 rpm. The linear speed can have a wider range since it depends also on the radius of the sliding track.

It has been proven that interactions take place between the materials in contact and the environment of the experiment. In fact, many studies confirm that chemical reactions due to gas consumption can change the way materials behave when submitted to tribological loadings [38-41]. Therefore, the environment is considered as an additional variable to be changed in order to vary the morphological measurements and the rheological data. Due to the location of the pin and the disk, it is impossible to change the gaseous environment of the contact without covering the entire top part of the tribometer (the supporting arm for the pin, the weights for applying the normal load, the force sensor and the counterweights). The designed solution is a cube to cover those parts. It is made of Poly (Methyl Methacrylate), i.e. PMMA, to maintain the possibility of filming the contact with the cameras being out. The gas chosen for this project are ambient air and dry argon (<2% relative humidity). The latter is wildly used with low-alloyed steel, especially in welding applications. It was chosen here for its chemical neutrality when exposed to steel. A low gas flow is maintained throughout the test to compensate for the leaks. The advantage of working with argon is that it is heavier than air so the contact is guaranteed to be immersed in the gas since it is at the bottom of the cabin. The temperature and humidity are not imposed in this project but they are recorded during each test. In this work, nine different tests were conducted to compare the effect of changing the different controllable conditions on the morphological and rheological measurements. The aim of the study is not to build a method to

automatically build a tribological scenario based solely on post-mortem views of the third body layer (although this can be considered as a long-term purpose of this preliminary work). As such, the nine different tests are chosen in order to generate a sufficiently large database of rheological and morphological information, but no conclusion is expected at that stage on the effect of the varied parameters (slip velocity, gaseous environment, slip distance) on the tribological scenario. Table 1 summarizes the different tests that were carried out.

Set	**Test**	**F(N)**	**R(mm)**	**ω(*rpm*)**	**V(mm/s)**	**Gas**	**L(m)**	**T(°C)**	**H(%)**
Set 1 Influence of V	1	10	20	6	**12.57**	Air	25	22	51
	2	10	15	6	**9.42**	Air	24	20	49
	3	10	10	6	**6.28**	Air	25	23	50
Set 2 Influence of gas	4	10	20	5	10.47	**Air**	18	22	44
	5	10	20	5	10.47	**Argon**	15	20	50
	6	10	20	5	10.47	**Argon**	17	23	50
Set 3 Influence of L	7	10	10	6	6.28	Air	**25**	27	42
	8	10	10	6	6.28	Air	**22**	23	49
	9	10	10	6	6.28	Air	**15**	21	51

Table 1. List of the experimental tests. F is the normal force, R is the distance of the contact to the center of the disc, ω is the angular speed, V is the sliding velocity, L is the final sliding distance, T is the temperature, H is the relative humidity.

III. RHEOLOGICAL DATABASE

In Fig. 2, illustrative experimental results obtained during test 4 are gathered. Fig. 2A shows the friction signal as well as discrete values of the average friction for each disc revolution. Three main stages can be inferred from this curve. For a sliding distance of 0-1 m, a rapid rise of the friction takes place, followed by a slower friction increase in a second stage (From 1 m to 10 m of sliding), accompanied by intense friction fluctuations and sharp peaks of tangential resistance to sliding. In a third stage (from 10 m to 18 m of sliding distance), friction appears to reach a plateau, with a continuously decreasing variability, and it can be considered that the contact has reached a steady state at the end of the test. These three stages are observed in the 9 tests reported here, with minor differences. Fig. 2B shows a zoomed view of the friction signal in the last meter of sliding, and exhibits a certain amount of periodicity at each lap, confirming that the local value of the friction coefficient at each location of the disc has reached a stable value. In the remainder of this paper, the transient stages (before 10 m of sliding) are not considered for the acquisition of morphological or rheological data, and the focus is put only on the steady state of sliding. No attempt is made on explaining in detail the tribological scenario during these transient states, because of the lack of sufficient data and because this is not the purpose of this study.

Figs. 2C-E show snapshots of the outlet of the contact taken by one of the cameras during the test 4, at three different times of the test. Fig. 2C is shot at the end of the first stage of the friction curve, and does not show any apparent degradation of the disc, while Fig. 2D (shot during the second stage of the friction curve) shows the appearance of a clear wear track. In Fig. 2E (taken in friction stage 3), the wear track is clearly established and has reached a mature state. This is confirmed by the SEM views taken at the end of the test and provided in Figs. 2F and 2G, for the pin and the disc respectively. The pin exhibits a central roughly-circular contact area with

elongated lighter and darker areas oriented in the sliding direction. Inlet and outlet areas are located just before and after the contact area, and the three areas are surrounded by a broad ejection zone full of wear debris. In several cases, a large third body agglomerate (called third body prow, Fig. 2F) is observed just before the inlet zone. The wear track on the disc is structured in a central contact area (usually about 1mm wide), two transition zones around it, and two ejection zones around them. Those main visual post-mortem attributes are observed for all the nine tests performed in this project.

A close observation of the videos acquired during the test reveals that the significant friction peaks observed in Fig. 2A correlate with large (i.e. several tenth of micrometers) off-track and vertical motion of the pin, indicating a temporary disturbance of the local contact conditions. The most likely explanation for this observation (which holds true for all the performed tests) is that the third body accumulates in front of the pin (in the so-called third body prow), and that this agglomerate eventually breaks down and passes under the pin. The disturbance created by this sudden large amount of third body spread on the wear track lasts long enough to be measurable on the cycle-averaged friction signal (blue squares in Fig. 2A), and sometimes during several cycles (green square). It is interesting to note (i) that such peaks seem to follow a certain periodicity, maybe related to a critical failure size of the third body prow, and (ii) that after the peaks the friction signal seems to return to its previous state, indicating that a certain memory of the contact conditions is kept despite this temporary disturbance. However, provided that the prow is located outside of the main contact area (Fig. 2F), it does not carry any significant load and is not responsible for any significant part of the friction force. It will thus be disregarded in the morphological database reported in the next section.

From these raw data, several descriptors are extracted in order to build the rheological database for further use in a Machine Learning context: Average friction coefficient, Standard deviation of friction coefficient, size of the contact area on the pin, width of the wear track, etc. However, for the sake of simplicity, only results related to the first one will be reported here. This choice is made because similar qualitative results are obtained for all descriptors, as will be elaborated later in the present paper. For the computation of the average friction, the third stage of the friction curve (taken after a sliding distance of 10 m) is considered for all tests. Results are reported in Table 2. They show a large variability in the quantitative results but do not seem to follow any clear trend, possibly revealing an important experimental variability.

Test	1	2	3	4	5	6	7	8	9
F	0.44	0.54	0.49	0.26	0.43	0.35	0.29	0.38	0.20

Table 2. Rheological database (restricted to one descriptor in this report); F stands for the average friction coefficient measured during steady state sliding (Stage 3)

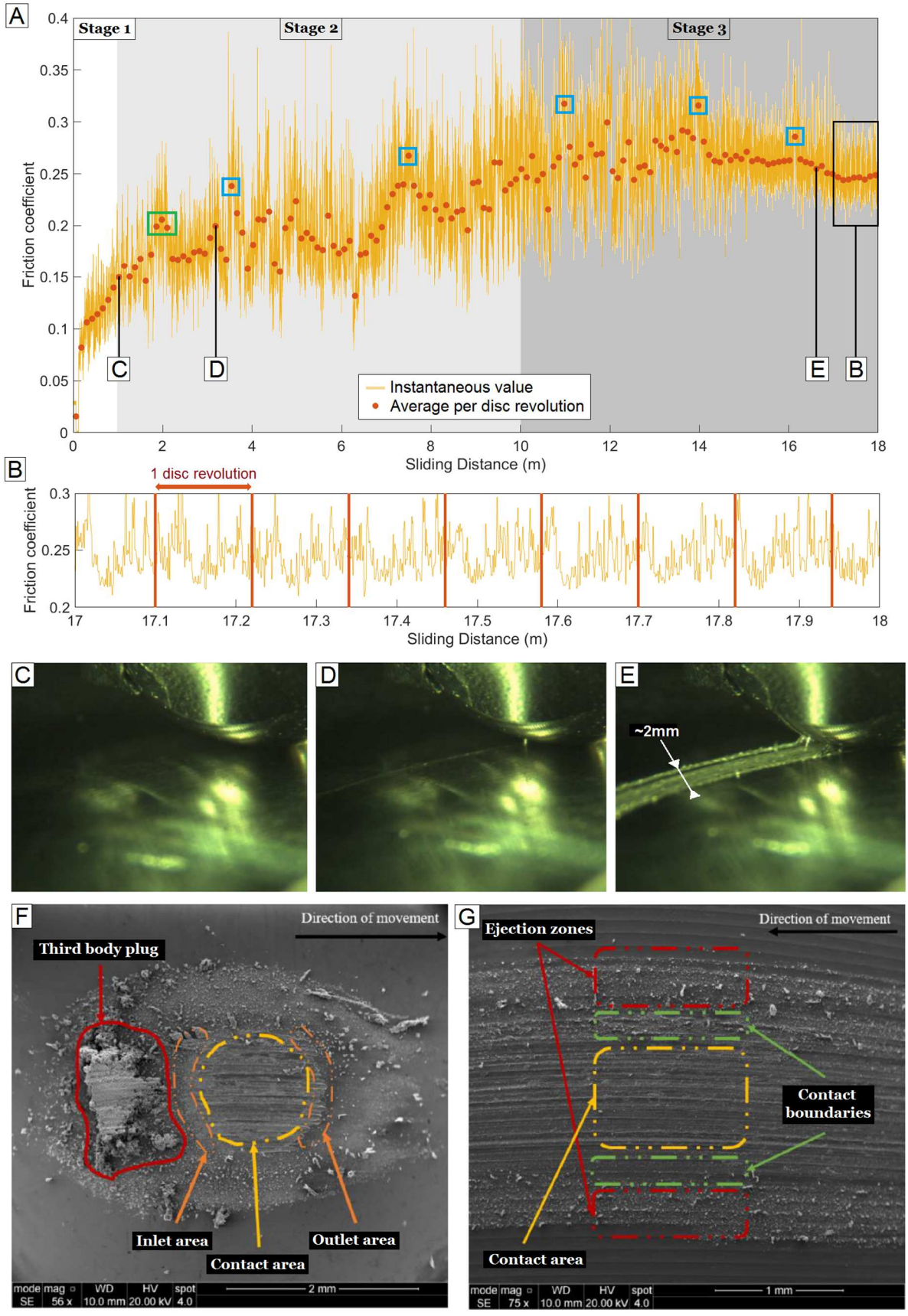

Fig. 2. Tribological measurements and observations for test 4. A. Friction signal; B. Zoom on the friction signal; C.-E. Snapshots of the wear track at the contact outlet during test; F. Post-mortem SEM view of the worn pin; G. Post-mortem SEM view of the worn disc

IV. MORPHOLOGICAL DATABASE

Conventional analysis of the third body particles has traditionally been performed based on the judgment of tribology experts. However, even though such interpretations are precise and accurate, it is not an effective method when having a large quantitative data set. Besides, the knowledge extracted from such a non-quantitative post-test analysis is difficult to transmit to non-experts, and consequently to use in other fields of the scientific or industrial community. Therefore it was decided to develop an image-processing algorithm that describes the third body particles created during a tribological test, using their microscopic images. This is performed by calculating a number of selected morphological descriptors. Using this method, the issue subjective to expert knowledge is expected to be reduced and the repeatability of the tribological analysis is expected to be improved. To limit as much as possible the disturbance of the third body particles at the end of the test, in order not to influence their shape and spatial organization (distribution and localization), the third body particles are left on the surface samples after contact opening. During the visualization procedure, the entirety of the sliding track on the disk and the contact surface on the pin is covered, in order to gather a maximum number of particles. This strategy does not take into account whether a considered particle was active (on the sliding track on the disk or on the contact zone on the pin) or passive (ejected) in the contact. In the case of an ejected particle, no information is available about the ejection time in relation with the contact's lifespan. In order to ensure a good reproducibility, the SEM settings (voltage, working distance, scanning velocity, resolution) are kept constant from one scan to the other, apart from contrast and brightness.

The morphology of an object refers to a description of its appearance, which may include its shape and texture. This study only focuses on shapes but does not take into consideration the texture aspects, although they represent a promising idea for future work. In order to be able to

calculate the morphological descriptors of each particle, a semi-automatic image-processing tool was developed in a Matlab environment. Such an algorithm has two intermediate steps that need to pass by. First, it must differentiate between the different particles present in the image and its background that is, in our case, either one of the surfaces of the pin or the disk or some other much larger particles. This separation is grey-level based. In all the microscopic images acquired during the study, the background is darker than the particles on it as can be seen in Fig. 3. Secondly, the physical edges of each particle (i.e. its geometrical contour) must be detected individually. In order to reach this goal the watershed segmentation method is applied, which is wildly used when studying medical [42-43] and material science [44-46] images. A region-based segmentation separates the entire area into disjoint portions. When reading an image it considers the values of the pixels as heights and transforms it into a topographical image. In a topographical image, there are three types of points:

- Points of minima: if a drop of water is placed on it, it will not move.
- Basins: If a drop of water is placed on it, it will fall to one single and known minima.
- Watershed lines: if a drop of water is placed on it, it can head to more than one minima. Such lines are the borders between two regions that are looked for.

This procedure is already programmed in Matlab. However, it is left to the user to prepare the proper matrix for its input. The raw SEM image needs to be treated before starting the watershed processing. The input image for the watershed function needs to be a grey-level image from where it can detect the appropriate minima for each particle. Fig. 3 illustrates the different transformations that the raw image (A.) goes through and the result of the watershed algorithm. After manual section of a given region of interest (B.), a filter is applied to clean the background noise and remove all the elements that are connected to the borders of the images (C.). Then, the image's histogram is adjusted to increase its contrast (D.). This filter saturates the bottom

1% and the top 1% of all the pixels values. A user-defined grey-level threshold value is used to make the image binary (E.). The resulting image may still contain some defaults referred to as 'holes'. A hole is a region of pixels that does not match its surroundings. A filter is used to correct those defaults (F.). At this point, a binary image is obtained, where the particles are in white and the background is in black. Only now can the watershed segmentation function be applied based on the Distance Transform Approach. This approach is regularly used along the watershed function [47-48]. It calculates simply the distance from any pixel to its nearest nonzero valued pixel. In a binary picture, the distance transform calculates the distance of every black pixel to its nearest white pixel. A white pixel will be replaced in the distance transform result by a zero valued pixel since its nearest white pixel is itself. The distance transform will result into transforming the binary image into a grey-level image (G.). The algorithm can then proceed to apply the watershed segmentation function (H.).

After these 7 steps three outcomes are possible: a correct segmentation, an over-segmentation (several regions are detected in a given particle) and an under-segmentation (several particles are merged in a given region). Human intervention is generally necessary at this stage, either to change the grey-level threshold and restart the segmentation (in the case of under-segmentation) or to merge several regions in order to form a single particle (in case of over-segmentation). This process leads to a correct labelling of several particles with clearly-defined contours (I.). Since the particles are not to be moved from where they lay at the end of the tests, the particles may set on each other or overlap. In this case, those particles are not taken into account.

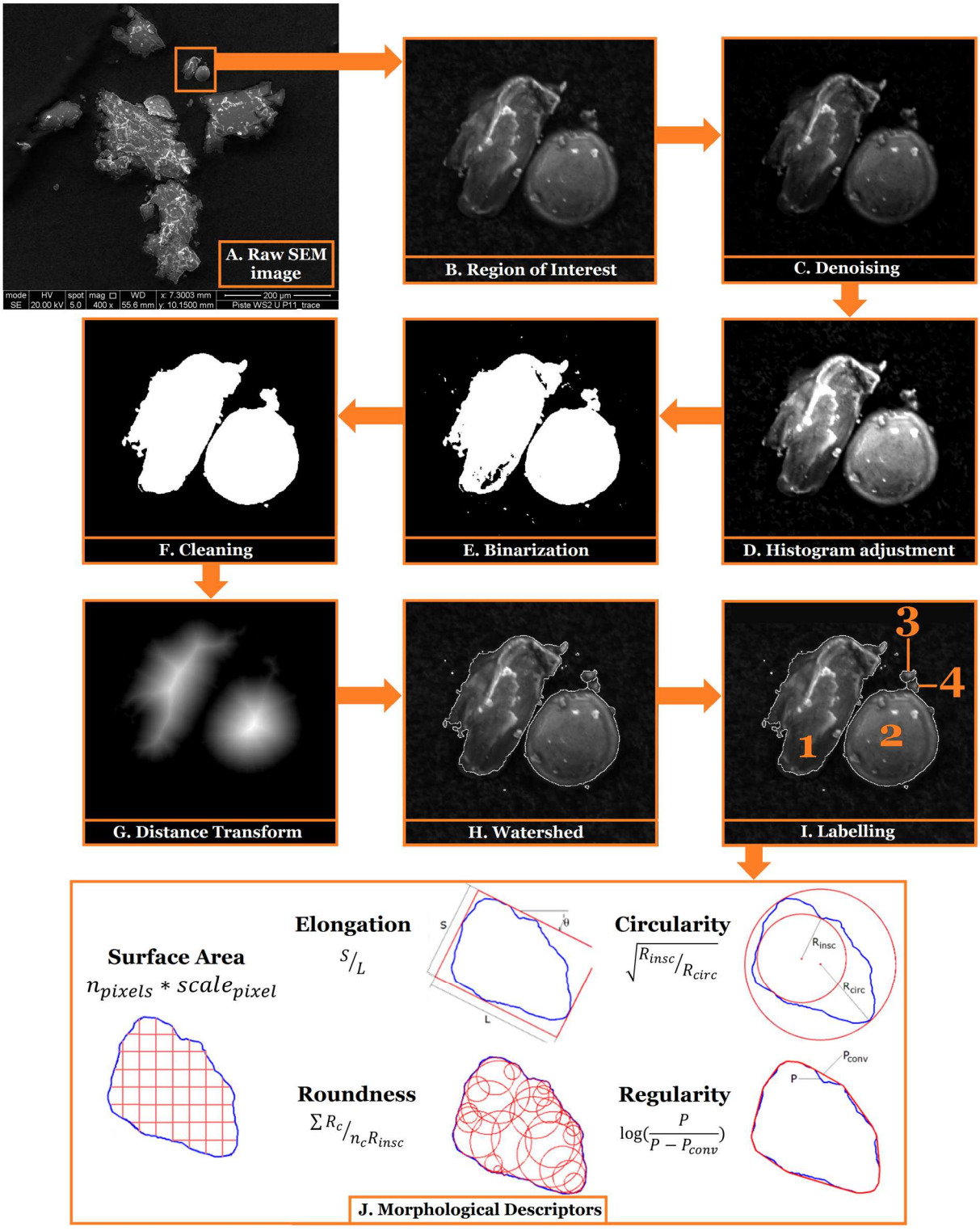

Fig. 3. Morphological extraction. A.-I. Image processing workflow; J. Extraction of morphological descriptors; S and L are the small and large dimensions of a rectangular enclosing box (S being minimized with respect to an angle θ), R_{insc} and R_{circ} are the inscribed and circumscribed radii of the particle, the scalars R_c are the radii of a number n_c of discs fitted on the particle contour, and P and P_{conv} are its perimeter and convex perimeter.

Many studies were conducted to determine the best descriptors to characterize the shapes of particles, especially in geological contexts (see for example a review in [49] and references therein). However, the number of selected descriptors should not overcomplicate the study and yet still give an acceptable idea about the shape of the particles. Five important aspects of the shape are considered crucial to judge the form of a particle, i.e., the projected surface area, the elongation, the circularity, the roundness, and the regularity [50]. They characterize morphological properties of the particle in a hierarchical way, focusing successively on several scales (from the largest to the smallest). These descriptors are shown in Fig. 3J:

-Area: this is the value of the apparent 2D surface. Mathematically, this descriptor is calculated by multiplying the area of a single pixel by the number of pixels contained within the detected contour. The surface area which is measured can be comparable to the real surface projected on a 2D plane orthogonal to the SEM electron beam axis. It does not provide a complete view of the 3D shape of the particle [51].

- Elongation: this descriptor defines the aspect ratio of the particle. The width of the particle in this study is calculated by minimization of an enclosing box with respect to an angle of rotation θ. The closer the elongation measurement is to one, the less elongated the particle is.

- Circularity: It calculates how similar the shape of the particle is to a circle by dividing the radius of the largest inscribed circle R_{insc} by the radius of the circumscribed circle R_{circ}. The closer the particle is to a circle, the closer to one the circularity factor.

- Roundness: It refers to the presence or the absence of sharpness in the particles edges. In the equation, the numbers R_c are the radii of a large number of circles that fit inside the particles and touch its outline at the same time (n_c is their number).

- Regularity: This term indicates the presence of either projections or indentations on the particle's surface. In this equation, P is the perimeter of the particle and P_{conv} is its convex perimeter. The more regular the particle, the larger the value of the regularity descriptor.

This procedure is applied on 1318 particles taken from the 9 tribological tests, and the results are summarized in Fig. 4. This figure shows the average value and the standard deviation of each of the five morphological descriptors for each of the 9 tests. It appears that elongation and circularity do not exhibit a large variability between the tests compared to the inter-particle variability, which is considerable. In contrast, surface area, roundness and regularity show more clear differences between the tests. Illustrative particles are provided in each case and indicate the meaning of each descriptor. It is evident that some descriptors are redundant in some way (for example, particles with large roundness are likely to have a large regularity, and particles with a small elongation are likely to have a small circularity), but still keep some intrinsic differences. There is also a clear correlation between the particle size, on one hand, and its roundness and its regularity, on the other hand. Small particles are more likely to be detected as regular and round when compared to large ones. It appears that, among those that were varied, the gaseous environment is the experimental parameter which is the most influential on the morphological properties of the particles. Argon (tests 5 and 6) leads to smaller, more rounded and more regular particles than room air (test 4). A more subtle trend can be observed on tests 7 to 9, where an increased duration of the test seems to have a similar effect (i.e. to reduce the size and to increase the roundness and the regularity of the particles). This trend is small when compared to the error bars, but could be interpreted as an evolution of the third body during the test.

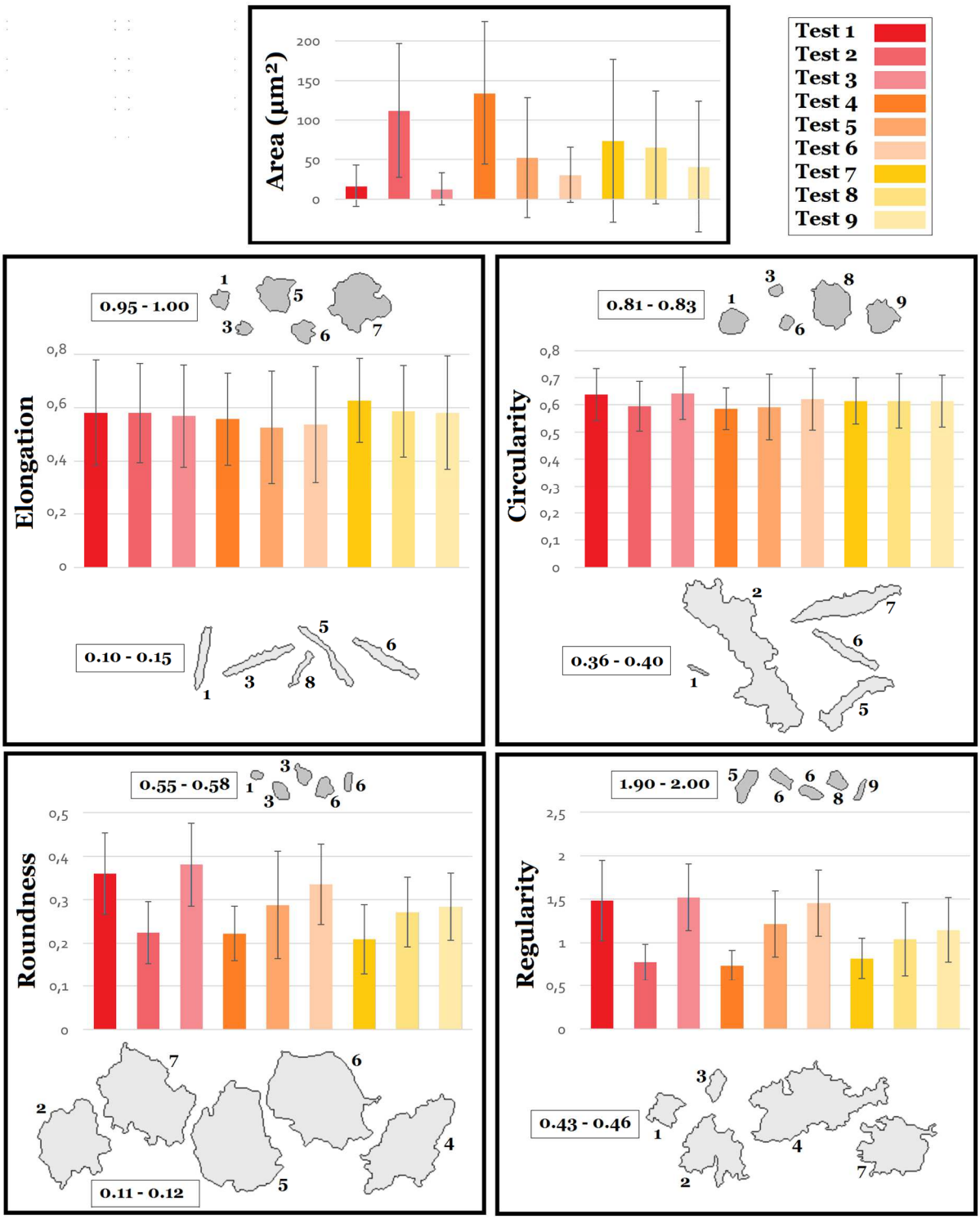

Fig. 4. Morphological database. Average value and standard deviation of the five morphological descriptors for each test (each bar corresponds to the statistics of particles taken from a given test, as detailed in the provided colour bar). Illustrative particles are provided in each case for high and low values of the descriptors, with the corresponding descriptor ranges. The small numbers near each illustrative particle indicate the test from which it originates, and the relative scales of the particles on their respective SEM images are preserved.

V. MACHINE LEARNING

1. Principles

Establishing a logical link between the rheology of the third body and its morphology would enable to enrich and strengthen existing tribological methods, and is therefore highly desirable. However, the size and the complexity of the databases make it almost impossible to achieve this goal manually, and an automated methodology is necessary. This is where machine learning comes into play.

Several studies have used machine learning in a tribological framework for various purposes, like for example to optimize metal composites for a better tribological behaviour of their surfaces [52] and to diagnose mechanical defaults through the vibration signal of a machine [53]. Decost and co-workers [54] worked on classifying materials from their microstructural features. However, the attempts to relate the morphological descriptors of the third body to its rheological parameters are not as numerous. Research reported in [55] aimed to diagnose the state of the mechanical contacts from the third body's morphological state. But the classification put in place did not take into consideration the rheological conditions of the creation of the third body and, as for morphological analysis of wear particles [20-21], the particles were transported away from the contact by a lubricant behaving as a carrier fluid. In a topic close to tribology, machine learning has also been proposed as a way to predict laboratory earthquakes from acquired seismic signals [56].

There are many methods to teach machines how to learn from their experiences. One of the main differences between those techniques is the feedback. The type of the feedback determines the learning mode. In the supervised learning, the feedback matches the action the agent chooses to apply depending on the set of percepts it receives. Therefore, the environment becomes the supervisor of the agent that corrects its errors [57]. However, in the unsupervised learning, the

agent does not receive any feedback. The agent starts with a collection of unidentified data and it builds its own strategy to describe this data by detecting similarities on its own without any interference from the outside world [58]. This study is restricted to supervised learning. The basis of supervised learning is the idea of learning from examples. In this type of machine learning, two main sets of data are used: Z and $Z*$, with Z the training set and $Z*$ the test set of respectively N and $N*$ elements of input-output pairs. The supervised learning algorithms take the training set Z={(X1;Y1),(X2;Y2),… ,(Xn;Yn)}, where Yi=f(Xi) and the function f is the unknown process to be predicted, and discover a resemblance function $\boldsymbol{h}$ to predict the unknown function $\boldsymbol{f}$. The learning process is a constant search for all the possible improvement implemented to the resemblance function $\boldsymbol{h}$. The efficiency of h is judged on the basis of its good performance on the test set $Z*$. The function $\boldsymbol{h}$, also called the hypothesis function, is said to generalize well if it predicts Y with the smallest error possible for the X in the test set. It is thus a very elaborate and multi-dimensional curve-fitting technique. There are two different types of problems in the supervised learning; Classification and Regression. When the output set Y is a number, the problem is called regression. However, the problem is called classification if the output set is a restricted series of values.

2. Classification problem

As a first step, the ability of a Machine Learning Algorithm to perform a classification task is investigated. The aim is to study the relationship between the shape of the particles and the rheological conditions of the tribological tests. The algorithm is therefore asked to associate the particle (input) to a test (output) where it might have been created. Hence, the focus is put on two typical classification algorithms (known also as classifiers): logistic regression and neural networks.

The logistic regression was developed as a statistical model in the late 1960s to study binary data (its units take only two possible values) especially in the medical field [59]. In spite of its label, logistic regression is used to solve classification problem [60-61] such as the text classification [62] and co-reference resolution [63]. By calculating probabilities using logistic (Sigmoid) equation, it measures the relationship between an output class and one or more independent input variables. Given a set of input parameters $X = [X_1, \dots, X_n]$ and a set of weights $\theta = [\theta_1, \dots, \theta_n]$, the hypothesis function for a logistic regression classifier is as follows:

$$h_\theta(X) = s(\theta \cdot X) = \frac{1}{1 - e^{-\theta \cdot X}} \quad (1)$$

In the presented study, the input parameters X_j are the five morphological descriptors of a given grain, and $h_{n\theta}(X)$ is the estimated probability for this grain to originate from a given tribological test Y_i. The dot · denotes scalar product. The weights θ_j are chosen to minimize a cost function expressed by:

$$J(\theta) = -\frac{1}{m}\sum_{i=1}^{m}\left[Y_i \cdot log\left(h_\theta(X_i)\right) + (1 - Y_i) \cdot log\left(1 - h_\theta(X_i)\right)\right] + \lambda \cdot \frac{1}{n}\sum_{j=1}^{n}\left[\theta_j^2\right] \quad (2)$$

The first term of this expression is called a cross-entropy function, and it evaluates the error made by the function h_θ when trying to predict a given output Y_i from a given input X_i, averaged on the m elements composing the database. The second term is called a Tikhonov regularization term [64], and aims at limiting the risk of over-fitting (i.e. perfect results of the algorithm on the database, but failure to generalize it to new cases). This minimization is performed by a Gradient Descent Algorithm [65], performed during a number of N descent steps. The whole approach hence depends on two purely numerical parameters: the regularization parameter λ, and the number of iterations N_i.

The artificial neural network (ANN) is a system of interconnected 'neurons'. Those connections have weights that are adjusted during the training phase of the learning. Usually the network has multiple layers of interconnected neurons. Each layer responds to a different combination of inputs that is the output of the layer before it. ANN are implemented in different domains such as handwritten digits recognition, the diagnosis of disease conditions [66], and face detection algorithms. The Input vector (the morphological descriptors in our case) goes through hidden layers, each of which being a set of values calculated the same way as the hypothesis function in the logistic regression. The last layer contains only one node, which is the output of the network. The analytical expression of the hypothesis function can thus be cumbersome in the case of a large number of layers and of nodes per layer. The number of hidden layers has been a subject of many researches to enhance the efficiency of the neural network algorithms. It was proven that a single hidden layer network with a finite number of nodes is capable of approximating "any function that maps one finite space to another" [67-68]. Therefore, the number of hidden layers is fixed to 1 in this preliminary study. The cost function has the same structure as that used in logistic regression. Its minimization is based on the classical backpropagation technique [69], with a number of 500 iterations. There are two numerical parameters: the regularization parameter λ, and the number of nodes in the hidden layer N_n. A scheme of the Neural Network is proposed in Fig. 5.

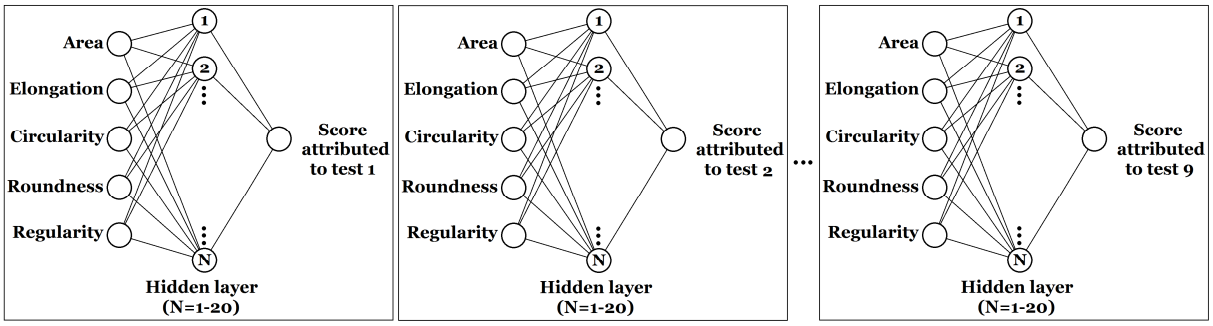

Fig. 5. Neural Network used for the classification problem

The results are presented for both classifiers. The morphological database is randomly divided into two parts: a training database (composed by 75% of the third body particles, and used to calibrate the weights of the Machine Learning algorithms), and a test database (composed by the remaining 25%, and used to evaluate the ability of the calibrated algorithms to predict accurately unknown data). In each case, a success rate is established, defined as the proportion of the third body particles that were successfully attributed to the test from which they actually originate. The results are gathered in Fig. 6, and show that increasing λ parameter does not always mean having better performing algorithms.

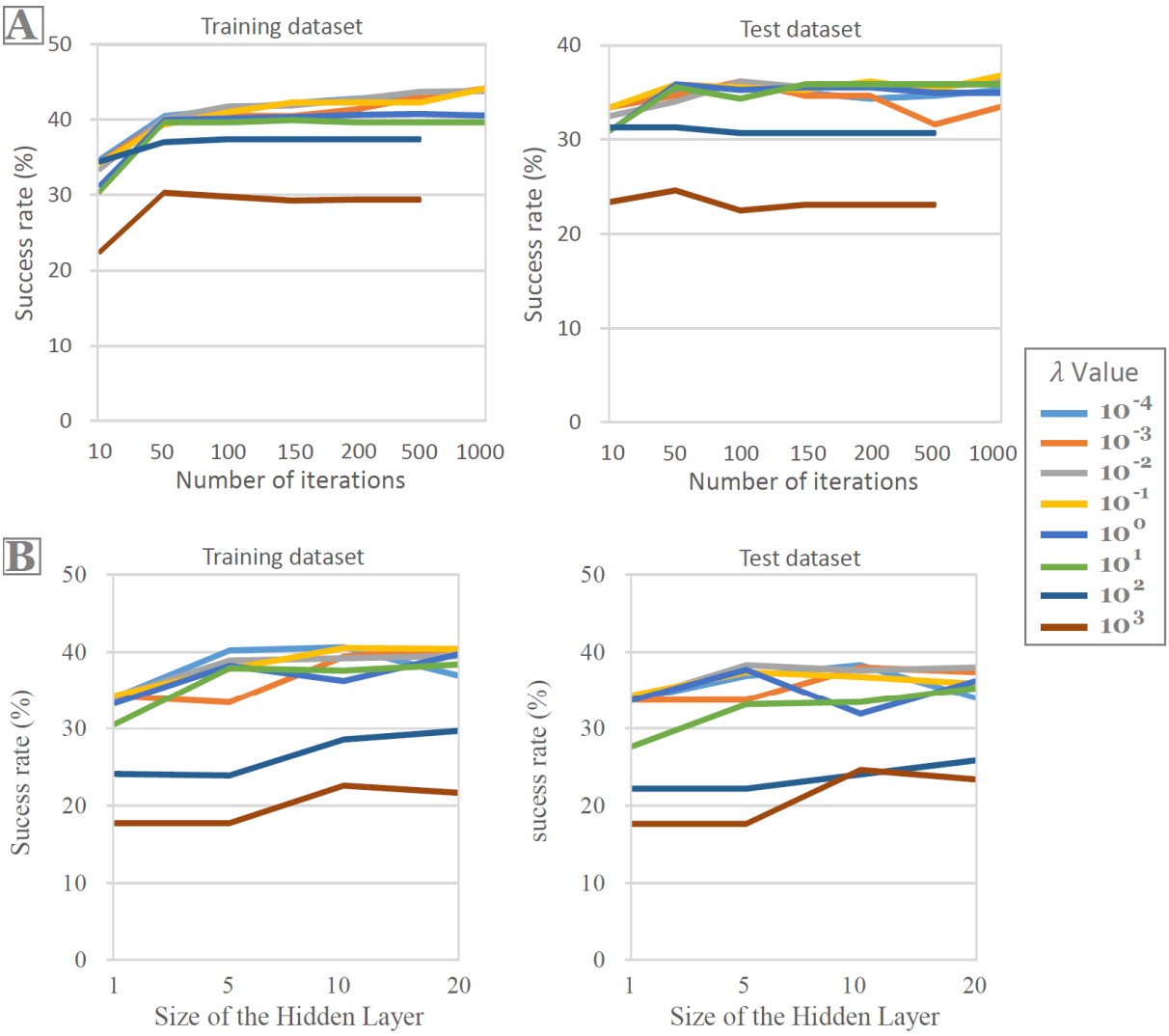

Fig. 6. Success rates of the classification algorithms. A. Logistic regression; B. Artificial Neural Network

In the case of the logistic regression, the success rate functions hit a plateau when the number of iterations N_i increases. The combination [λ=0.01; N_i=500] guarantees the best success rates during both the training phase (44%) and the test phase (35%). When using Artificial Neural Networks, increasing λ makes the algorithm better until a certain threshold is reached. For λ=1 or λ=100 the success rate of the algorithm is much lower than when λ is smaller. However, when changing the size of the hidden layer, the success rate of the algorithm hits a plateau at a number of 5 nodes. For the combination that ensures the highest success rate on the test database, which is [λ=0.01; N_n=5], the network results into having 39 % and 38 % as success rates when applied on the training and the test databases respectively.

Fig.7 shows the plot of the tests attributed by the logistic regression algorithm to the particles in the training database. More specifically, each column corresponds to the particles originating from a given test, and the colours in this column indicate what test is attributed to these particles by the machine-learning algorithm. Results obtained for the test database and results obtained for both databases by the neural network algorithm are very similar and are not shown here. It is evident from this figure that the machine learning predictions are not equally good for all the tests. Tests 2, 3, and 6 lead to rather good predictions (more than 50% of success), tests 1 and 7 lead to mildly satisfactory predictions (between 30% and 50% of success), and tests 4, 5, 8 and 9 are very poorly predicted (less than 10% of success). It is interesting to note that a majority of particles from test 1 are labelled as belonging to test 3, indicating that the algorithm has difficulties to discriminate properly between these two tests. Likewise, particles from test 5, 8, and 9 are in majority labelled as originating from test 6.

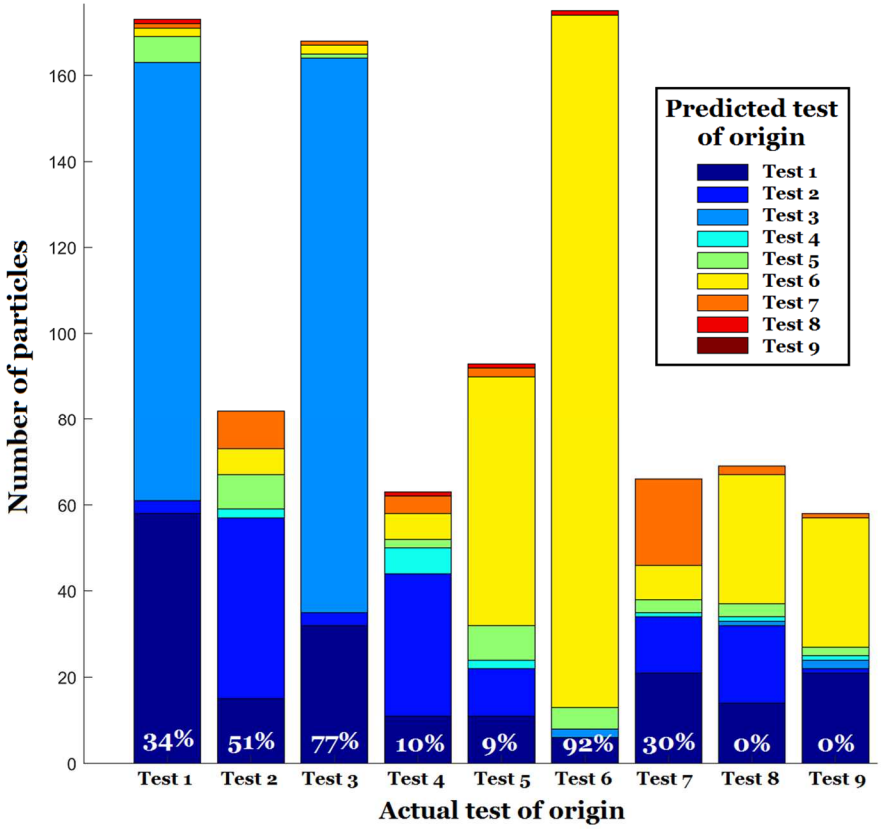

Fig. 7. Predictions of the Logistic Regression Algorithm in the case λ=0.01 and N_i=500, as a function of the actual test of origin of the particles (training dataset only); Percentages in white indicate the success rate per test of origin

A more detailed view is provided in Fig.8, which gives the detailed score attributed to each particle for it originating from each test. This information is hidden in Fig.7 since only the best score is retained for each particle in order to label it with a predicted test of origin. In Fig.8, the particles (on the vertical axis) are first binned by test of origin, and then sorted in each bin by the score that is attributed to them for their actual test of origin. This graph is very instructive, because it shows that particles with a good score in test 1 also have a good score in test 3. Tests 5 and 6 are also very well correlated (which makes sense, since they share the same experimental conditions, i.e. argon atmosphere), as well as tests 2, 4 and 7. Particles from tests 8 and 9 have low scores for all tests. The second column of Fig.8 provides the corresponding morphological descriptors for all particles. The typical "profile" of the particles of some tests appear clearly. For example, particles with high scores in tests 5 and 6 are characterized by a low surface area, a high roundness and a very high regularity. However, particles with high

scores in tests 2, 4 and 7 are characterized by a large surface area, a small roundness and a small regularity. Particles well-identified from tests 1 and 3 tend to have a low surface area, a high roundness, and a low regularity. For all the particles, the values of elongation and circularity do not show any identifiable pattern, and do not seem to discriminate at all between the different tests. This is confirmed by Table 3, which provides the success rates obtained when removing a morphological descriptor from the database. The most discriminating descriptor is, by far, the regularity.

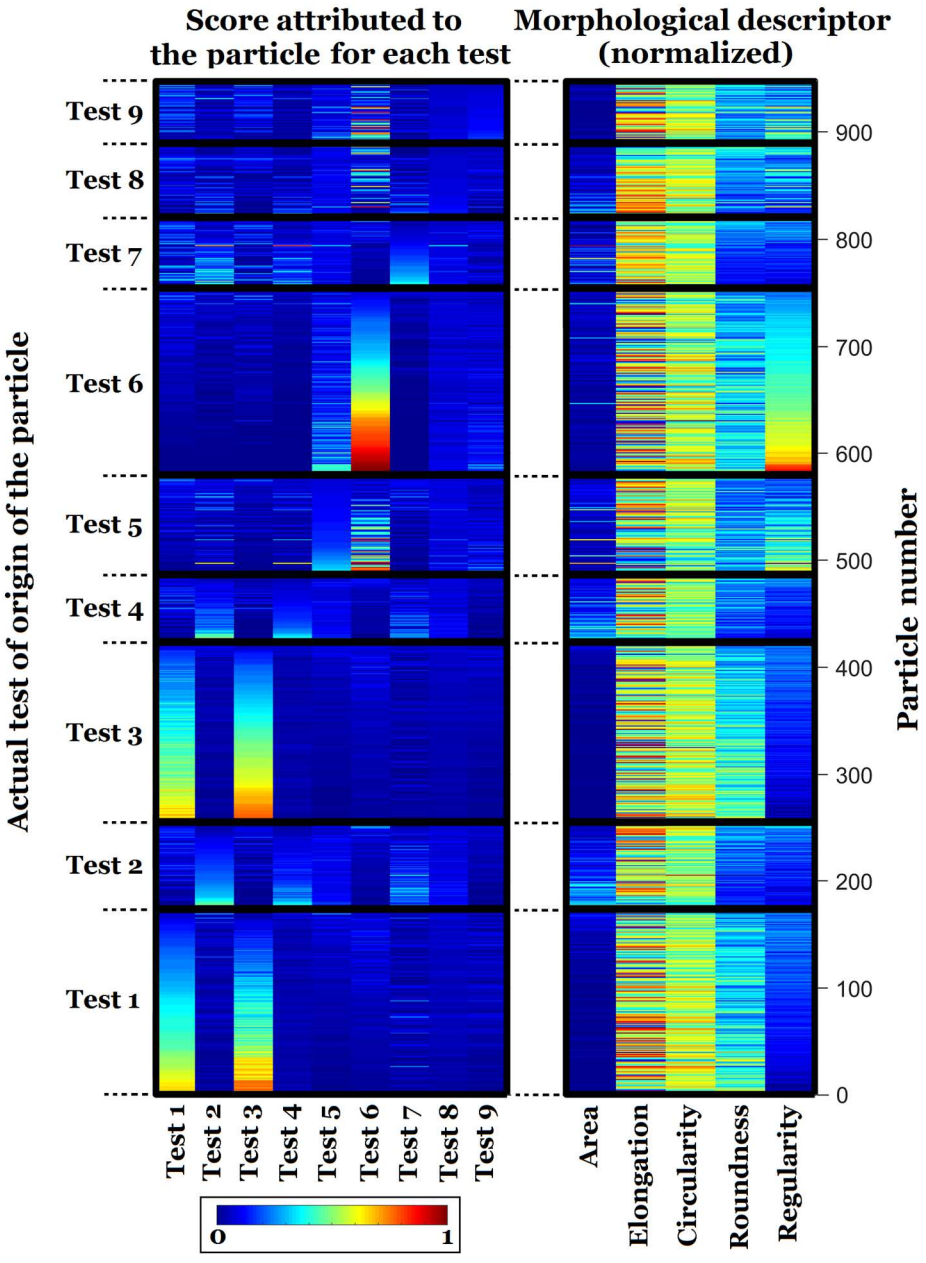

Fig. 8. Score attributed to each particle by the Logistic Regression Algorithm in the case λ=0.01 and N_i=500 (training dataset only), and corresponding morphological descriptors. Particles (on the vertical axis) are first sorted by test of origin, and then by the score obtained for their test of origin; the height of each bin is proportional to the number of particles

	All descriptors	Without surface area	Without elongation	Without circularity	Without roundness	Without regularity
Training dataset	44%	42%	44%	43%	41%	32%
Test dataset	36%	36%	36%	36%	34%	27%

Table 3. Success rates obtained by the Logistic Regression Algorithm in the case λ=0.01 and N_i=500 when considering a truncated dataset, i.e. when removing one morphological descriptor.

The results provided by Logistic Regression and by Neural Network are similar, but a more detailed comparison is provided in Fig.9. The scores attributed to each particle for each test are reported in nine separate graphs, allowing to compare the prediction from both algorithms. Both training and test datasets are included, and the particles originating from each considered test are highlighted in red. In general, the scores provided by both algorithms correlate very well, for both datasets. However, a general pattern seems to characterize each algorithm. For low scores, data points are generally located above the one-to-one slope, indicating that the Neural Network attributes larger scores than Logistic Regression. However, for large scores, the inverse observation can be made. In general, the range of the scores is larger for Logistic regression, indicating that this algorithm seems to discriminate more between different particles. It is striking, for example, that the Neural Network attributes very low scores to absolutely all particles for tests 7, 8, and 9, preventing any particle from these tests to be accurately labelled.

The fact that the scores attributed by both methods are so strongly correlated might also indicate that both of them are adequately extracting all the relevant information from the database, but that their results saturate because of the limited amount of such information. It clearly points

towards the fact that the chosen descriptors fed in the morphological database do not provide a complete picture of the morphology of the grains. In addition to a larger amount of elements, a qualitative enrichment of the database, including different descriptors, is expected to improve the results and reveal differences in performance between the two ML approaches, which in the present case can hardly be ranked.

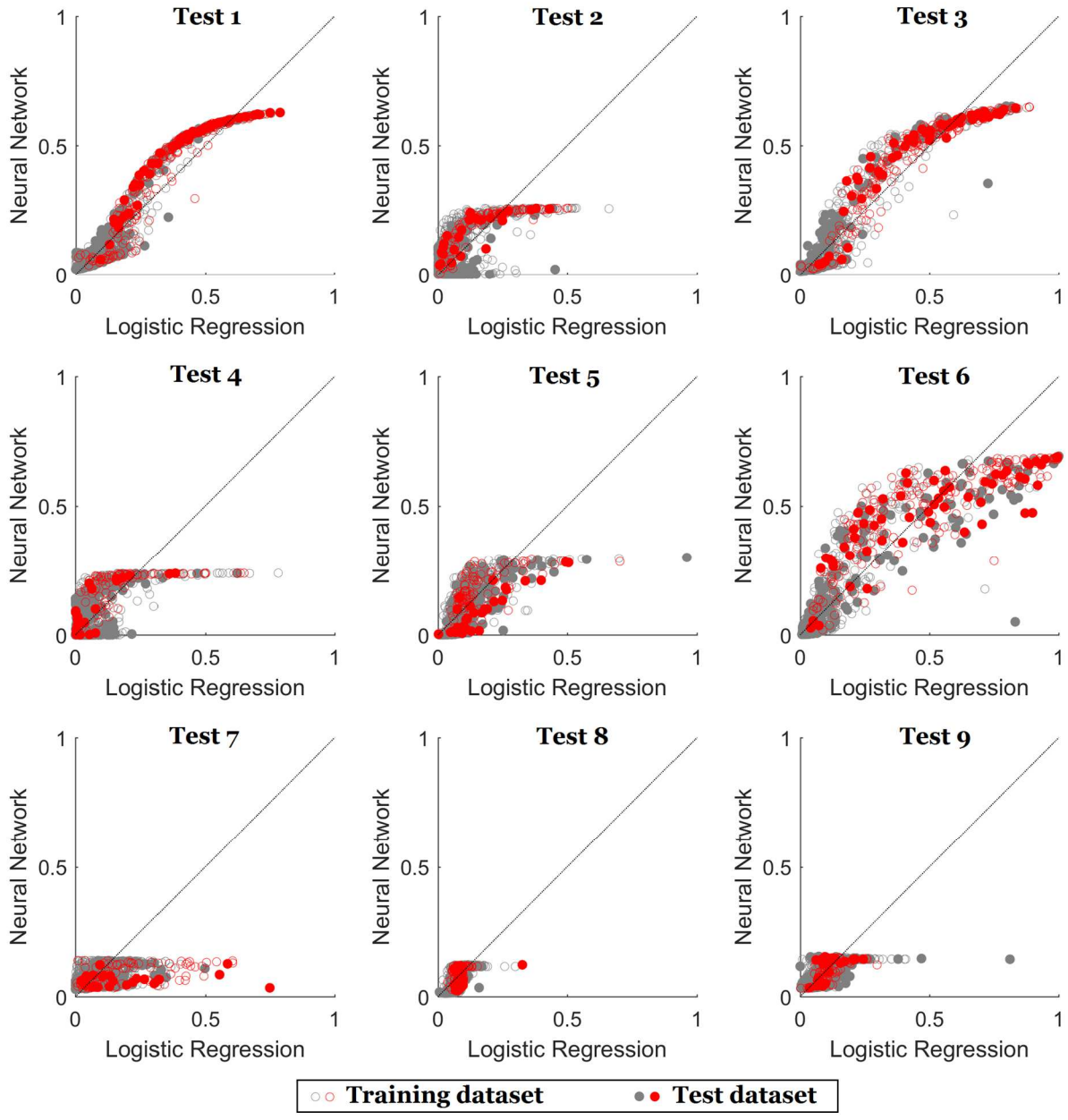

Fig. 9. Comparison of the scores attributed to each particle by the Logistic Regression Algorithm (in the case λ=0.01 and N_i=500) and by the Artificial Neural Network (in the case λ=0.01 and N_n=5) to particles for all tests. Dots in red mark the particles which actually originated from the considered test, dots in grey mark all the other particles.

3. Regression problem

In the previous sub-section, the output of the machine-learning algorithms is the number of the test from where the particle is generated. In this sub-section, as a second step, machine learning capability to predict the rheological measurements from the morphological descriptors of the particles is investigated. This changes the nature of the problem because it is no longer a classification objective but a regression one. More specifically, the algorithms would not predict a class to which the grains may belong to, but a numerical quantity. The database contains all the particles from the tests. As previously, it is divided into two databases: the training (75 %) and the test database (25 %). The input part of the algorithm is the morphological database, and the output is the rheological one (restricted in this preliminary study to the value of the average friction provided in Table 2). The chosen approach is polynomial regression, with the following hypothesis function:

$$h_\theta(X) = \sum_{k=1}^{p} \sum_{j=1}^{n} \theta_{jk} \cdot X_j^k \tag{3}$$

The cost function used in that case has the following expression:

$$J(\theta) = -\frac{1}{m} \sum_{i=1}^{m} \left(Y_i - log\left(h_\theta(X_i)\right) \right)^2 + \lambda \cdot \frac{1}{n \cdot p} \sum_{k=1}^{p} \sum_{j=1}^{n} \theta_{jk}^2 \tag{4}$$

In this expression, Y_i is the coefficient of friction measured in the test of the particle i and the vector X_i contains its morphological descriptors, p is the polynomial order of the method, n is the number of morphological descriptors, and m is the number of particles in the database. The $n \cdot p$ weights are determined by a Gradient Descent Algorithm as in the previous sub-section. To estimate the success of the prediction, the success quantifier is the average absolute error on the friction coefficient (i.e. the average on all the particles of the absolute value of the difference between the predicted and actual friction coefficients).

The results of this approach are gathered in Fig.10. Fig.10A shows the predicted coefficients of friction for all the particles, in the case of a polynomial of order 2 and for two different values of the regularization parameter λ. The prediction appears to be rather poor. It predicts well the friction coefficients for the particles coming from some tests (e.g. tests 3 and 5 for $\lambda = 0.001$ and tests 1, 6 and 8 for $\lambda = 1$), but it is probably due to chance since the prediction for the other tests are not as good. Increasing the value of λ clearly reduces the variance of the predictions, making them closer to the average value of the friction coefficient. As shown in Fig. 10B, the average absolute error (which serves as a success quantifier) remains close to 0.065 for all values of λ (except for $\lambda > 100$) and for all orders of the method. It should be noted that the same method was applied on an extended version of the rheological database (including more rheological descriptors for each of the nine tribotests, not reported here), and led to very similar results.

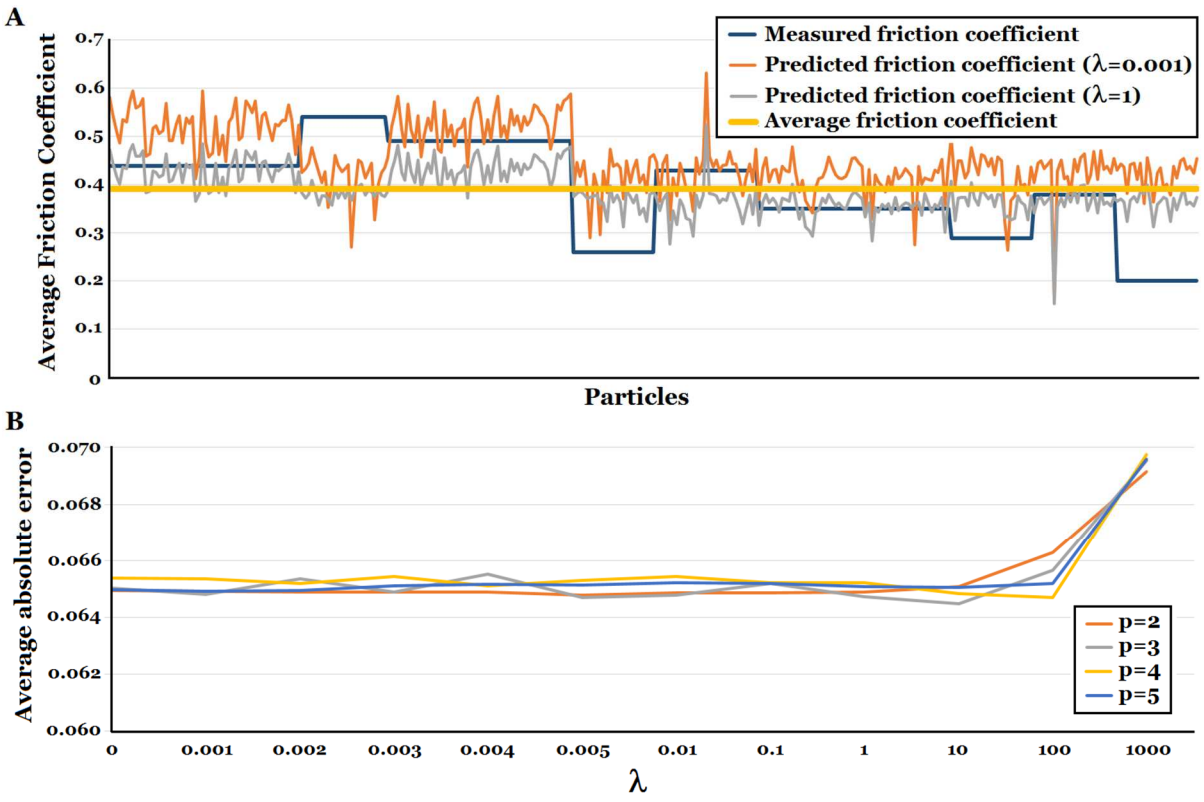

Fig. 10. Results of the regression approach. A. Comparison between the measured and predicted friction coefficient for p=2; B. Average absolute error as a function of the regularization parameter and the polynomial order.

VI. DISCUSSION AND CONCLUSIONS

In the classification problem, the success rates obtained by both methods are modest when compared to similar applications of machine learning techniques reported in the literature. The work of Podsiadlo and Stachowiak 2005 [30] reported for example an impressive success rate of 97% when applying Support Vector Machines (SVM) to the classification of wear particles coming from the same tribotest, i.e. same contact conditions such as speed, load, or environment, but at different test durations. The key of success in this case was an appropriate method of data reduction and the choice to account for particles textures. This success rate, however, was obtained for an optimal choice of the parameters of the Machine Learning algorithm, and could drop below 75% in other cases. Other studies reported good results, such as the work of Stachowiak et al. 2008 [31] and Yuan et al. 2016 [32]. The former reached a success rate of 97% when classifying particles from three different wear mechanisms (fatigue, adhesive and abrasive wear) based only on simple shape descriptors. In that case, it appeared that the three categories could be easily discriminated by such simple properties as surface area and elongation. In most cases it could even have been done manually. Yuan et al. [32] also obtained good results (between 75% and 83% of success) when classifying particles in four rather distinct categories (e.g. spherical, irregular, etc.), using only contour descriptors.

A very interesting case was provided in Wang et al. 2019 [33]. Based only on morphological features extracted from pictures collected using ferrography, they first applied Neural Networks to classify various kinds of particles in five classes. This part of their work was similar to the present work in many aspects, and led to a satisfactory level of prediction. However, detailed analyses showed that the Neural Network was very predictive for three different classes (success rate larger than 80%), but was unsuccessful for two other classes (36% and 45% of success for Severe Sliding and Fatigue Wear, respectively). This is in good agreement with the

results reported here, since it indicates that a given algorithm can give very different results with the same training, depending on the database: some cases are simply more difficult to discriminate based on the available data. In the work by Wang et al. [33], a second pass on the same grains but using a Convolutional Neural Network (CNN) on raw images instead of extracted features restored a success rate larger than 80% for the most critical classes. It indicates that the discriminating information was there, but that it could not be extracted by simple shape features. This raises the following important question: can the visual information allowing to discriminate between different classes of particles be always encoded in cleverly designed morphological descriptors? If the answer is yes, then future challenges will consist in building such descriptors, but if the answer is no, then image-based Deep Learning might be the only viable solution.

Based on the rather scarce literature on the topic, it can be hypothesized that the classification results reported here are less good than expected for two related reasons. The first reason is that the tests are too numerous (9, instead of 3 to 5 reported in [30-33]) and too similar (only minor experimental parameters are modified) to produce particles that could be easily discriminated based on simple descriptors (this can be observed in Fig. 4, where the error bars are larger than the inter-class variability). The second reason is that the geometric information that would be necessary to perform this discrimination is missing in the chosen descriptors. Considering these two reasons, it appears that the choices made in the construction of the database should be proportioned to the difficulty of the task assigned to the trained algorithm. For example, the chosen descriptors might have been sufficient for the dissimilar particles of study [32], but were not for those of the present study. The strategy to build a sufficiently rich database for a given task remains to be investigated. However, despite being less good than some other reported results, the predictions obtained in the present study remain way above a purely random

assessment (it would be one successful prediction over nine). As such, they are equally disappointing and promising. They demonstrate that, in some cases, the morphology of the grains possess a signature that is closely linked to the tribological conditions where they were formed. As such, this work brings a first quantitative validation to the common practice of tribological experts, who have been used for decade to interpret tribological scenarios based on the observation of the particles on worn surfaces. It is striking, for example, to consider the prediction results of test 6. As shown in Fig. 8, the morphological signature of this test is clearly the Regularity descriptor. This is confirmed by Table 3, where the absence of this descriptor leads to a drop of the success rate. It is thus very likely that the lowest success rate of some other tests is related to the fact that some of the visual features characterizing their particles was absent from the morphological database. The choice of considering nine different classes of particles (instead of three to five in similar studies) coming from rather similar tests (same wear type, only varying minor operational parameters) also certainly led to the limited success rates. It can thus be foreseen that future enrichment of the database could only improve the prediction capability of the algorithms. This extended database could include more descriptors of the particles contours, but also of their texture and chemistry, and take advantage of different observation tools in a "multi-mode" framework (optical microscopy, topography measurements, etc.). The idea that contour and texture descriptors are equally necessary to distinguish different classes of particles (contour descriptors for some classes, texture descriptors for some others) was already proposed in several studies [26, 31], and is a promising research avenue. A quantitative enrichment of the database, including a much larger number of particles, could also improve the training results. As demonstrated in [33] however, Deep Learning might be the only solution for the most complicated cases.

The outcome of the regression problem, in contrast, seems to indicate that very little should be expected in terms of prediction of the rheology of a tribosystem based on the morphology of its generated particles, if the methodology used in this work is reproduced as such. There is, however, a simple explanation for this disappointing result. As shown in Fig.2B, the friction coefficient is a measured quantity that strongly varies in space (variation during a given revolution of the tribometer), but also in time (variation from one revolution to the next). It is certainly the same for any other rheological descriptor. This variability has been lost in the chosen methodology, by considering rheological descriptors averaged both in space and time. It makes sense that morphological properties of particles, which are local by definition, should fail to predict a quantity which is averaged on a whole tribotest. This observation means that the results of the regression approach could be considerably improved by considering local rheological parameters (e.g. instantaneous friction coefficient) instead of averaged ones. This promising research avenue will be explored in future studies.

The results obtained in this work did not allow to discriminate between different Machine Learning approaches. In the classification problem, Logistic Regression and Neural Networks led to similar results, albeit with some minor differences (Fig.9), and a parametric study showed that the influence of the methods numerical parameters was not dramatic in the investigated range (Fig.8). This is true as well for the regression approach, where the regularization parameter and the polynomial order did not appear to have sharp effects on the average absolute error (Fig.10). At this point, it seems clear that the priority is the enrichment of the database (by adding morphological descriptor and making the rheological descriptors more local, while removing any sampling bias), more than a search for the most suited Machine Learning approach.

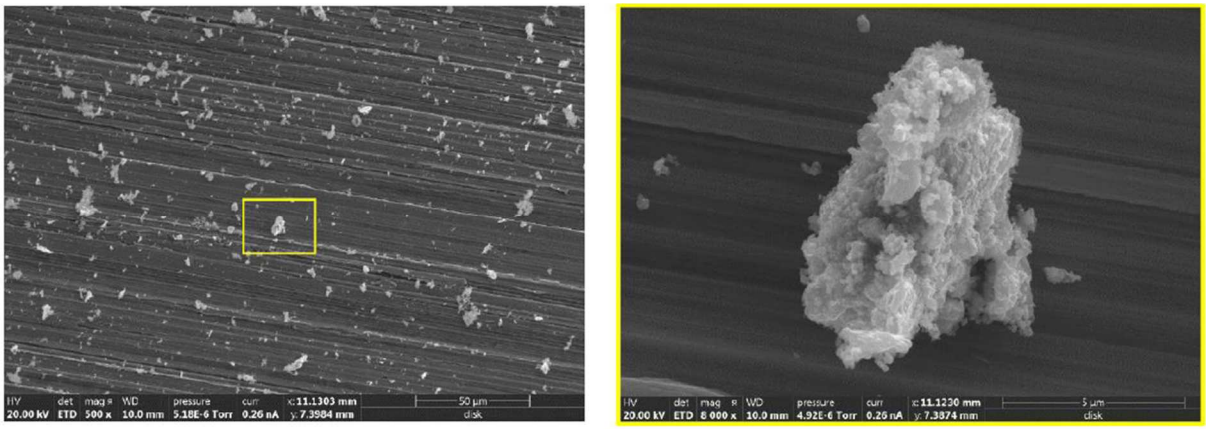

Fig. 11. SEM views of the same third body particle with different magnifications

One important question that seems to arise from this work is illustrated in Fig.11, and is related to the definition of a third body particle. This definition seems, at the moment, extremely subjective. It depends on the personal assessment of the observer, but also on the chosen scale of observation. Fig.11 indeed shows that what would appear as a clearly identified third body particle could also be seen as an agglomerate of a large number of individual particles if sufficiently magnified. At the moment, the expert judgement is thus still needed when building the morphological database. This is a good point if the mission assigned to Machine Learning is to improve our understanding of tribological phenomena, because human eye needs in that case to be part of the process in order to bring some interpretation to the Machine Learning results (connecting them, for example, with numerical simulations). But if Machine Learning is to be assigned the role of a black box making operational decisions (say, by automatically modifying in real time an industrial process based on some changes detected on the surfaces after observation of the third body particles), then the objectivity of the method should be seen as an ultimate goal. In the context of an increasing use of and trust in automated Machine Learning algorithms in countless applications, these two purposes (scientific and operational) might request rather distinct strategies.

DISCLOSURE

The authors acknowledge that this study contains original material, as a result of a purely academic study without any kind of private funding or conflict of interest. Its publication has been approved tacitly by the responsible authorities at the institutes where the work has been carried out.

ACKNOWLEDGMENTS

This work is supported by the LABEX MANUTECH-SISE (ANR-10-LABX-0075) of Université de Lyon, within the program “Investissements d’Avenir” (ANR-11-IDEX-0007) operated by the French National Research Agency (ANR).

REFERENCES

[1]. Godet, M. (1984) 'The third-body approach: A mechanical view of wear', *Wear*. Elsevier, 100(1–3), pp. 437–452

[2]. Lofficial, G. and Berthier, Y. (1987) 'Load carrying in slow reciprocating mechanisms', *Tribology Series*. Elsevier, 12, pp. 281–290

[3]. Dowson, D., University of Leeds. Institute of Tribology. and Institut national des sciences appliquées de Lyon. (1996) *The third body concept : interpretation of tribological phenomena : proceedings of the 22nd Leeds-Lyon Symposium on Tribology, held in the Laboratoire de mécanique des contacts, Institut national des sciences appliquées de Lyon, France, 5th-8th September 1995*. Elsevier.

[4]. Denape, J. (2014) 'Third body concept and wear particle behavior in dry sliding conditions', *Tribological aspects in modern aircraft industry*, 640, pp. 1–12

[5]. Cundall PA, Strack ODL. (1979). "A discrete numerical model for granular assemblies", *Géotechnique*, 29(1), 47–65

[6]. Mollon, G., Richefeu, V., Villard, P., and Daudon, D. (2015). "Discrete modelling of rock avalanches: sensitivity to block and slope geometries", *Granular matter*, 17(5), 645-666

[7]. Seve B, Iordanoff I, Berthier Y. (2001) "A discrete solid third body model: influence of the intergranular forces on the macroscopic behavior", *Tribol Interface Eng Ser,* 39:361–8.

[8]. Renouf M, Fillot N. (2008), "Coupling electrical and mechanical effects in discrete element simulations", *Int J Numer Methods Eng*, 74:238–54.

[9]. Champagne, M., Renouf, M., and Berthier, Y. (2014), "Modeling wear for heterogeneous bi-phasic materials using discrete elements approach", *ASME J Tribol*, 136(2):021603

[10]. Mollon, G. (2015). "A numerical framework for discrete modelling of friction and wear using Voronoi polyhedrons", *Tribology International*, 90, 343-355

[11]. Mollon, G. (2016) 'A multibody meshfree strategy for the simulation of highly deformable granular materials', *International Journal for Numerical Methods in Engineering*. John Wiley and Sons Ltd, 108(12), pp. 1477–1497

[12]. Mollon, G. (2018), "A unified numerical framework for rigid and compliant granular materials", *Computational Particle Mechanics*, 5(4), 517-527

[13]. Mollon, G. (2019), "Solid flow regimes within dry sliding contacts", *Tribology Letters*, 67:120

[14]. Hunt, T. M. (1993) *Handbook of wear debris analysis and particle detection in liquids*. Elsevier Applied Science.

[15]. Surapol, R. and BJ, R. (1995) 'The classification of wear particle shape', *Lubrication Engineering*, 51(5), pp. 432–437.

[16]. Podsiadlo, P. and Stachowiak, G. . (1997) 'Characterization of surface topography of wear particles by SEM stereoscopy', *Wear*. Elsevier, 206(1–2), pp. 39–52.

[17]. Roylance, B. J. and Raadnui, S. (1994) 'The morphological attributes of wear particles — their role in identifying wear mechanisms', *Wear*, 175(1–2), pp. 115–121

[18]. Xu, K., Luxmoore, A. R. and Deravi, F. (1997) 'Comparison of shape features for the classification of wear particles', *Engineering Applications of Artificial Intelligence*. Pergamon, 10(5), pp. 485–493

[19]. Roylance, B. J. and Hunt, T. M. (1999) *The wear debris analysis handbook*. 1st ed. Oxford UK: Coxmoor Pub. Co

[20]. Roylance, B. J., Williams, J. A. and Dwyer-Joyce, R. (2000) 'Wear debris and associated wear phenomena—fundamental research and practice', *Proceedings of the Institution of Mechanical Engineers, Part J: Journal of Engineering Tribology*, 214(1), pp. 79–105

[21]. Cho, U., and Tichy, J. A. (2000), "A Study of Two-Dimensional Binary Images of Wear Debris as an Indicator of Distinct Wear Conditions", *Tribology Transactions*, 44(1), 132-136

[22]. Anderson, D. P. (1982) 'Wear Particle Atlas. Revised'. Available at: https://apps.dtic.mil/docs/citations/ADA125512

[23]. Kowandy, C. et al. (2007) 'Correlation between the tribological behaviour and wear particle morphology—case of grey cast iron 250 versus Graphite and PTFE', Wear. Elsevier, 262(7–8), pp. 996–1006

[24]. Kaye, B. H. (1993) 'Applied Fractal Geometry and the Fineparticle Specialist. Part I: Rugged boundaries and rough surfaces', Particle & Particle Systems Characterization. John Wiley & Sons, Ltd, 10(3), pp. 99–110

[25]. Thomas, A.; Davies, T. & Luxmoore, A., Computer image analysis for identification of wear particles, Wear, Elsevier BV, 1991, 142, 213-226

[26]. Peng, Z. & Kirk, T.,Computer image analysis of wear particles in three-dimensions for machine condition monitoring, Wear, Elsevier BV, 1998, 223, 157-166

[27]. Myshkin, N.; Kwon, O.; Grigoriev, A.; Ahn, H.-S. & Kong, H.,Classification of wear debris using a neural network, Wear, Elsevier BV, 1997, 203-204, 658-662

[28]. Umeda, A.; Sugimura, J. & Yamamoto, Y.,Characterization of wear particles and their relations with sliding conditions, Wear, Elsevier BV, 1998, 216, 220-228

[29]. Laghari, M. S. & Boujarwah, A.,Wear Particle Texture Classification Using Artificial Neural Networks, International Journal of Pattern Recognition and Artificial Intelligence, World Scientific Pub Co Pte Lt, 1999, 13, 415-428

[30]. Podsiadlo, P. & Stachowiak, G.,Development of advanced quantitative analysis methods for wear particle characterization and classification to aid tribological system diagnosis, Tribology International, Elsevier BV, 2005, 38, 887-897

[31]. Stachowiak, G. P.; Stachowiak, G. W. & Podsiadlo, P.,Automated classification of wear particles based on their surface texture and shape features, Tribology International, Elsevier BV, 2008, 41, 34-43

[32]. Yuan, W.; Chin, K.; Hua, M.; Dong, G. & Wang, C.,Shape classification of wear particles by image boundary analysis using machine learning algorithms, Mechanical Systems and Signal Processing, Elsevier BV, 2016, 72-73, 346-358

[33]. Wang, S.; Wu, T.; Shao, T. & Peng, Z.,Integrated model of BP neural network and CNN algorithm for automatic wear debris classification, Wear, Elsevier BV, 2019, 426-427, 1761-1770

[34]. Sliney, H. E. (1978) 'Dynamics of Solid Lubrication as Observed by Optical Microscopy', A S L E Transactions, 21(2), pp. 109–117

[35]. Scharf, T. W. and Singer, I. L. (2003) 'Quantification of the thickness of carbon transfer films using Raman tribometry', Tribology Letters, 14(2), pp. 137–146

[36]. Wahl, K. J., Chromik, R. R. and Lee, G. Y. (2008) 'Quantitative in situ measurement of transfer film thickness by a Newton's rings method', Wear, 264(7–8), pp. 731–736

[37]. Descartes, S. and Berthier, Y. (2002) "Rheology and flows of solid third bodies: Background and application to an MoS1.6coating", Wear, 252(7–8), pp. 546–556.

[38]. Descartes, S., Godeau, C., and Berthier, Y. (2015), "Friction and lifetime of a contact lubricated by a solid third body formed from an MoS1.6 coating at low temperature", Wear, 330-331, 478-489

[39]. Gouider, M. et al. (2004) 'Mass spectrometry during C/C composite friction: carbon oxidation associated with high friction coefficient and high wear rate', Wear. Elsevier, 256(11–12), pp. 1082–1087

[40]. Kasem, H. et al. (2007) 'Interdependence between wear process, size of detached particles and CO2 production during carbon/carbon composite friction', Wear. Elsevier, 263(7–12), pp. 1220–1229

[41]. Colas, G., Saulot, A., Godeau, C., Michel, Y. and Berthier, Y. (2013). "Describing third body flows to solve dry lubrication issue - MoS2 case study under ultrahigh vacuum", Wear, 305(1-2), 192-204

[42]. Pham, D. L., Xu, C. and Prince, J. L. (2000) 'Current Methods in Medical Image Segmentation', Annual Review of Biomedical Engineering. Annual Reviews 4139 El Camino Way, P.O. Box 10139, Palo Alto, CA 94303-0139, USA , 2(1), pp. 315–337

[43]. Atta-Fosu, T. et al. (2016) '3D Clumped Cell Segmentation Using Curvature Based Seeded Watershed', Journal of Imaging. Multidisciplinary Digital Publishing Institute, 2(4)

[44]. Waggoner, J. et al. (2013) '3D Materials Image Segmentation by 2D Propagation: A Graph-Cut Approach Considering Homomorphism', IEEE Transactions on Image Processing, 22(12), pp. 5282–5293

[45]. Myasnikov, E. V. (2017) 'Hypersepctral image segmentation using dimensionality reduction and classical segmentation approaches', Computer Optics, 41(4), pp. 564–572.

[46]. Ando, E., Hall, S.A., Viggiani, G., Desrues, J., and Besuelle, P. (2012), "Grain-scale experimental investigation of localized deformation in sand: a discrete particle tracking approach", Acta Geotechnica, 7(1), 1-13

[47]. He, P., Fang, K. L. and Liu, X. H. (2013) 'Improved watershed algorithm based on morphology and distance transform', in Applied Mechanics and Materials, pp. 1071–1075

[48]. Win, K. Y., Choomchuay, S. and Hamamoto, K. (2017) 'Automated segmentation and isolation of touching cell nuclei in cytopathology smear images of pleural effusion using distance transform watershed method', in Second International Workshop on Pattern Recognition. SPIE, p. 104430Q

[49]. Blott, S. J. and Pye, K. (2008) 'Particle shape: A review and new methods of characterization and classification', Sedimentology, 55(1), pp. 31–63.

[50]. Mollon, G. and Zhao, J. (2012) 'Fourier-Voronoi-based generation of realistic samples for discrete modelling of granular materials', Granular Matter, 14(5), pp. 621–638

[51]. Mollon, G., and Zhao, J. (2013), "Generating realistic 3D sand particles using Fourier descriptors", Granular Matter, 15(1), 95-108

[52]. Thankachan, T., Soorya Prakash, K. and Kamarthin, M. (2018) 'Optimizing the Tribological Behavior of Hybrid Copper Surface Composites Using Statistical and Machine Learning Techniques', Journal of Tribology. American Society of Mechanical Engineers, 140(3), p. 031610

[53]. Meng Li (2011) 'An intelligent fault diagnosis system of rolling bearing', in Proceedings 2011 International Conference on Transportation, Mechanical, and Electrical Engineering (TMEE). IEEE, pp. 544–547

[54]. Decost, B. L. and Holm, E. A. (2015) 'A computer vision approach for automated analysis and classification of microstructural image data', Computational Materials Science. Elsevier B.V., 110, pp. 126–133

[55]. Laghari, M. S., Memon, Q. A. and Khuwaja, G. A. (2007) 'Knowledge Based Wear Particle Analysis', (January), pp. 998–1002.

[56]. Rouet-Leduc, B. et al. (2017) 'Machine Learning Predicts Laboratory Earthquakes', Geophysical Research Letters. Blackwell Publishing Ltd, 44(18), pp. 9276–9282

[57]. Kearns, M. J. and Vazirani, U. V. (1994) An introduction to computational learning theory. MIT Press

[58]. Xu, R. and Wunsch, D. C. (2008) Clustering.

[59]. Cox, D. R. (David R. (1977) The analysis of binary data. Chapman and Hall.

[60]. Hosmer, D. W. and Lemeshow, S. (2000) Applied logistic regression. Wiley

[61]. El-Habil, A. M. (2012) 'An Application on Multinomial Logistic Regression Model', Pakistan Journal of Statistics and Operation Research, 8(2), p. 271

[62]. Nigam, K., Lafferty, J. D. and McCallum, A. (1999) 'Using Maximum Entropy for Text Classification'.

[63]. Kehler, A. (1997) 'Probabilistic Coreference in Information Extraction'.

[64]. Rong-En Fan et al. (2008) 'LIBLINEAR: A Library for Large Linear Classification', The Journal of Machine Learning Research, 9, pp. 1871–1874

[65]. Mayooran, T. (2018) 'A Gradient-Based Optimization Algorithm for Ridge Regression by Using R', International Journal of Research and Scientific Innovation (IJRSI, Volume V (Issue IV).

[66]. Kalyan, K. et al. (2014) 'Artificial neural network application in the diagnosis of disease conditions with liver ultrasound images.', Advances in bioinformatics. Hindawi, 2014,

[67]. Hornik, K. (1991) 'Approximation capabilities of multilayer feedforward networks', Neural Networks. Pergamon, 4(2), pp. 251–257

[68]. Alexeev, D. V. (2010) 'Neural-network approximation of functions of several variables', Journal of Mathematical Sciences. Springer, 168(1), pp. 5–13

[69]. Rumelhart, D. E., Hinton, G. E. and Williams, R. J. (1985) 'Learning Internal Representations by Error Propagation'.