



Statistical Modeling and Monitoring of Geometrical Deviations in Complex Shapes With Application to Additive Manufacturing

Riccardo Scimone, Tommaso Taormina, Bianca Maria Colosimo,
Marco Grasso, Alessandra Menafoglio, Piercesare Secchi

This is an Accepted Manuscript of an article published by Taylor & Francis in Technometrics on 31 Aug 2021, available online: <http://www.tandfonline.com/10.1080/00401706.2021.1961870>.

This content is provided under [CC BY-NC-ND 4.0](https://creativecommons.org/licenses/by-nc-nd/4.0/) license



Statistical modeling and monitoring of geometrical deviations in complex shapes with application to Additive Manufacturing

Riccardo Scimone

MOX, Dipartimento di Matematica, Politecnico di Milano
Center for Analysis, Decision and Society, Human Technopole

Tommaso Taormina

Dipartimento di Meccanica, Politecnico di Milano

Bianca Maria Colosimo

Dipartimento di Meccanica, Politecnico di Milano

Marco Grasso

Dipartimento di Meccanica, Politecnico di Milano

Alessandra Menafoglio

MOX, Dipartimento di Matematica, Politecnico di Milano
and

Piercesare Secchi

MOX, Dipartimento di Matematica, Politecnico di Milano
Center for Analysis, Decision and Society, Human Technopole

March 11, 2021

Abstract

The industrial development of new production processes like additive manufacturing (AM) is making available novel types of complex shapes that go beyond traditionally manufactured geometries and 2.5D free-form surfaces. New challenges must be faced to characterize, model and monitor the natural variability of such complex shapes, since previously proposed methods based on parametric models are not applicable. The present study proposes a methodology that applies to complex shapes represented in the form of triangulated meshes, which is the current standard for AM data format. The method combines a novel bi-directional way to model the deviation between the reconstructed geometry (e.g., via x-ray computed tomography) and the nominal geometry (i.e., the originating 3D model) with a profile monitoring approach

for the detection of out-of-control shapes. A paradigmatic example consisting of an egg-shaped trabecular shell representative of real parts produced via AM is used to illustrate the methodology and to test its effectiveness in detecting real geometrical distortions.

Keywords: complex shape; statistical process control; geometrical defects; additive manufacturing; industry 4.0

1 Introduction

New production paradigms and advanced inspection capabilities have made novel kinds of complex shapes increasingly available in the industry 4.0 scenario. By exploiting new levels of design freedom, e.g., via additive manufacturing (AM) technologies, it is now possible to produce parts characterized by lightweight structures whose complexity goes beyond the one of traditionally manufactured shapes and simple 2.5D free-form surfaces. Examples of such novel geometries include biomimetic designs, topologically optimized components and lattice structures like the ones shown in Fig. 1. On the other side, a new generation of non-contact metrology systems (e.g., structured lights, laser scanners, X-Ray CT) allows to inspect and measure complex shapes that can be reconstructed starting from very large, unstructured and noisy 3D point clouds.



Figure 1: Examples of complex shapes produced via metal additive manufacturing: a) lightweight bracket for space applications (Jorge (2018)), b) topologically optimized space antenna support¹ and c) rocket engine demonstrator²

As a counterpart to the enhanced functionalities enabled by this novel level of shape complexity, practitioners have to face new significant challenges for the development and adoption of appropriate quality modeling and statistical quality monitoring methods.

Usually, the quality of geometrical shapes is monitored with feature-based approaches, which consist of focusing on relevant quality indicators, often related to functional or aes-

¹https://www.eos.info/case_studies/additive-manufacturing-of-antenna-bracket-for-satellite

²<https://www.etmm-online.com/3d-printing-a-rocket-engine-a-886960/>

thetic performances, and monitor them with time. This procedure is consistent with the standard industrial practice of Statistical Quality Monitoring (SQM), where traditional control charts on macro- or micro-geometrical descriptors (e.g., diameters, roughness, lengths or distances) are basically assuming these features as representative of part functionality. The approach is summarized in Fig. 2a, where three representative features, namely f , h , Ra , are shown as examples of critical indicators to be monitored with time. Feature-based approach can be applied even to complex surfaces and will be considered as benchmark competitor in this paper, too. In the case of latticed free-form shapes as the one shown in Fig. 1, synthetic indicators as the overall printed volume (i.e., the volume included in the “latticed” free-form structure) can be considered as synthetic quality indicator to be measured with time. This approach will be considered as representative of the industrial practice.

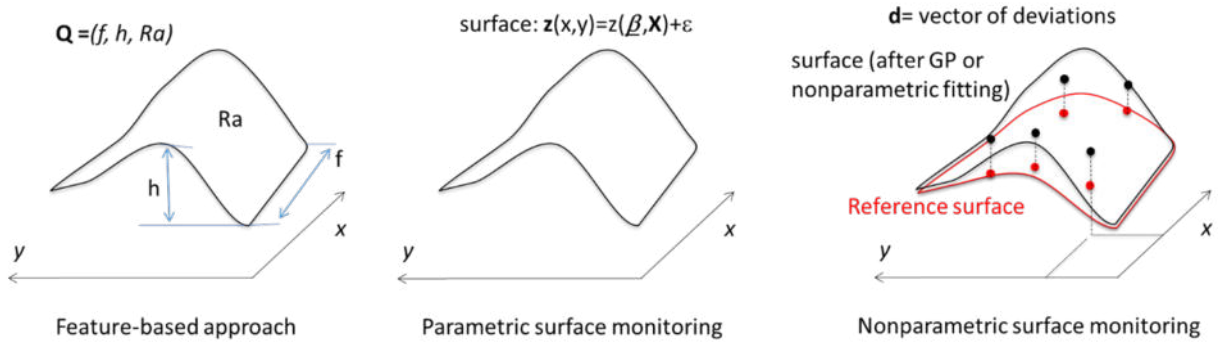


Figure 2: from feature-based monitoring (a) to parametric (b) and non-parametric (c) surface monitoring

With the increasing advent of free-form or complex shapes and a new generation of measurement systems (e.g., non-contact sensors as light scanners or structured light), the dense and rich content of the whole point cloud can be usefully considered for quality monitoring. In this case, surface monitoring represents a natural extension of profile monitoring to efficiently detect undesired changes of the final shape. In this case, two main streams of approaches have been proposed in the literature. In the first case, a parametric model of the surface geometry is considered. For each new surface, the model parameters are estimated and monitored with time to detect undesired shape changes (Colosimo et al. (2014)). In a second stream of research, surface point cloud is firstly fitted via Gaussian processes (del Castillo et al. (2014), Wang et al. (2014)) or non-parametric smoothing (Zang and

Qiu (2017), Zang and Qiu (2018)). Then, for each new surface, the difference between the smoothed geometry and the reference shape (usually the average surface computed in Phase 1) is computed on a set of locations acting as reference position or landmarks. At each landmark, the deviation is measured as the Euclidean distance between the corresponding points observed on the fitted and the reference surfaces.

A last class of approaches proposed in the literature do not consider pre-processing via surface smoothing but simply compute the deviations of point clouds with respect to the nominal CAD shape by simply measuring the shortest distance of each element in the point cloud to the nominal CAD surface. Then, the distribution of this deviation set can be summarized in the quantile-quantile (Q-Q) plot and then monitored via profile monitoring (Wells et al. (2013)). More recently, a new procedure has been suggested which consists of dividing the point cloud volume into Region of Interests (ROIs), compute the mean of the deviations observed in each ROIs and monitor them via multivariate generalized likelihood ratio (GLR) control chart (Stankus and Castillo-Villar (2018)).

With reference to these existing literature, two main limiting assumptions have to be overcome when very complex 3D shapes as the one considered in this paper are of interest. First, with few exceptions (del Castillo et al. (2014)), most of the parametric and non-parametric models are based on the assumption of **2.5D surfaces**, i.e., surfaces where one coordinate can be modeled as a function of the other two (Colosimo et al. (2014), Wang et al. (2014), Zang and Qiu (2017), Zang and Qiu (2018)), as shown in Fig 2b. This assumption can hold even after coordinate transformations (e.g., cylindrical coordinates) but basically restrict the set of possible shapes to be modeled with time. Many complex shapes obtained via additive manufacturing take full advantage of the design freedom allowed by this technology as in the paradigmatic example of the latticed egged shape we are dealing with.

Second, “**structured**” data set are usually considered, where data are assumed to be measured on a fixed raster or are somehow remapped to obtain a “**one-to-one**” correspondence of surface data for control charting. Structured data can be obtained using a specific measurement system (i.e., coordinate measuring machine or structured light) or can be gained after surface smoothing registering the position of the surface on some landmarks or the

mean deviation observed in some fixed ROIs (Stankus and Castillo-Villar (2018)). When large, unstructured and noisy point clouds are available on complex 3D shapes, the definition and identification of landmarks or ROIs can be a very difficult or even intractable task. This is why we will not assume this task as required in our newly proposed solution.

This study follows the route of analyzing the deviation between a measured shape and a nominal model but it introduces novel contributions aimed at overcoming the simplistic assumptions that prevent the applicability of the method proposed in the literature to very complex 3D shapes, as the ones produced via 3D printing or additive manufacturing (Colosimo et al. (2018)). Assuming large, unstructured and noisy point clouds are measured on complex 3D shapes, we consider a metric that does not require landmarks, one-to-one correspondence or identification of ROIs on a fully 3D (and not simply 2.5D) volume.

Moreover, we assume that one or both the objects to be compared can be represented as point clouds or triangulated mesh, as shown in Figure 3. As a matter of fact, when the deviation between the inspected item and its nominal shape is of interest, usually the first object is a point cloud while the latter is usually the triangulated mesh derived by the CAD design.

In our paper, we will exploit a solution based on the Hausdorff distance (Hausdorff (1914)), a computationally tractable and flexible metric to compute distances between two shapes despite of the fact that they are represented as unstructured point clouds or triangulated meshes (as shown in Fig. 3). This extraordinary flexibility will take our method close to any applicative domains, despite of the specific convention used to measure shape deviations.

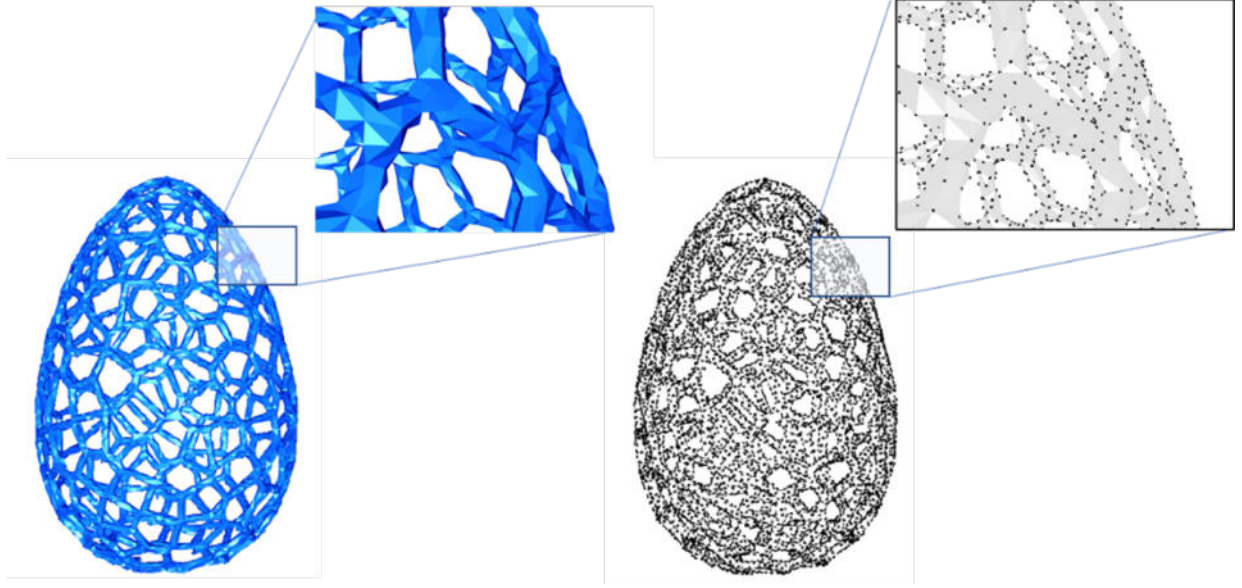


Figure 3: Types of data set assumed in our proposed method to represent the final geometry: triangulated mesh (left) or unstructured point clouds (right), obtained considering the mesh vertices.

The Hausdorff distance is a metric applied by several authors in a wide range of applications (Zhou and Wang (2009), Zhao et al. (2005), Mémoli and Sapiro (2004), Aspert et al. (2002), Kwan-Ho Lin et al. (2001)). It is a very general notion of metric given for any couple of subsets of a given metric space, thus including couples of shapes, surfaces, meshes or point clouds in the usual 3D Euclidean space, regardless of their complexity or mutual relationship. As stated in Alt et al. (2003), the Hausdorff distance represents the most natural distance measure for couples of geometrical objects where no point-to-point correspondence is available, a characteristic that makes it particularly suitable to determine differences between complex shapes, as the ones we are focusing on. It is also the best-known metric between subsets of a metric space (Eiter and Mannila (1999), Conci and Kubrusly (2018)). As far as the computation of distances between sets, other notions of distances have been proposed (for an extended overview, the reader is referred to Eiter and Mannila (1999), Gardner et al. (2014), Conci and Kubrusly (2018) and the references therein). However, the Hausdorff distance has a much more general field of validity being a very general notion of metric applicable to any couple of shapes, surfaces, meshes or point clouds in the usual 2D or 3D Euclidean space, regardless of their complexity thanks to its computationally efficient tractability. This is also the reason why it is the default distance

notion adopted in various software tools, like CloudCompare and MeshLab, in addition to being implemented in a large number of libraries (like the OpenCV library for image processing in C and Python), and other packages for Matlab and R.

While the Hausdorff distance has been used in different contexts, as in jump detection of regression surfaces (Kang and Qiu (2014)), or in the field of image analysis to study nanoparticles aggregation (Sikaroudi et al. (2018)), to the best of our knowledge it has never been considered in the framework of statistical quality modeling and monitoring of complex shapes.

A second novel contribution of the proposed methodology consists of exploiting a specific feature of the Hausdorff distance, whose computation is based on two different contributions, which depend on the direction of the comparison. When object A and B have to be compared, the Hausdorff distance is the maximum of two different distance maps: the first measuring deviations observed on A with respect to B and the second one measuring deviations observed on B with respect to A. In our newly proposed method, we exploit this bi-directional nature of the Hausdorff-based distance, to enhance the capability of our method to detect out-of-control shapes and provide diagnostic aids when an alarm signal is issued.

Eventually, our proposed method is based on summarizing these two directional deviation maps in their empirical probability density functions (PDFs) and then using profile monitoring as basic framework to detect the out of control. In particular, we propose a profile monitoring method where the probability density functions of the two deviation maps are modeled and monitored via a functional variant of the Principal Component Analysis (PCA) known as Simplicial Functional PCA (SFPCA, Hron et al. (2014)). Menafoglio et al. (2018) showed that the SFPCA applied to the profile monitoring of PDFs generalizes the profile monitoring of Q-Q curves leading to better performances especially when the underlying assumptions of the Q-Q plot based approach are violated. Two couples of control charts are designed, one for each distinct PDF, to monitor the natural variability of serially produced complex shapes in terms of deviations from their common nominal geometry. Due to the asymmetrical nature of the captured information, some types of out-of-control geometrical distortion may be signaled by one control chart only and some other distor-

tions may be signaled by both of them. This, combined with the physical interpretability of the information content of each deviation map, is shown to ease the identification and diagnosis of signaled anomalies.

The proposed approach addresses the class of problems where the complex shape is represented in terms of a 3D point cloud or a triangulated mesh. In the latter case, we approximate the meshes with the point clouds of their vertices and advocate the use of Bernstein polynomial estimators to fit the PDFs of the resulting deviation maps before applying the SFPCA (Menafoglio et al. (2018)). The proposed approach can be thus applied to any shape, regardless of its geometrical complexity and without requiring distributional assumptions for the measured deviation from a nominal model. This makes the method generally applicable to a wide range of problems. Nowadays, large, noisy unstructured point clouds measured via non-contact sensors (e.g., laser scanners, X-Ray CT) represents the standard rather than the exception in the Industry 4.0 scenario. This is why we do believe that different interesting applications of the proposed approach can be used in different contexts, ranging from shape classification to quality optimization (i.e., selecting the process parameters to minimize the deviation from the nominal). Besides our motivating I4.0 application, point clouds analysis on high-dimensional manifolds is also frequent in many different application domains, from computational biology to medical image analysis, from material science image-based classification to financial data. This is why many possible extension of our newly proposed contribution can be seen in different directions. However, to illustrate the methodology and to show its benefits we present a paradigmatic example. It consists of an egg-shaped trabecular shell (hereafter denoted as “egg” for sake of simplicity) representative of real parts produced via AM like the ones shown in Fig. 1. Different copies of the part were 3D-printed via Fused Deposition Modeling (FDM) and their geometry was reconstructed via x-ray computed tomography. A Phase I implementation of the proposed control charting method is presented and its performances are demonstrated by means of a simulation study inspired by the real egg-shape case study. **A comparison against a benchmark approach representative of the common industrial practice for 3D-printed complex shapes is also discussed.**

The remaining part of this work is organized as follows: Section 2 describes the rationale

behind the choice of the motivating case study, Section 3 describes the methodology, while a simulation study is proposed in Section 4. Results of the application of the method to a dataset of real parts, produced via AM, constitute Section 5 and Section 6 concludes the paper.

2 A paradigmatic example

The case study here presented was chosen as paradigmatic example of the challenges imposed by novel classes of shape complexity. It consists of an egg-shaped trabecular structure that recalls topologically optimized and lightweight shapes enabled by AM design freedom (Fig. 4). As in the case of free-form surfaces, the egg-shaped surface cannot be simply parametrized: due to such complexity, the family of model-based methods is not applicable, which motivates the investigation of different approaches. Moreover, its geometry cannot be fully reconstructed via 3D laser scan, but only by means of X-ray computed tomography, which makes the egg an archetype of shape complexity that goes beyond free-form surfaces. The shape was designed starting from a Voronoi tessellation of an egg surface and creating a trabecular structure with struts placed along the boundaries of tessellated regions (Galimberti (2018)). The egg was 60 mm tall with a maximum diameter of 40 mm and it was produced via polymer FDM, a process where a filament of thermoplastic material is heated, extruded through a nozzle tip and deposited onto a platform to build parts layer by layer (Gibson et al. (2014)). The X-ray Computed Tomography (CT) of the 3D-printed egg was used to reconstruct a triangulated mesh of its geometry. Both the reconstructed shape of the egg and its nominal model were represented in STL (Standard Triangulation Language) format, which is the current standard for AM processes.

The X-ray CT methodology represents the only way to reconstruct both outer and inner surfaces of a 3D object. The CT measurement results in a voxel dataset, that can be also represented as a stack of greyscale images. The surface of the manufactured part corresponds to the grey value transition between the low-density air surrounding the part and the high-density material of the part itself. This surface, also called iso-grey-value-surface, can be extracted as a polygonal mesh. Due to the very high CT measurement resolution, the vertices of the extracted mesh are so dense that passing from a representation in terms

of vertices and edges (mesh) to vertices only (point cloud) yields no or negligible loss of information about the reconstructed geometry.

Fig. 4 shows the CAD model representing the nominal shape of the egg (a), one copy of the 3D-printed parts (b) and the reconstructed shape of the 3D-printed part via x-ray CT (c). Due to the intrinsic “signature” of the FDM process, a natural variability of the 3D-printed shape with respect to the nominal is present.

Fig. 5 shows a detail of the mesh of a 3D-printed part generated by the X-ray CT, which highlights the high level of detail at which the geometry is reconstructed. Fig. 5 also explains the terminology that is used in the rest of the paper. The lattice structure consists of nodes and struts, where a node represents a joint where two or more struts meet, and a strut is a structural component that connects two nodes (Syam et al. (2018)). The terms “face”, “edge” and “vertex” are instead used to indicate the geometrical elements of each polygon the mesh is composed of.

Fig. 6 shows a detail of the superimposition between the reconstructed mesh and the nominal one aligned via Iterative Closest Point (ICP) (Besl and McKay (1992)). Details about the alignment operations as pre-processing step are provided in Section 3.

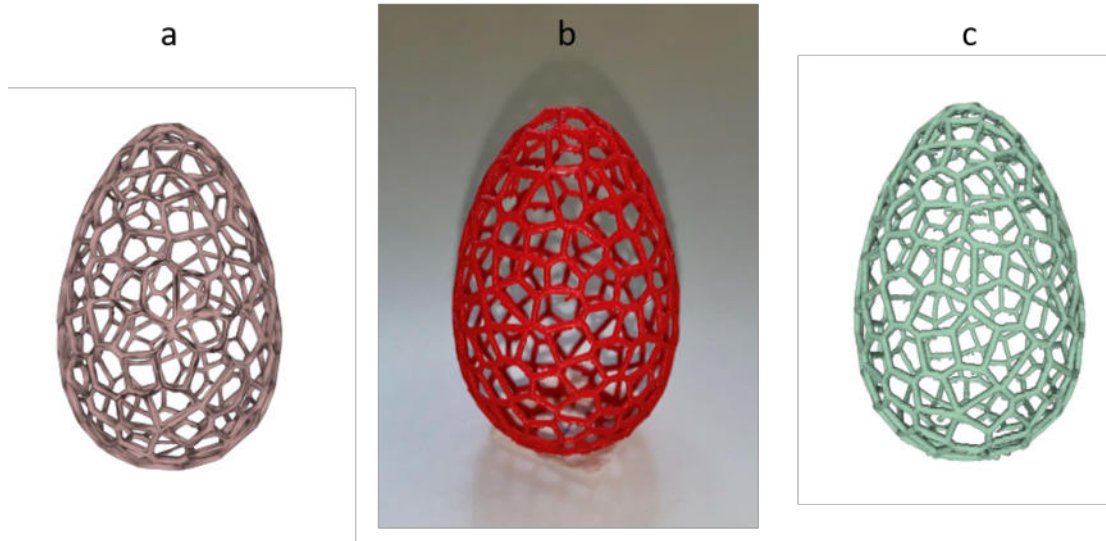


Figure 4: Nominal CAD model (left panel), 3D-printed part (central panel) and reconstructed geometry of the 3D-printed part via x-ray CT (right panel).

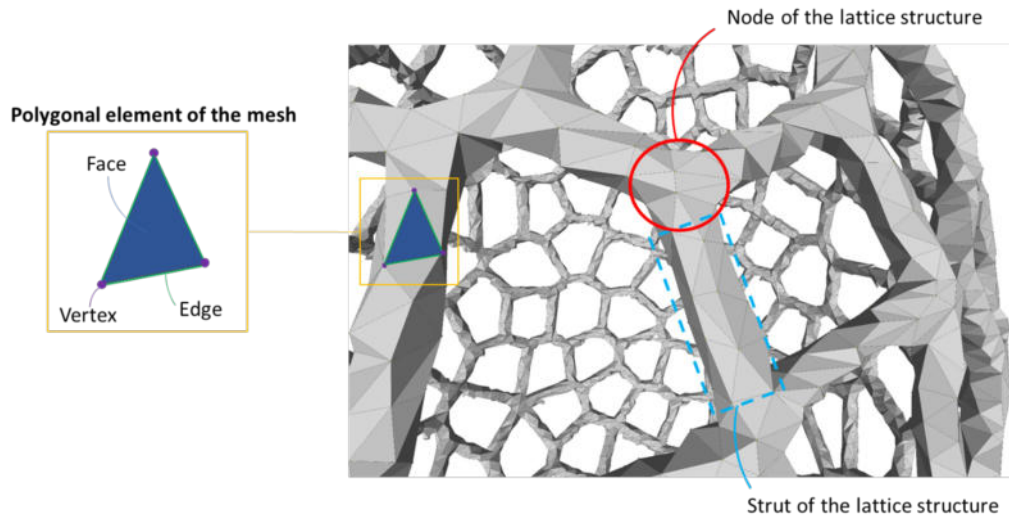


Figure 5: Detail of the mesh of a 3D-printed part reconstructed via X-ray CT

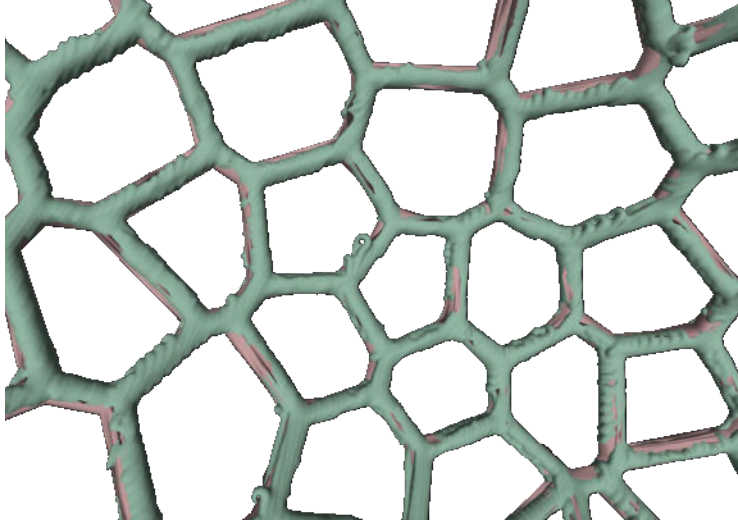


Figure 6: Geometrical mismatch between the nominal shape (pink) and the real shape reconstructed via x-ray CT (green, superimposed).

In this study, seventeen copies of the same egg were produced by means of the same FDM system and the same polymeric material. The dataset includes fifteen in-control shapes characterized by good geometrical quality and lack of defects, and two shapes with two different kinds of geometrical distortions.

Two examples of different types of out-of-control geometrical distortions are shown in Figures 7 and 8. In particular, Fig. 7 shows the reconstructed shape of one egg where several local irregularities during the deposition of the extruded filament produced irregular geometries of the struts and several regions where excessive material was present. This egg labeled as out-of-control was produced with a pigment of the polymer filament different from the one used for other eggs. The pigment of the filament may yield slightly different behavior when the filament is heated up and extruded, with consequent effects on the final quality of the part (Valerga et al. (2018)). Fig. 8 instead shows an example of an egg with a missing strut, indicated by a red line. Differently from the example shown in Fig. 7, in this case the defect was artificially injected: the strut is present in the nominal geometry but it was removed from the 3D-printed egg.

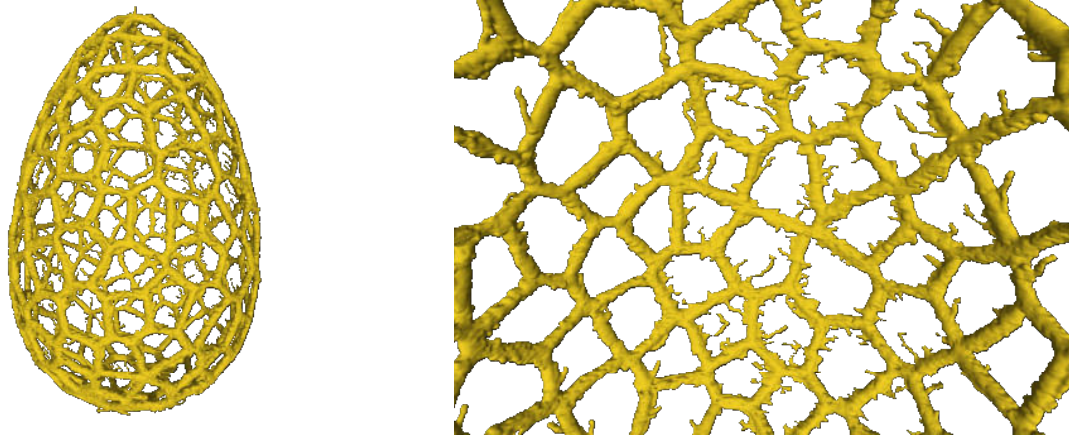


Figure 7: Defectiveness caused by the presence of exceeding mass around the struts of the as-printed part.

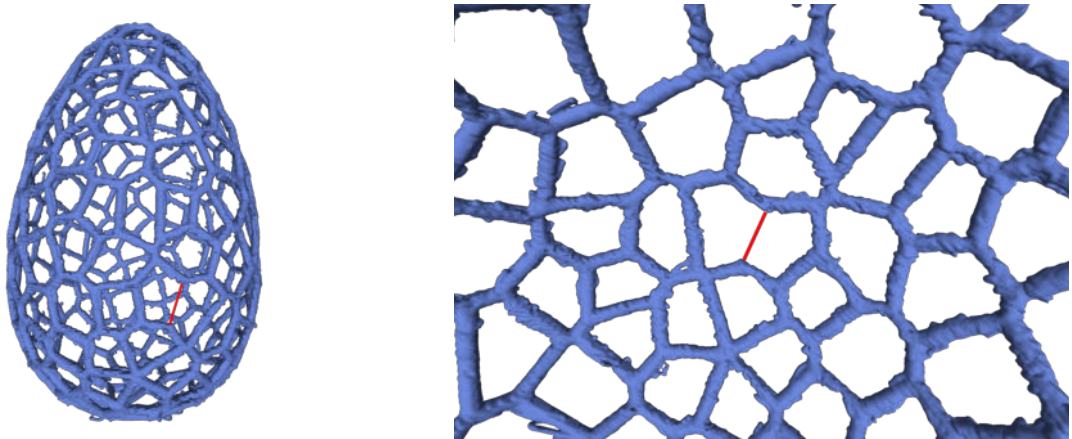


Figure 8: Defectiveness caused by the absence of a strut with respect to the nominal model.

The two different kinds of defectiveness shown in Fig. 7 and Fig. 8 are representative of anomalies that are critical for the statistical process monitoring of complex shapes. The presence of excessive material connected to the struts is representative of a real anomaly

in the filament deposition, which affects the geometrical accuracy and quality of the final part. The example entailing a missing strut is representative of real applications where a local distortion may compromise the mechanical performances of the overall part. Both these defective shapes should be signaled as soon as possible to identify non conforming parts and to detect an out-of-control state of the production process. Generally speaking, a variety of defects could actually occur in additive manufacturing processes in general and FDM in particular. They include local geometrical distortions (typically located in the presence of complex geometrical features, like not fully supported overhang surfaces or thin walls), pores, cracks and delamination, etc. In FDM, geometrical errors and dimensional inaccuracies are quite common. Geometrical errors are usually related to a local lack or excess of material caused by various factors. An excess of material could consist of the so-called “stringing” effect: the result is the presence of a burr in different locations of the part. The example shown in Figure 7 is representative of this kind of defect. It can be caused by the adopted process parameters and the temperature control of the nozzle. A lack of material could be caused by a lack of local adhesion between two layers, a disrupted material flow or a misaligned nozzle (Gibson et al. (2014)). The defect shown in Figure 8 was selected as representative of this second category. Other local and global dimensional inaccuracies may be obtained by varying the part orientation or using non optimal process parameters, especially in the presence of complex shapes (Górski et al. (2013)).

3 Methodology

We denote by $\{S_j\}_{j=1,\dots,N}$ a dataset of meshes, each representing a real shape produced on the basis of the same nominal model P , which is a mesh as well. By *mesh*, we mean that each S_j is defined by a finite set of vertices $\{v_1^j, \dots, v_{B_j}^j\} \subset \mathbb{R}^3$, which are called vertices of the mesh, coupled with a set of triangles $\{T_j^j, \dots, T_{R_j}^j\}$, which are the faces of the mesh. The edges of the triangles form the set of the edges of the mesh and, of course, the vertices of each triangle need to be among the vertices of the mesh. This is exactly the case of our paradigmatic example in Section 2: $N = 17$ shapes are manufactured on the basis of the same CAD model, and from each shape a mesh is obtained, via X-ray CT.

In order to identify in-control and out-of-control conditions in the shapes, we here aim to develop a methodology to characterize the variability of the dataset $\{S_j\}_{j=1,\dots,N}$. Clearly, a first, crucial, assumption is that the meshes associated to the manufactured shapes are a faithful representation of them or, at least, that they contain all the information about the structural properties which are of interest for process control. Ideally, we would like to represent each mesh S_j through descriptors capturing the full information content on the differences between S_j and the prototype P . To this end, as detailed in Section 3.1, each S_j will be represented by a couple of deviation maps. This requires a preliminary discretization of each mesh, i.e., to extract a point cloud from each S_j and from P . Of course, a second assumption to be made is that the point clouds extracted from the meshes are dense enough to avoid the loss of information about the geometry of the objects themselves. For example, if the point clouds $\{v_1^j, \dots, v_{B_j}^j\}$, $j = 1, \dots, N$ of the vertices of the S_j s are deemed to be dense enough, one may choose them as discretizations: this is the approach followed in the present work. Alternatively, a uniform sampling on the faces of the meshes could be performed. In both cases, from each S_j and P one obtains point clouds $S_j^* = \{p_1^j, \dots, p_{M_j}^j\}$ and $P^* = \{p_1^P, \dots, p_{M_P}^P\}$. The computation of the aforementioned deviation maps requires the computation of distances between the points of the point clouds, so that a preliminary alignment is needed. The proposed method can be summarized by the following steps.

0. Align each mesh S_j to the prototype P : the well-known ICP algorithm (Besl and McKay (1992)) can be used in this phase. Extract then S_j^* and P^* from each S_j and P respectively. From now on we will drop the apex \cdot^* and, with a little abuse of notation, with S_j we will denote the point cloud $\{p_1^j, \dots, p_{M_j}^j\}$ and, analogously, with P we will denote the point cloud $\{p_1^P, \dots, p_{M_P}^P\}$
1. Compare the clouds $\{S_j\}_{j=1,\dots,N}$ and the prototype P , through appropriate deviation maps. Build a summary of these maps attaining a balance between the complexity and richness of the representation and the ability to perform the statistical analysis.
2. Explore and analyze the variability structure of the new dataset with suitable methods
3. Use the results of the analysis to produce a control chart scheme to identify in-control or out-of-control conditions.

The ICP algorithm has become a standard in the framework of 3D shape alignment. The registration approach adopted in this study involves a first rough alignment based on manually selected landmarks, followed by the fine registration via ICP. This approach is implemented in various software tools for mesh manipulation and analysis, like the open source Meshlab platform (Cignoni et al. (2008)) used in this study. However, such alignment operation is known to be affected by the initialization condition. Because of this, a sensitivity analysis with respect to the choice of landmarks is presented in the Appendix.

In the following, we shall assume that the S_j s are already aligned to the prototype (i.e., that Step 0 has already been performed), and we also assume that any residual misalignment error is enclosed into the overall natural variability of the monitored shapes, and such error is assumed to be small with respect to the shape variations among different parts. The following subsections describe the remaining steps.

3.1 Step 1: Hausdorff distance, distance maps and PDFs

To deal with the problem of measuring differences between two aligned clouds, we refer to the general definition of Hausdorff distance between two subsets of a metric space (Henrikson (1999)).

Definition 3.1. *Let X, Y be two closed, bounded, non-empty subsets of a metric space (U, d) . Their Hausdorff distance is defined as*

$$d_H(X, Y) := \max \left\{ \sup_{x \in X} \inf_{y \in Y} d(x, y), \sup_{y \in Y} \inf_{x \in X} d(x, y) \right\}$$

This definition implicitly defines two *deviation maps*, namely

$$d_Y : X \rightarrow \mathbb{R}^+, d_Y(x) := \inf_{y \in Y} d(x, y)$$

and

$$d_X : Y \rightarrow \mathbb{R}^+, d_X(y) := \inf_{x \in X} d(x, y)$$

On this basis, one can see the Hausdorff distance as

$$d_H(X, Y) := \max \left\{ \sup_{x \in X} d_Y(x), \sup_{y \in Y} d_X(y) \right\}.$$

Note that d_Y associates to each $x \in X$ its distance from Y , while d_X acts analogously associating to each point $y \in Y$ its distance from X . Clearly, these maps carry a richer information content on the mutual relation between X and Y than the Hausdorff distance $d_H(X, Y)$, which could face huge fluctuations when single points are added to one of the sets X or Y . The method proposed in this work grounds on the key observation that, in general, $\sup_{x \in X} d_Y(x) \neq \sup_{y \in Y} d_X(y)$. This fact, beside being reasonable from the mathematical viewpoint (e.g., d_X and d_Y do not have the same spatial domain), suggests that both deviation maps should be analyzed to fully characterize the geometrical differences between the nominal model and the reconstructed object. **In fact, the couple (d_Y, d_X) characterizes the geometrical differences between X and Y because the Hausdorff distance (which is directly derived from these maps) has the good property that it never vanishes unless the sets under study coincide exactly (Seal and Bhowmick (2017)) - i.e., $d_H(X, Y) = 0$ if and only if $X = Y$. This is actually an immediate corollary of the properties of distance functions in metric space. As a consequence, if the hypotheses regarding the accuracy of the extracted point clouds, stated at the beginning of Section 3, are valid, we may assume that the couple of distance maps obtained substituting S_j and P for X and Y codifies all the relevant information about the geometrical deviations between the manufactured object and the prototype. These considerations justify the analysis and control framework which is outlined in the following.**

For a dataset of point clouds, the definition of the deviation maps is extended as follows. Seeing each cloud (including the prototype) as a subset of the Euclidean space equipped with the usual Euclidean metric, the pair of deviation maps (d_{S_j}, d_P^j) for the j -th object is defined as

- $d_{S_j} : P \rightarrow C_j$, associating to each $p \in P$ its distance from S_j : $d_{S_j}(p) = \inf_{s \in S_j} d(p, s)$
- $d_P^j : S_j \rightarrow K_j$, associating to each $s \in S_j$ its distance from P : $d_P^j(s) = \inf_{p \in P} d(p, s)$

where d is the Euclidean metric, and C_j, K_j can be taken to be intervals in \mathbb{R} . Each object can thus be represented by the corresponding pair of distance maps (d_{S_j}, d_P^j) , which are real functions over a spatial domain.

For each cloud S_j one gets two finite samples of deviations, hereafter denoted by vectors

D_{S_j} and D_P^j (named deviation vectors). Note that these vectors typically have a high-dimensionality (if based on a dense or a large point cloud). Moreover, the vectors might have different dimension for different objects. In fact, although pairs $\{(D_{S_j}, D_P^j), j = 1, \dots, N\}$ represent the geometric discrepancy between the produced objects and the nominal model, statistical process control on this basis appears unfeasible, because of the dimensionality and complexity of these data.

To allow for a comparison of the geometrical discrepancy between produced objects and the nominal model, one thus needs to build appropriate summaries of (D_{S_j}, D_P^j) . This may be obtained via sample statistics (sample mean, sample variance, quantiles), at the expense of a significant loss of information. To attain a balance between the information content being preserved and the need to simplify the deviation vectors, we summarize (D_{S_j}, D_P^j) through their (estimated) probability density functions (PDFs), denoted by f_{S_j} and f_P^j , similarly as proposed by Menafoglio et al. (2018) in the context of metal foam monitoring. Note that these PDFs can be estimated by working directly on the elements of D_{S_j} and D_P^j (i.e., the distances) or on any monotone transformation of distances; in this work, all the analyses are carried out on natural logarithms of the distances. **The logarithmic transformation was chosen because of its properties of reducing the skewness of data being densely concentrated near zero. The condition where a lot of distance values are near zero seems reasonable, as it represents a situation where large parts of the manufactured objects matches the prototype and residual small regions are affected by anomalies. Of course, this choice is not prescriptive, and a practitioner may find that another kind of transformation (e.g. a power transformation) is more suitable in a different context.**

In the following, we shall consider a Bernstein polynomial estimator to obtain a smooth estimate of the PDFs f_{S_j} and f_P^j from the deviation vectors, following the approach of Menafoglio et al. (2018). **The Bernstein estimator for the cumulative distribution function (CDF) F from the empirical CDF F_N is defined as**

$$\hat{F}(t) = \sum_{k=0}^N F_N(k/N) b_{k,N}(t).$$

This allows one to obtain an explicit estimator for the PDF as

$$\hat{f}(t) = N \sum_{k=0}^{N-1} [(F_N((k+1)/N) - F_N(k/N)) b_{k,N}(t)]$$

where $b_{k,N}$ is the k -th Bernstein polynomial of order N . Bernstein estimators were chosen because they are well suited to estimating density with a compact interval as support, which is exactly our case, since we are dealing with continuous shapes. On the other hand, kernel-based estimators may face instability at the borders (Chen et al. (2000); Babu et al. (2002); Silverman (2018)). Moreover, these estimators are consistent and smooth (Babu et al. (2002)).

Independently of the chosen smoothing method, this step leads to the definition of two datasets $\mathbf{f}_S = \{f_{S_j}, j = 1, \dots, N\}$ and $\mathbf{f}_P = \{f_P^j, j = 1, \dots, N\}$, each containing the N PDFs of the deviation vectors. We remark that the datasets \mathbf{f}_S and \mathbf{f}_P allow for different geometric interpretations, intimately linked to the *directional* nature of deviation maps d_{S_j} and d_P^j . For instance, in regions of S_j characterized by excessive material (e.g., in the case of Fig. 7), the deviation map d_P^j is likely to highlight large deviations. On the contrary, regions in S_j presenting a lack of material (as in the missing strut in the example shown in Fig. 8), would be highlighted by the deviation map d_{S_j} , as discrepancies in d_P^j are measured only for points belonging to the mesh S_j (i.e., missing parts are not accounted for in this deviation map). Clearly, when the j -th produced object presents both excess and lack of material, both deviation maps are expected to highlight a geometrical discrepancy between S_j and P .

3.2 Step 2: Exploring the variability of geometrical discrepancies

To explore in-control variability and detect possible out-of-control conditions, we here follow the approach of Menafoglio et al. (2018), and we use the Simplicial Functional Principal Component Analysis (SFPCA, Hron et al. (2014)), which represents an extension to PDFs of Principal Component Analysis (PCA). PCA is a well known and widely used technique to explore and interpret the variability of an $N \times p$ multivariate dataset. This technique grounds on the spectral analysis of the covariance matrix of the dataset, which is well defined as long as the dataset can be embedded in a separable Hilbert space $(H, \langle \cdot, \cdot \rangle, +, \cdot)$ (Ramsay and Silverman (2005)). This is also true when H is infinite dimensional, which is indeed the case of functional and PDF data, since both sample means and covariance operators (i.e., the infinite-dimensional analogue of the sample covariance matrix) can be

defined. Moreover, the interpretation of the Principal Components and of the associated eigenvalues of the covariance operator is analogous to the Euclidean case (Ramsay and Silverman (2005), Horváth and Kokoszka (2012)).

A proper Hilbert space setting for PDF data has been defined and studied in several works (Egozcue et al., 2006; Gerard van den Boogaart et al., 2010; Pawłowsky-Glahn et al., 2014): it is the Hilbert space B^2 , an infinite dimensional generalization of the Aitchison's simplex for finite-dimensional compositional vectors (Aitchison, 1986; Barceló-Vidal, 2001; Filzmoser et al., 2018), with its operation of inner sum, product by a scalar and inner product. Moreover, a PCA procedure has been proposed, under the name of Simplicial Functional Principal Component Analysis (Hron et al., 2014) and its practical application to an example of Profile Monitoring and Statistical Process Control is extensively described in Menafoglio et al. (2018); we refer the reader to these works for a full technical description of SFPCA. In the present work, we apply SFPCA separately to the two datasets \mathbf{f}_S and \mathbf{f}_P , and use the results of both analyses to build a control framework, as further detailed in Section 3.3.

After performing SFPCA on \mathbf{f}_S , we obtain a sequence of $N - 1$ couples eigenvalues-eigenfunction $(\lambda_i^{(S)}, \zeta_i^{(S)})$, $i = 1, \dots, N - 1$, of the sample covariance operator, where the sequence of eigenvalues is non-increasing and the eigenfunctions are mutually orthogonal. Moreover, for each $f_{S_j} \in \mathbf{f}_S$ we can define, as in the Euclidean case, its scores along the i -th Principal Component as $z_{ji}^{(S)} = \langle f_{S_j} - \bar{f}_S, \zeta_i^{(S)} \rangle_{B^2}$, where \bar{f}_S is the mean of \mathbf{f}_S computed in B^2 . Finally, as usually done, a number $K^{(S)} \in \{1, \dots, N\}$ of Principal Components to be retained is chosen, on the basis of a threshold on the minimum fraction of variance explained by the PCs. Analogously, the same procedure can be applied to \mathbf{f}_P , obtaining a set of $K^{(P)}$ scores z_{ij}^P along the first $K^{(P)}$ principal component $\zeta_i^{(P)}$. In the present work, the selection of $K^{(S)}$, $K^{(P)}$ is always performed as to explain the 98% of the total variability of the dataset \mathbf{f}_S , \mathbf{f}_P , respectively, as suggested in Menafoglio et al. (2018).

3.3 Step 3: Profile monitoring of geometrical discrepancies

To monitor the objects S_j , $j = 1, \dots, N$, represented through the datasets of densities \mathbf{f}_S and \mathbf{f}_P , we shall rely on the results of the previously described SFPCA procedure, applied

separately on each dataset. For brevity, we detail the method referring to the dataset \mathbf{f}_S , analogous development being applied to \mathbf{f}_P . In the PCA-based approach to profile monitoring, it is standard to consider two different statistics, the Hotelling T^2 statistic and the Q statistic (Colosimo and Pacella (2010)), generating the T^2 and the Q chart respectively. Both statistics have been defined for density data (Menafoglio et al. (2018)): chosen a number $K^{(S)}$ of retained components, we have

$$T_j^{2(S)} = \sum_{i=1}^{K^{(S)}} \frac{(z_{ji}^S)^2}{\lambda_i^{(S)}}$$

$$Q_j^S = \left\| f_{S_j} - f_{S_j}^* \right\|_{B^2}^2 = \sum_{i=K^{(S)}+1}^{\infty} z_{ji}^2$$

where $f_{S_j}^* = \bar{f}_S + \sum_{i=1}^{K^{(S)}} z_{ji} \zeta_i^{(S)}$ is the reconstruction of f_{S_j} obtained considering the first $K^{(S)}$ principal components. As detailed in Colosimo and Pacella (2010) and Menafoglio et al. (2018), the T^2 statistic allows one to identify anomalies in the PDFs based on the scores along the retained components, while the Q statistic detects anomalies in the reconstruction error, leading to a complementary insight with respect to the one provided by the T^2 statistic. These statistics lead us to two couples of control charts, $(T^{2(S)}, Q^S)$ for \mathbf{f}_S and $(T^{2(P)}, Q^P)$ for \mathbf{f}_P respectively. We assume that an element is out-of-control if at least one of the four charts raises an alarm since, as discussed in Section 3.1, the datasets \mathbf{f}_S and \mathbf{f}_P may carry complementary information. The couples $(T^{2(S)}, Q^S)$ and $(T^{2(P)}, Q^P)$ cannot be assumed to be independent, however we assume that $T^{2(S)}$ is independent from $Q^{(S)}$ and that $T^{2(P)}$ is independent from $Q^{(P)}$, as done in (Colosimo and Pacella (2010)). Hence, if an overall type I error probability equal to α^* is desired, we consider a Bonferroni correction setting a type I error probability $\alpha' = \alpha^*/2$ for each couple of charts, and finally we exploit the independency between the charts within each couple by setting a type I error probability $\alpha = 1 - \sqrt{1 - \alpha'}$ for each one of the four charts. In the present work, we always set $\alpha^* = 1\%$. As regards the control limits, in each T^2 chart we follow the Phase I standard procedure for the case when the size of the preliminary samples is unitary (i.e. the case of *individual observations*, as detailed in Montgomery (2009)), setting the Upper Control Limit as

$$UCL_{\alpha}(T^2) = \frac{(N-1)^2}{N} q_{Beta(1-\alpha, K/2, (N-K-1)/2)}$$

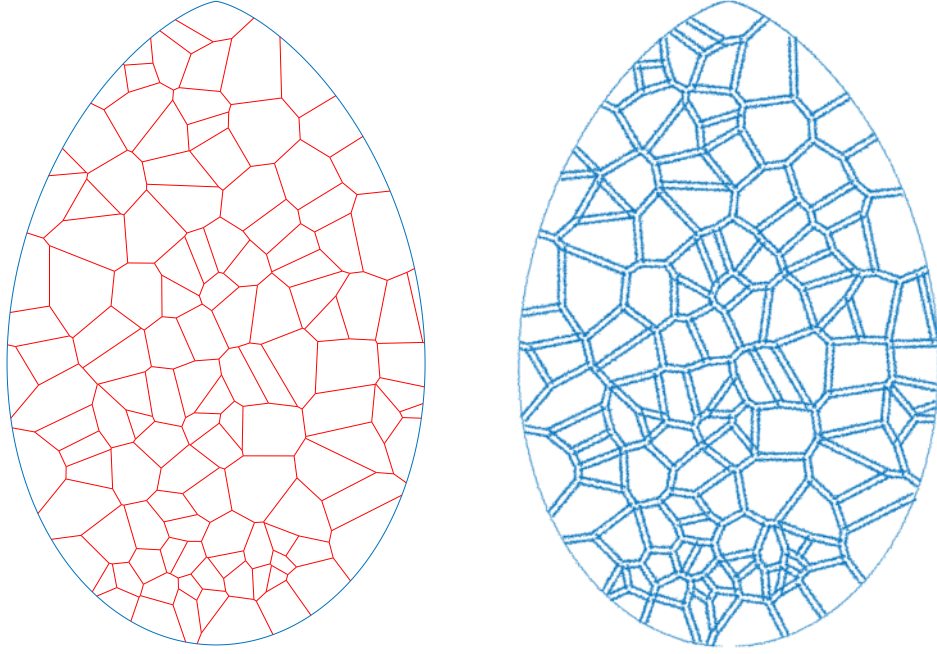
where N is the sample size, K the number of retained principal components ($K^{(S)}$ and $K^{(P)}$ for \mathbf{f}_S and \mathbf{f}_P respectively), and q_{Beta} denotes the quantile of the Beta distribution. Concerning the two Q charts, several studies on an approximate distribution of the Q statistic have been carried out (Jensen and Solomon (1972), Nomikos and MacGregor (1995)). When the size of the preliminary samples is unitary, the appropriate control limit is given by (Jensen and Solomon (1972), Zhu et al. (2007)), as

$$UCL_\alpha(Q) = \theta_1 \left\{ 1 - \theta_2 h_0 \left(\frac{1 - h_0}{\theta_1^2} \right) + \frac{\sqrt{z_\alpha(2\theta_2 h_0^2)}}{\theta_1} \right\}^{\frac{1}{h_0}},$$

z_α being the $(1 - \alpha)$ -quantile of the standard normal distribution, $h_0 = 1 - \frac{(2\theta_1\theta_3)}{3\theta_2^2}$, and $\theta_r = \sum_{j=K+1}^{\infty} \lambda_j^r$, $r = 1, 2, 3$, where K is the number of retained components and λ_j is the eigenvalue of the j -th component. The Q control chart is aimed at detecting possible deviations affecting PCs orthogonal to the ones retained and included into the estimation of the T^2 statistic. The Q control chart is commonly used as a complementary control chart for the error due to PCA reconstruction (see Colosimo and Pacella (2010) and literature cited therein), with the aim of preventing any information loss about the process.

4 Testing the power of the method: a 2D simulation study

In order to test the performance of the method, we propose a simulation framework based on the generation of 2D trabecular structures including both in-control and out-of-control realizations of shapes having a common nominal model. Random realizations of the simulated shape were constructed by means of a random Voronoi tessellation of the inner region of an egg-shaped curve, resembling a 2D version of the 3D-printed shapes presented in Section 2. A 2D model was used as it enables the simulation of complex shapes in a computationally feasible way without losing generality for testing the performances of the proposed method. Fig. 9a shows the original tessellation, where each strut is represented by a segment. Fig. 9b shows one realization of a simulated shape, where each strut is represented in terms of a point cloud. The realization in Fig. 9b was used as nominal shape for the comparison against all other realizations, both in-control and out-of-control.



(a) Analytical model

(b) Nominal model

Figure 9: Left panel: basic analytical model of all the simulated trabecular structures that were part of the study. It is completely defined by the analytical expression of the boundary and by the coordinates of the end points of the segments, representing struts. On the right panel an example of point cloud, which is actually the nominal model used in the simulation

In order to realize the nominal shape in Fig. 9b, two steps have been followed. First, each segment i in the analytical model was doubled, with a width w_i , generating the trabecular structure; all the resulting segments were then discretized considering a **fixed sampling density**, i.e., a constant number of equispaced points for unit length. This was achieved by assigning $n = 1000$ points to the longest segment in the analytical model in Fig. 9a, and all other segment were then discretized keeping the same density. Next, each point j in each segment i was perturbed orthogonally with respect to the segment i it belongs to, and independently from any other point, by a quantity

$$d_{ij} \sim N(0, \sigma_{ij})$$

Note that the parameter w_i controls the dimension of the struts, while σ_{ij} controls the regularity of the trabecular surface. For the generation of the nominal model, the parameters were set as constant over the entire structure, i.e., $w_i = \mu_{w_0} = 0.04$ and $\sigma_{ij} = \bar{\sigma} = 0.004$.

To mimic the actual variability of the production process, we simulated the in-control realizations by introducing a random perturbation on the parameters w_i and σ_{ij} . In particular, to simulate a new shape, we followed the same steps as described for the nominal model, but (i) the average width of the strut i was randomly and independently determined by sampling w_i from

$$w_i \sim N(\mu_{w_0}, \theta),$$

where $\theta = 0.008$, (ii) the “roughness” parameter σ_{ij} was randomly sampled according to

$$\sigma_{ij} \sim U(\bar{\sigma} - \epsilon, \bar{\sigma} + \epsilon),$$

where U denotes the uniform distribution and $\epsilon = 0.002$.

In order to simulate a dataset of real shapes, it is worth noting that each w_i is sampled independently for each strut in each realization of the shape, while σ_{ij} is sampled once for each realization and kept fixed for all struts in a given realization. In other words, the trabecular width w_i is a source of variability which adds randomness to struts dimension *within* a realization of the shape, while σ_{ij} introduces randomness *between* two different realizations.

Three scenarios representative of out-of-control geometrical distortions were generated by injecting three different kinds of defects in the simulation model, namely a variation of the strut width at different severity levels, the presence of missing struts, and the presence of irregularities on the trabecular surface. These three scenarios are described in Section 4.1, Section 4.2 and Section 4.3, respectively, while the corresponding results are shown in Section 4.4. In all the scenarios, 100 realizations of the 2D shapes were generated in each simulation run. 95 of them were generated by using the same parameters μ_{w_0} , θ , ϵ , $\bar{\sigma}$ and hence they are representative of in-control shapes, while the 5 remaining realizations simulate the out of control situation characterizing each scenario.

4.1 Scenario I: Shift in struts dimension

In the first scenario, the five out-of-control shapes were generated by injecting a shift $\Delta w > 0$ in the parameter that controls the mean strut width: so the width of each strut of an out-of-control shape is sampled from a normal distribution with mean $\mu_{w_1} = \mu_{w_0} + \Delta w$, all other parameters being kept fixed. The shift Δw was generated at 4 levels of severity, as reported in Table 1 in Section 4.4. Distances are computed with respect to the nominal model displayed in Fig. 9b. In Fig. 10 two examples of in-control and out-of-control shapes are provided.

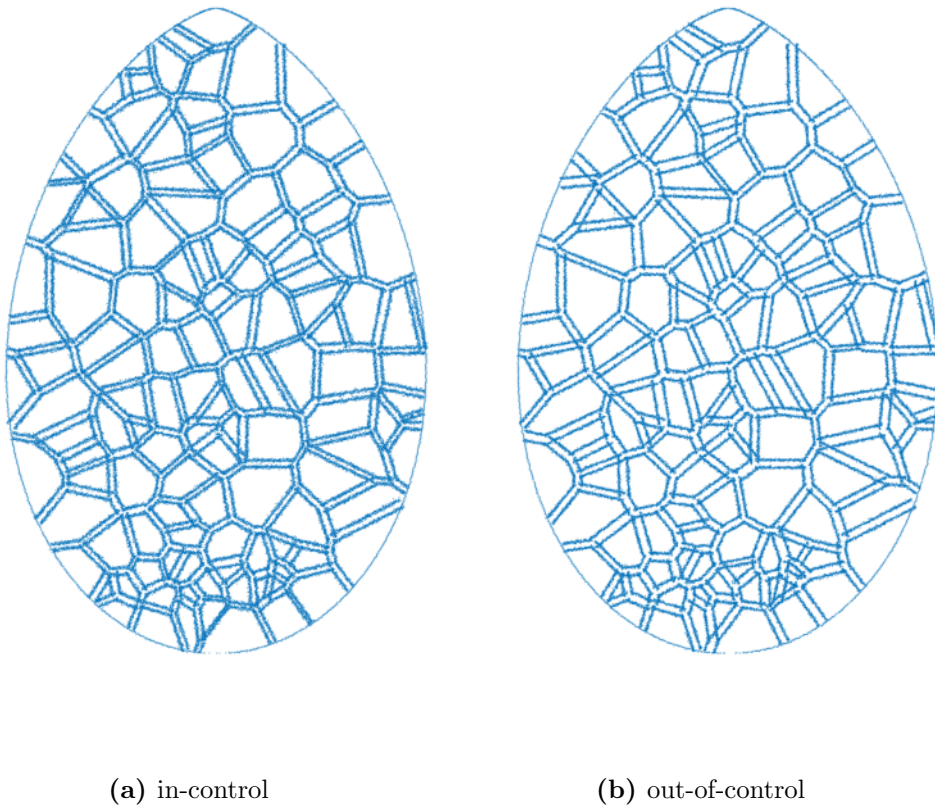


Figure 10: On the left panel an in-control shape: note the irregularity in the trabecular dimension, in comparison with the regularity of the nominal model shown in Fig. 9a. The right panel shows a out-of-control shape, with $\mu_{w_1} = 0.05$ versus $\mu_{w_0} = 0.04$ of the in-control shape.

4.2 Scenario II: missing struts

In Scenario II the injected defect of the five out-of-control shapes consisted of removing struts from the initial structure at random: for each defective shape, η struts are randomly chosen and eliminated, independently of all other shapes. As in the previous case, we simulated different severities, corresponding to different values of η . We show in Fig. 11 point clouds affected by this kind of anomaly, at the four simulated severities ($\eta = 1, 2, 3, 4$).

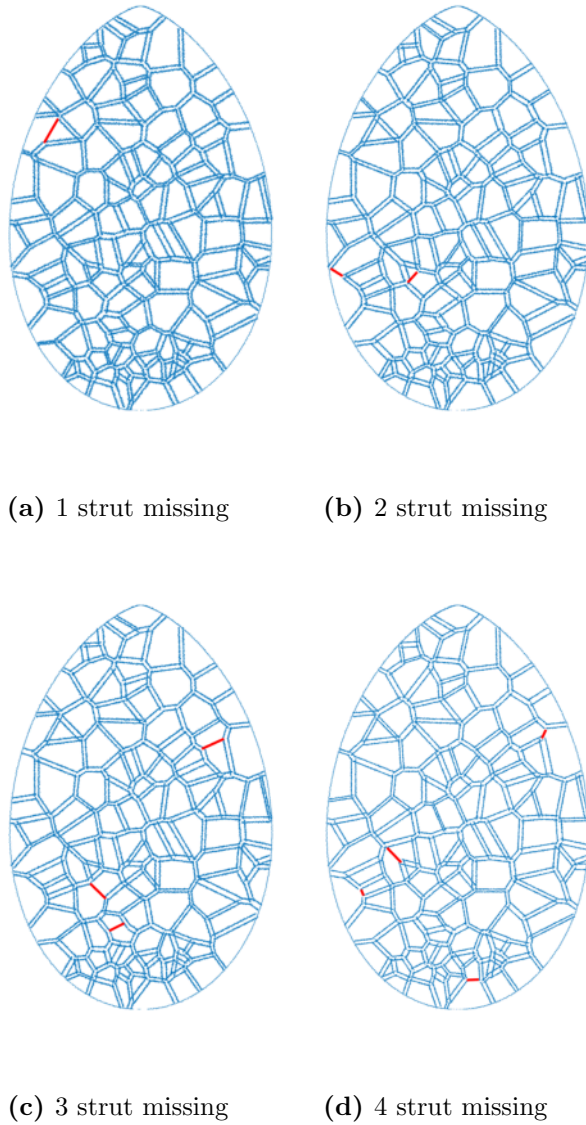
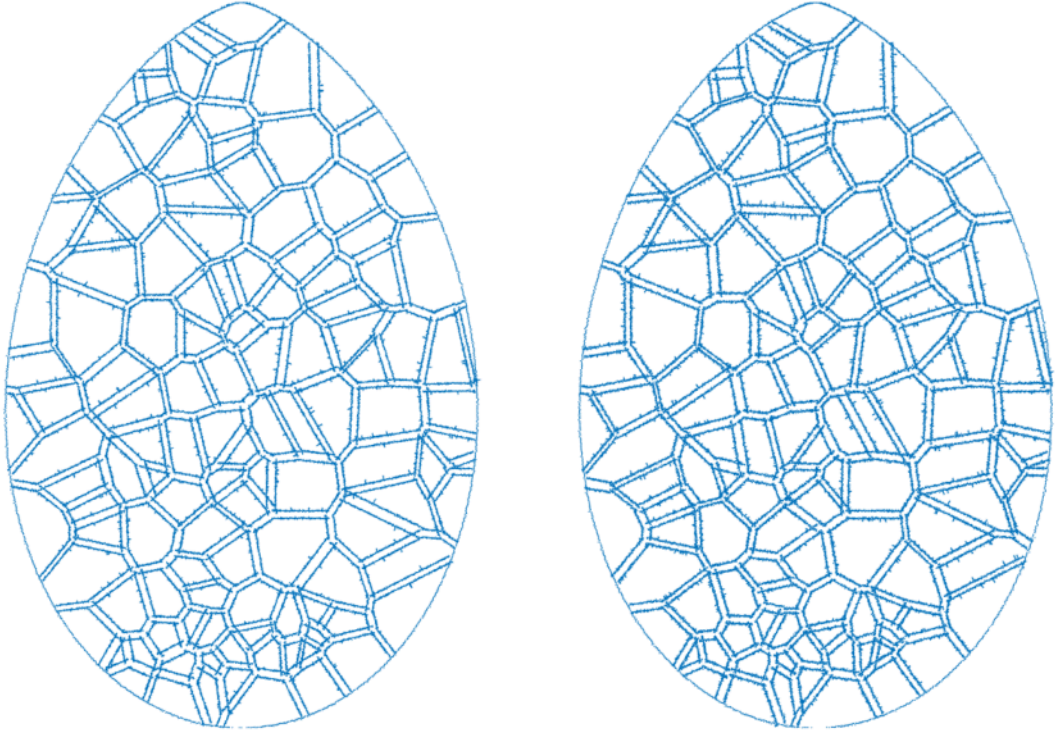


Figure 11: Realization of the shape with 1, 2, 3, 4 missing struts, which are marked in red.

As shown in Figure 11, the simulated defects are very local and difficult to identify from a visual inspection of the shape, since it does not modify substantially the overall structure. However, local defects like these are particularly critical in real life applications, as they may introduce weakness points in the structure, affecting its functionality. Each realization is simulated following exactly the same procedure of the in-control shapes, described in Section 4, and missing struts are chosen randomly.

4.3 Scenario III: irregularities on the trabecular surface

In this case, the five out-of-control elements are characterized by irregularities on the struts. This scenario was driven by the observation of real anomalies in 3D-printed shapes, similarly to the ones shown in Section 2. This kind of irregularities may be representative of an irregular material deposition, and hence they indicate a departure from an in-control production process. Irregularities are modeled as additional segments on the principal struts, as shown in Fig 12.



(a) Low severity

(b) High severity

Figure 12: Simulated out-of-control shapes of Scenario III, at different severities

As can be seen from Fig. 12, each additional segment is attached at a random location of a strut. The severity of the defect can be controlled by varying the maximum number ν of irregularities within a single strut. For example, the shape in Fig. 12a has been generated by setting $\nu = 6$ while 12b corresponds to $\nu = 10$. Given ν and a shape realization, for each strut the actual number of irregularities is sampled from a discrete uniform distribution $\mathcal{U}(0, 1, \dots, \nu)$. The length of each irregularity is a random fraction of the length of the strut: in particular, this fraction is sampled from a uniform distribution $\mathcal{U}(1/16, 1/20)$. Also in this case, we simulated four different severities, corresponding to different values of ν reported in Table 1 in Section 4.4.

4.4 Results

The aim of this Section is twofold: (i) estimating the detection power of the proposed method, and (ii) checking that the obtained results are coherent with the different geometric characteristics of the simulated defect, and hence that the analysis of both deviation maps is necessary.

In Fig. 13, Fig. 14, Fig. 15, we show an estimate of the detection power $1 - \beta$ of the method, where β is the type II error probability. The estimate is based on a 100 simulation run for each severity level in each scenario, for Scenario I, II and III respectively. For each scenario, three functions are plotted: the detection power of the method (panel (c)), and the detection power of the single couple (T^2, Q) of charts applied either to the \mathbf{f}_P densities (panel (a)) or \mathbf{f}_S densities (panel (b)). In Table 1 we summarize the severities for each scenario. In every simulation run, we set the nominal type I error probability α^* to 1%. Note that the first severity level of Scenario I corresponds to an in-control situation where all elements are conformal: hence, this level has also been placed as first level for the other Scenarios. Of course, the estimate corresponding to it is the empirical type I error probability.

Fig. 13 shows that in scenario I both $(T^{2(P)}, Q^P)$ and $(T^{2(S)}, Q^S)$ contribute to the overall detection power: this is reasonable, since the injected defect does not concern excessive or missing material, but a general over-sizing effect of all the struts that affects both \mathbf{f}_P and \mathbf{f}_S densities, as shown in Fig. 16. On the contrary, Fig. 14 shows that, in Scenario II, where we are simulating defects due to lacks of material, only the analysis of \mathbf{f}_S contributes to the identification of defective shapes: this is due to the fact that higher values of distances are associated to the points in the reference model belonging to the struts which are missing in the realized out-of-control shape, while the distances of the points of the realized shapes from the reference model cannot present any anomaly, as shown in Fig. 17. As regards Scenario III, where we simulate irregularities due to excessive material, Fig. 15 shows an opposite situation: only the analysis of \mathbf{f}_P allows identifying defective shapes, since the points associated to high values of distances are the ones belonging to the irregular additional segments in the realized out-of-control shapes (see Fig. 18).

The method seems to be highly sensitive when missing or additional structures are con-

cerned: the power functions regarding Scenario II and III are quite far from zero even when just one strut is missing or few irregularities are added. This may be due to the fact that such anomalies do not cause big shifts of the mean, median or in general of the structure of the distance density, but rather they introduce small peaks in the density over values of distance where in-control densities are monotonically decreasing and, more important, vanishing. The scalar product in B^2 (defined in Section 3), which fully defines the geometrical features of the space and consequently of the principal components, is quite sensitive to this kind of difference in the densities, because it is based on (log-) ratios of probabilities, varying sharply when one of the density vanishes while the other does not.

Scenario	Varying parameter	Severity I	Severity II	Severity III	Severity IV
I	Δw	0	0.005	0.01	0.015
II	η	1	2	3	4
III	ν	2	4	6	10

Table 1: Different severity levels for the scenarios

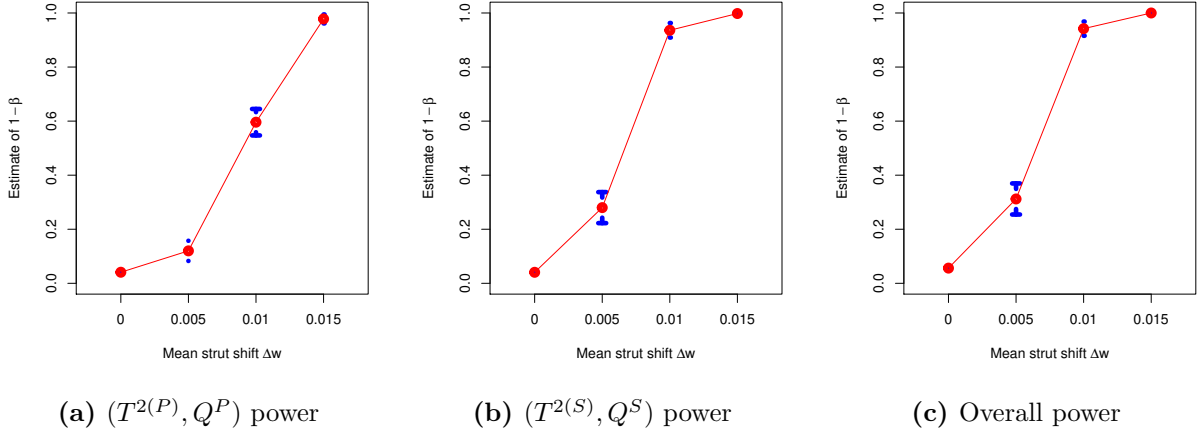


Figure 13: Plot of $1 - \beta$ for the couple $(T^{2(P)}, Q^P)$ (a), for $(T^{2(S)}, Q^S)$ (b) and for the method (c), for Scenario I, with 99% Confidence Intervals, computed on 100 simulation runs for each severity level. The different severities are $\Delta w = 0, 0.005, 0.01, 0.015$.

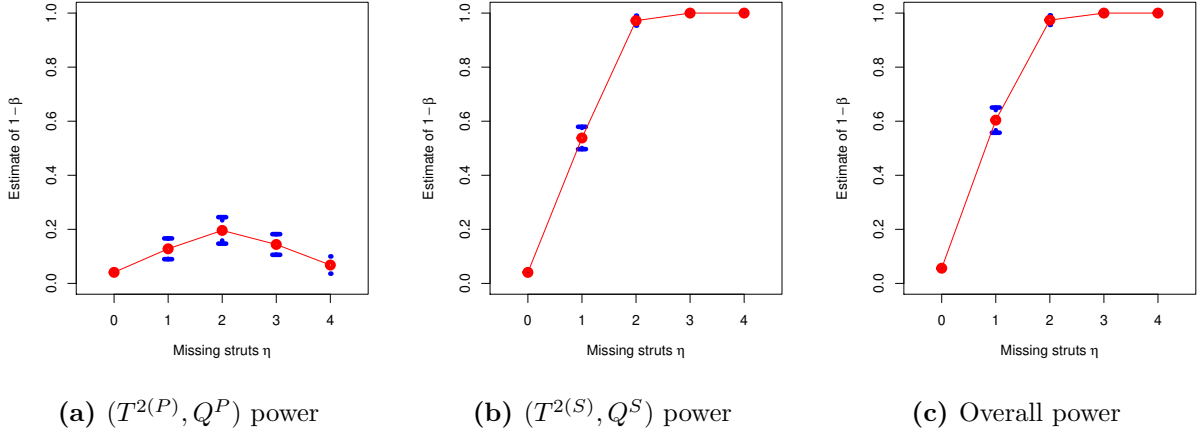


Figure 14: Plot of $1 - \beta$ for the couple $(T^{2(P)}, Q^P)$ (a), for $(T^{2(S)}, Q^S)$ (b) and for the method (c), for Scenario II, with 99% Confidence Intervals, computed on 100 simulation runs for each severity level. The different severities are $\eta = 0, 1, 2, 3$ or 4 missing struts.

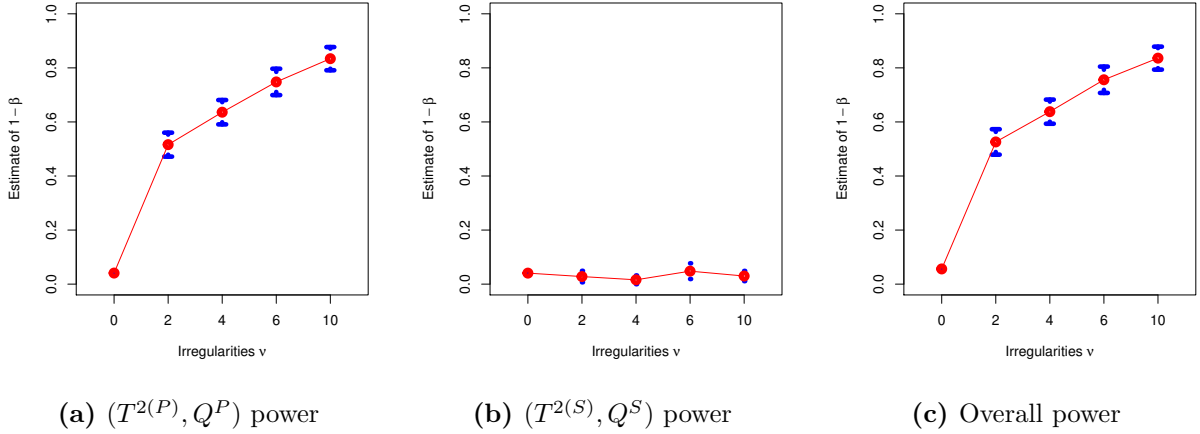


Figure 15: Plot of $1 - \beta$ for the couple $(T^{2(P)}, Q^P)$ (a), for $(T^{2(S)}, Q^S)$ (b) and for the method (c), for Scenario III, with 99% Confidence Intervals, computed on 100 simulation runs for each severity level. The different severities are given by $\nu = 0, 2, 4, 6, 10$, the maximum number of additional segments attached at each strut.

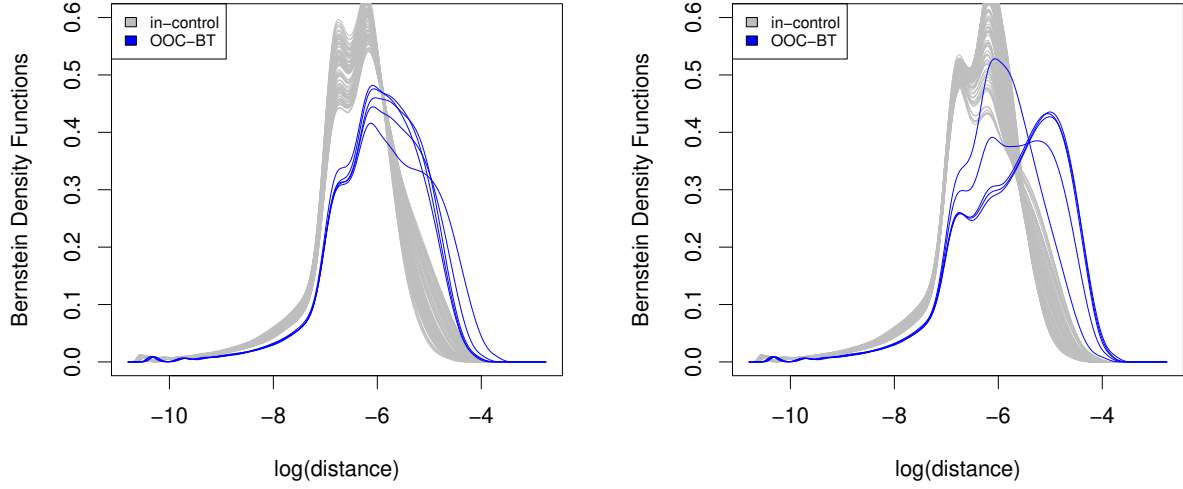


Figure 16: An example of dataset \mathbf{f}_P (left) and \mathbf{f}_S (right) from a simulation run of Scenario I, with the highest out-of-control severity, $\mu_{s_1} = 0.055$. Densities corresponding to defective shapes are drawn in blue.

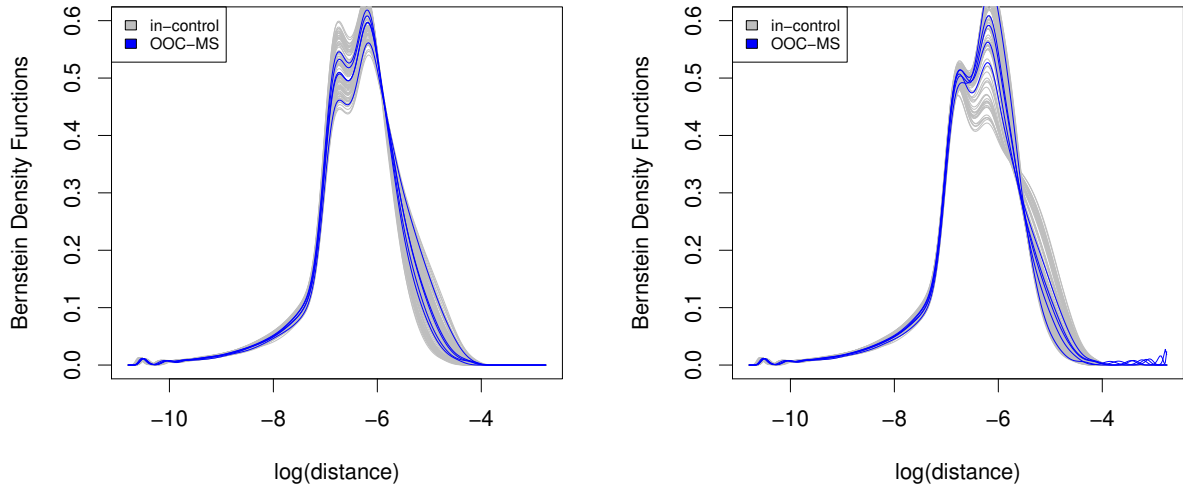


Figure 17: An example of dataset \mathbf{f}_P (left) and \mathbf{f}_S (right) from a simulation run of Scenario II, where the out-of-control shapes, whose densities are drawn in blue, have $\eta = 2$ missing struts. Note the behavior of the right tail of the densities in \mathbf{f}_S representing defective shapes.

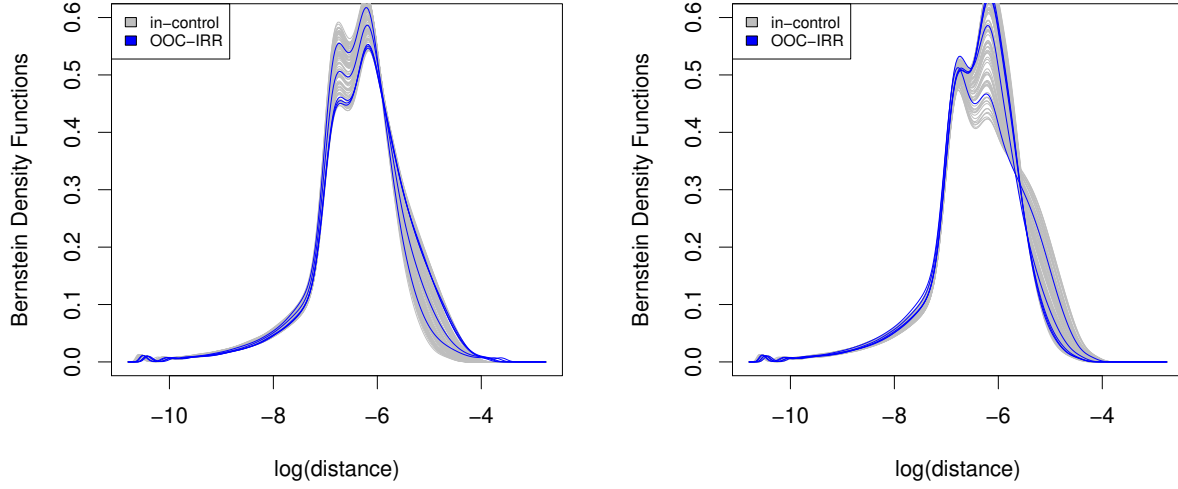


Figure 18: An example of dataset \mathbf{f}_P (left) and \mathbf{f}_S (right) from a simulation run of Scenario III where the out-of-control shapes, drawn in blue, are characterized by $\nu = 10$. Note again the tails of the blue densities in the chart on the left.

An aspect which is worth investigating is how the performance of the method is affected by variations in the density of the clouds – in our case, this corresponds to a variation of the parameter n mentioned at the beginning of Section 4, which was set to 1000 and kept fixed in the previous scenarios. One could guess that the performance of the method remains stable as long as the sampling remains uniform and does not become coarse to the extent that a loss of information about some geometrical features of the original shape is caused, or the convergence of the Bernstein estimators is prevented. We show an example in Figure 19, which is relative to a simulation run identical to Scenario II, apart from the sampling density, which was reduced to $n = 500$. As expected, apart from a slight variation at the beginning of the curve, the detection power appears stable.

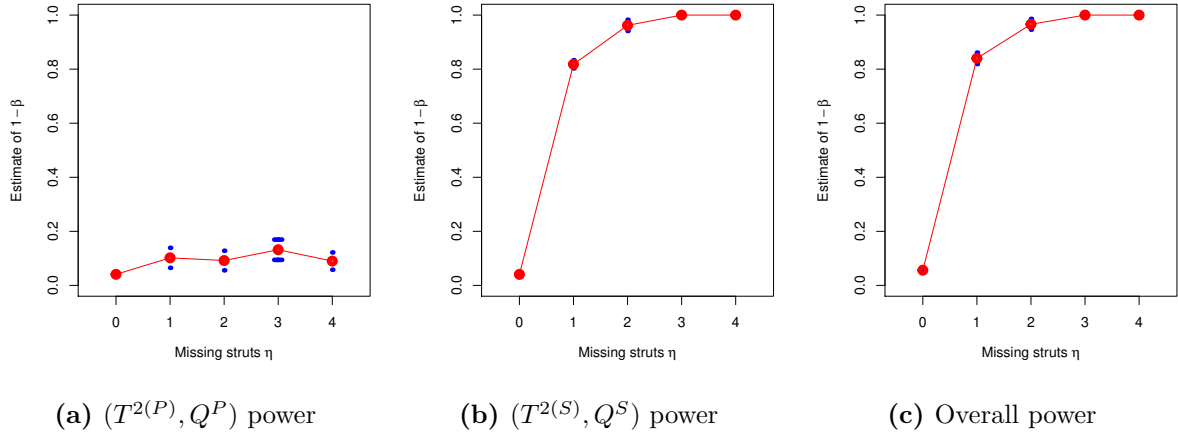


Figure 19: Plot of $1 - \beta$ for the couple $(T^{2(P)}, Q^P)$ (a), for $(T^{2(S)}, Q^S)$ (b) and for the method (c), for Scenario II with reduced sampling density $N = 500$, with 99% Confidence Intervals, computed on 100 simulation runs for each severity level. The different severities are $\eta = 0, 1, 2, 3$ or 4 missing struts

5 Application to the real case study

5.1 Experimental settings

The eggs were 3D-printed by using an FDM Sharebot 42 system. The material was PLA (Polylactic Acid), a thermoplastic polymer commonly used in FDM, which provides a good surface finishing with the further advantage of being a non-toxic material. No supporting structure was needed since the geometry is self-supporting, despite the presence of low-angle overhang struts. The filament was extruded at a temperature of $^{\circ}\text{C}$ 230 leading to an extruded diameter of 0.40 mm. In the following, the out-of-control egg characterized by local excess of material is referred to as OOC-EM (out-of-control exceeding material) and the egg with the missing strut as OOC-MS (out-of-control missing strut). The X-ray computed tomography of all the eggs was performed by using a North Star IX 25 system, with a voxel size of about 52 microns.

Fig. 20 shows the densities \mathbf{f}_P (left panel) and \mathbf{f}_S (right panel) obtained by smoothing the PDFs using Bernstein polynomials.

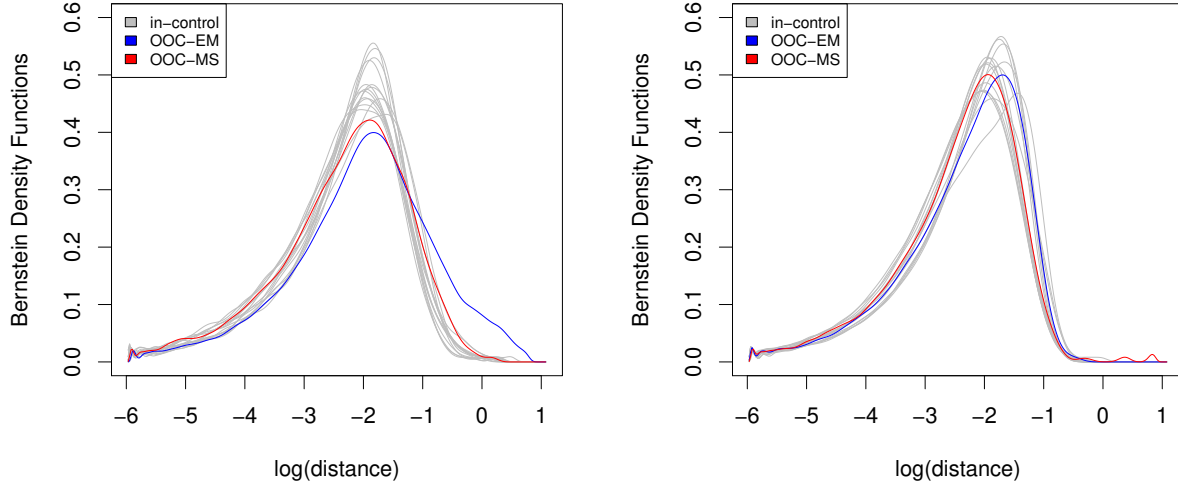


Figure 20: Dataset \mathbf{f}_P , on the left, and \mathbf{f}_S , on the right. The Bernstein estimators are smooth and well suited to estimate densities with compact support. In control elements are shown with gray lines while out-of-control elements are shown in red (OOC-MS) and blue (OOC-EM). The only PDF which appears to be clearly different from the others at first glance is the one representing OOC-EM, in the dataset \mathbf{f}_P

Fig. 20 (left panel) shows that when the PDF of the distances of the real shape from the nominal shape, i.e., \mathbf{f}_P , is considered, the presence of excessive material (OOC-EM) inflates the right tail of the PDF. Indeed, locations of the real shape with exceeding material exhibit a larger distance from the nominal shape, modifying the overall PDF. On the contrary, the presence of excessive material has no effect on the PDF of the distances of the nominal shape from the real shape, i.e., \mathbf{f}_S (Fig. 20, right panel). Indeed, in this case, no distance value can be associated to points that do not belong to the nominal shape, and hence the contribution of regions of the real shape characterized by exceeding material is filtered out. An opposite effect regards the other type of out-of-control distortion, i.e., OOC-MS. In this case, the effect is captured by the \mathbf{f}_S density only, where the missing strut caused an inflation of the right tail of the PDF, whereas no evident effect on the \mathbf{f}_P density was observed. These two out-of-control examples represent two opposite and complementary conditions, which highlight the importance of considering the directional nature of the Hausdorff distance measures to estimate deviations from a nominal model.

Fig. 21 shows the plots of the mean density perturbed along the direction of the first Principal Component, which explains more than the 98% of the total variance, for both datasets. Fig 21, left panel, shows the perturbations on the first PC on the \mathbf{f}_P dataset, whereas Fig. 21, right panel, shows the perturbations on the \mathbf{f}_S dataset. Fig. 10 shows the boxplots and individual values of the scores for the projections of the \mathbf{f}_P (left panel) and \mathbf{f}_S (right panel) densities onto the direction of their first PC respectively.

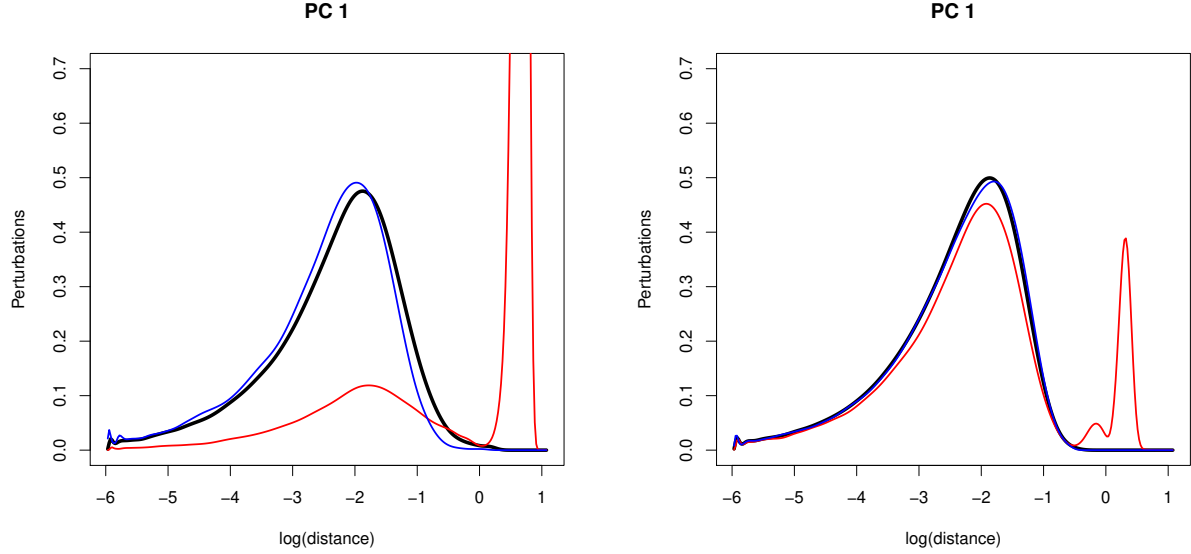
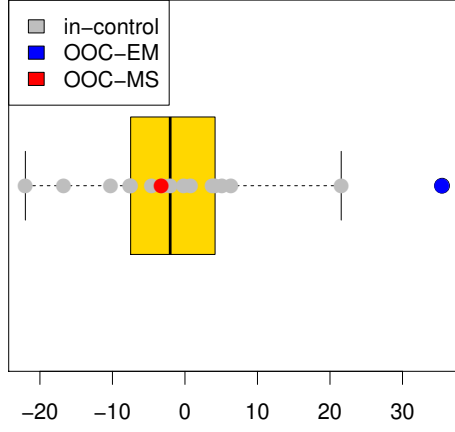
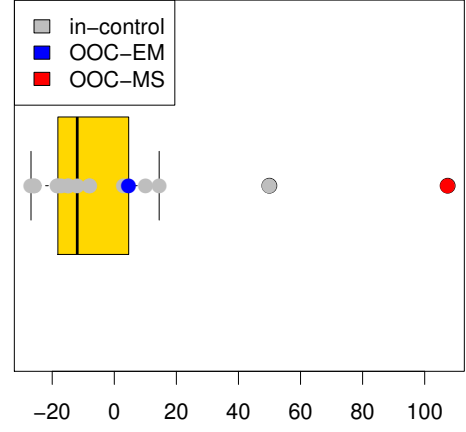


Figure 21: Mean density (black) perturbed by $3\sqrt{\lambda_1}\zeta_1$ (red) and $-3\sqrt{\lambda_1}\zeta_1$ (blue), for the dataset \mathbf{f}_P on the left, and for the dataset \mathbf{f}_S on the right. The perturbations are performed using the operation of sum and product defined for B^2 , as detailed in Section 3

In both the cases shown in Fig. 21, the effect of the two out-of-control shapes affects the right tail of the distribution. For the first PC of the \mathbf{f}_P densities, the perturbation is driven by the anomaly of the OOC-EM shape, whereas the perturbation along the first PC for the \mathbf{f}_S densities is driven by the OOC-MS shape, as shown in Fig. 20. Fig. 21 and Fig. 22 shows that the first PC is sufficient to capture such effect for both the out-of-control cases and both the densities.



(a) dataset \mathbf{f}_P scores



(b) dataset \mathbf{f}_S scores

Figure 22: Boxplot of the scores, along the first PC, of the element of the dataset \mathbf{f}_P (a) and \mathbf{f}_S (b). The positive shift of the defective realization of the shapes can be appreciated.

Fig. 23 shows the T^2 control chart and the Q control chart, in relation to the scores of the SFPCA applied to the \mathbf{f}_P ((a) and (b)) and \mathbf{f}_S ((c) and (d)). The type I error for each chart has been set starting from $\alpha^* = 0.01$, as detailed in Section 3.

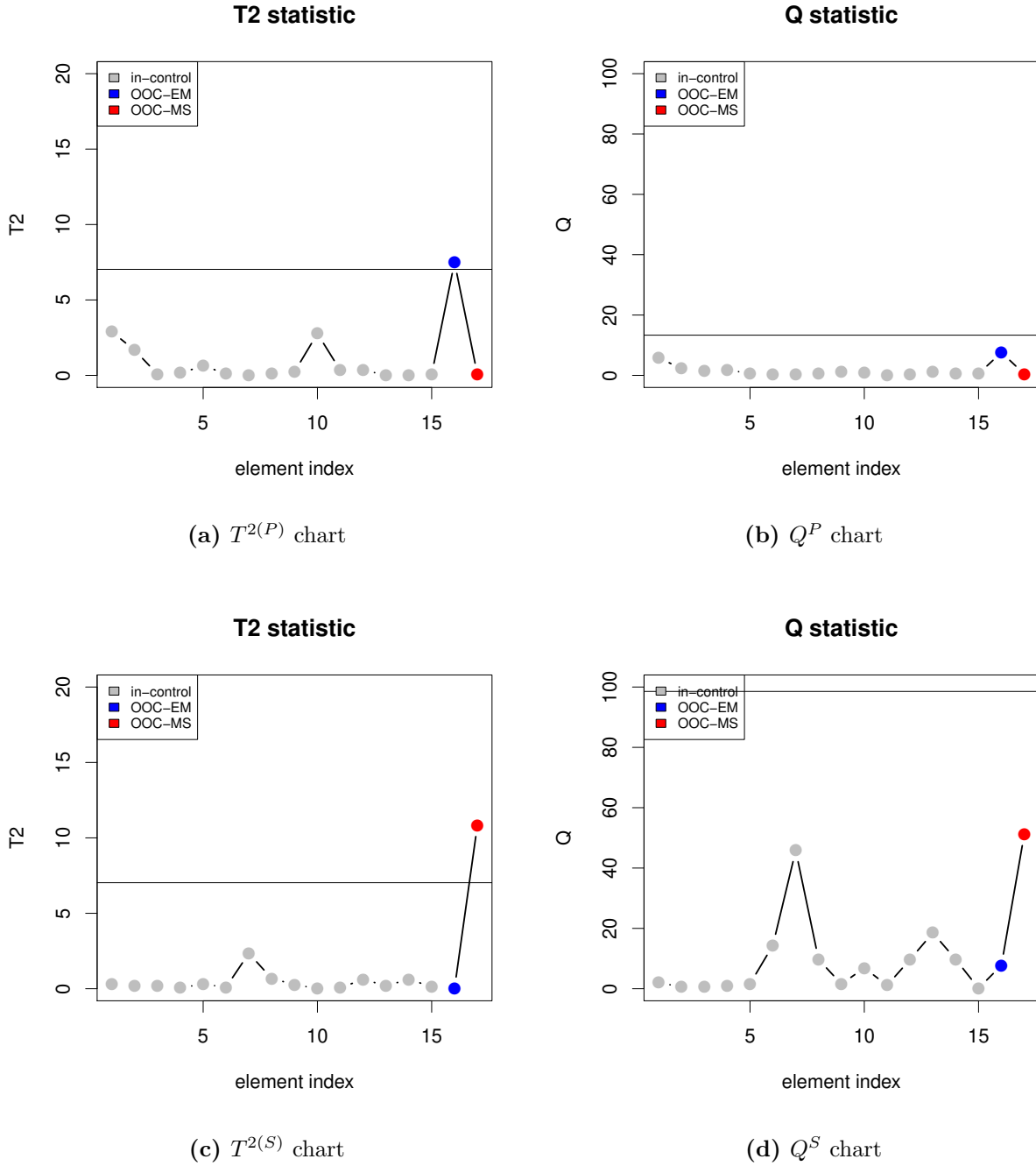


Figure 23: Control charts for \mathbf{f}_P (a, b) and for \mathbf{f}_S (c, d). In both cases, the number of PCs used to compute the statistic is 1 (more than 98% of explained variance). The straight black line represents the UCL in all cases.

Fig. 23 confirms that both the defective shapes are signaled as out-of-control by the proposed approach, but each couple (T^2, Q) of control charts signals only one out-of-control

depending on the effect of the geometric distortion on the corresponding deviation map. Due to the nature of the defective shapes considered in this study, the two couples of control charts signal complementary out-of-control contaminations of the Phase I dataset. The $T^{2(P)}$ control chart signals a contamination caused by excessive material whereas the $T^{2(S)}$ control chart signals a contamination caused by a missing strut, coherently with the interpretations given to the perturbations in Fig. 21. Generally speaking, the use of two distinct couples of control charts eases the interpretation and diagnosis of signaled observations. Moreover, if a shape exhibits a geometrical distortion involving both excess and lack of material in different locations, it is expected to be signaled by both the couples, since it affects both the \mathbf{f}_P and \mathbf{f}_S deviations.

Although, as discussed in the Introduction, it is not immediate to identify “standard” competitors for the proposed approach when the target objects are in a complex manifold configuration as in the presented real case study, it is possible to refer to the common practice adopted in industry. In the framework of additively manufactured complex shapes, the most common approach consists of computing individual and global descriptors, like the overall volume of the manufactured shape (that can be derived by standard software for mesh analysis) or the density of the part (that can be measured via the Archimede’s method). By monitoring such global descriptors it is possible to determine the presence of internal porosity or anomalous lack or excess of material. In order to assess the benefits of the proposed approach against such industrial practice, the volume of each 3D-printed shape was estimated through a mesh reconstruction to obtain a “watertight” mesh. To this aim, the screened Poisson surface reconstruction approach was applied (Kazhdan and Hoppe (2013)). The estimated volume of the mesh was then used to design a univariate control chart for the shapes included into the real case study. The result is shown in Fig. 24.

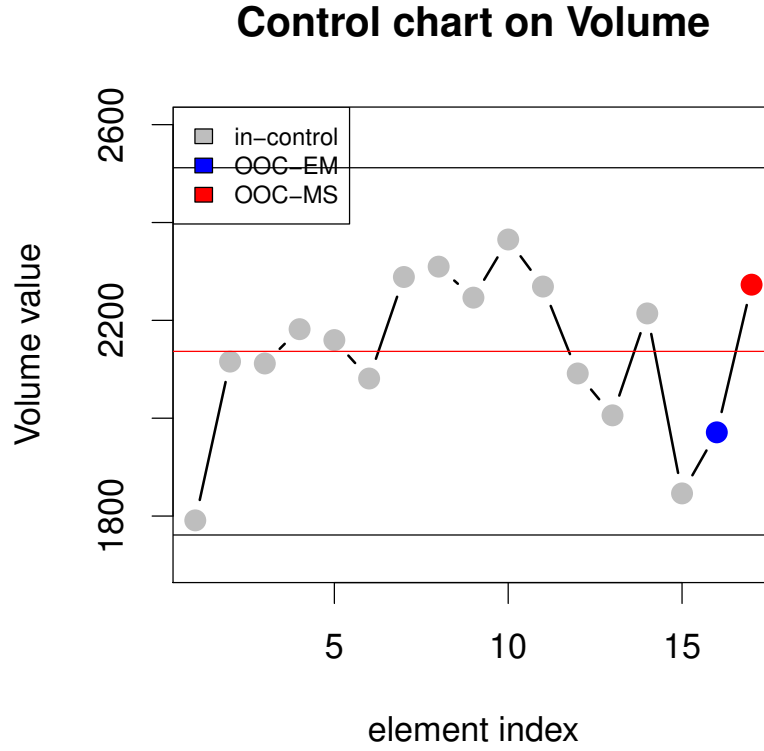


Figure 24: Individual control chart for the “volume” descriptor. The red line represent the mean volume, while the control limits are set considering a deviation of $\pm z_{0.99}\sigma$, σ being the volume standard deviation.

Fig. 24 shows that relying on a global descriptor, despite being representative of the common industrial practice, is not suitable to identify out-of-control deviations like the ones discussed in this study. Indeed, the local variability introduced by the local defects represents a minor contribution to the overall natural variability of the shapes’ volume.

6 Conclusion

The continuous development of digital manufacturing makes available novel levels of shape complexity and big-data unstructured point clouds in the Industry 4.0 scenario. This imposes novel challenges in the design and use of statistical quality modeling and monitoring methods, as traditional approaches may not be applicable. The method proposed in this study is aimed at identifying anomalous geometrical structures and out-of-control depar-

tures from a natural shape variability by investigating the deviation between each inspected shape and its nominal (prototype) model. It applies to full 3D geometry reconstructions in the form of triangulated shapes transformed into a point cloud format.

Differently from methods previously proposed in the literature, we showed that the map of deviation between the inspected shape and the nominal one has an intrinsic directional nature. Coping with such directional property is of fundamental importance to avoid information losses in the characterization of the natural shape variability, and to properly capture any out-of-control deviation pattern. The results showed that the parallel monitoring of two deviation maps, one projected on the nominal shape and one projected on the inspected shape, allows one to signal geometrical distortions that could not be detected considering one deviation map only. The study also proposed a method to synthesize the information content enclosed in the two deviation maps in the form of two marginal PDFs, leading to a profile monitoring scheme that can be applied to any complex shape for which a reference prototype (i.e., the nominal model) is available.

This approach is particularly suitable for additive manufacturing applications, where a near net shape, whose complexity is enabled by a completely new design freedom, is produced layer by layer starting from a nominal CAD model. The proposed approach was tested considering a paradigmatic example, consisting of an egg-shaped trabecular shell. However, the method is completely general in nature, and hence it can be extended to even more complex shapes involving both external and internal geometrical features. **X-ray CT is the only non-destructive inspection method that allows the reconstruction of both outer and inner surfaces of a 3D object. Due to the high resolution of this kind of measurement, passing from a shape representation in terms of vertices and edges (mesh) to vertices only (point cloud) is expected to yield negligible loss of information. Nevertheless, the proposed approach could be generalized to exploit the entire information enclosed by the mesh representation, by including both edges and vertices into the analysis.** Even if the approach originates in the context of statistical quality monitoring of 3D printed complex shapes, it can be applied to any unstructured noisy point clouds or meshes on complex manifolds, arising in many different application contexts, as computational biology and medicine (e.g., Lila et al. (2016), Sangalli et al. (2014)). Future studies will investigate the

performances of the method in the presence of different kinds of shapes and its possible extension to Phase II control chart implementations. Another relevant future improvement of the proposed technique regards the inclusion of a spatial model approach to preserve the spatial dependency information that is lost when PDF curves are used as synthetic signatures of the deviation maps. This is expected to enhance the capability of detecting small shifts affecting only locally the overall shape, and hence it represents a natural and interesting extension of the proposed approach. Eventually, it is worth noting that different types of defects could be present, at the same time, in different locations of the part. This is expected to cause shifts in the monitored statistics, and we believe that addressing explicitly the bi-directional nature of deviations from a nominal shape enhances the capability of detecting complex and hybrid defectiveness conditions. Therefore, future experimentation could deepen the performance of the method for different geometrical errors and different shapes.

7 Acknowledgments

This work has been supported by ACCORDO Quadro ASI-POLIMI “Attività di Ricerca e Innovazione” n. 2018-5-HH.0, collaboration agreement between the Italian Space Agency and Politecnico di Milano.

A Sensitivity of the ICP algorithm and number of landmarks

The registration approach adopted in this study involves a first rough alignment based on manually selected landmarks, followed by the fine registration via ICP. The algorithm is known to be affected by the initialization condition, which is the result of the first landmark-based step. By placing any manufactured shape with the same orientation during the X-ray CT scan and by keeping fixed the number and location of initial landmarks for any measured shape, the initialization condition can be made repeatable. Under these premises, the residual misalignment error is assumed to be repeatable from one shape to another. In this

study, we carried out a sensitivity analysis to quantify the effect of varying initial conditions, i.e., different number and location of manually selected landmarks, together with their contribution to the overall natural variability of the 3D-printed shapes included in the real case study. We tested four different configurations of landmarks, consisting respectively of 4, 8, 12 and 16 manually selected pairs of vertices in the measured and prototype meshes. The landmarks were always selected to be about equally spaced apart on the entire surface of the shape. Four in-control 3D-printed shapes were randomly selected for this analysis and the same location of landmarks was applied to each of them. The average Hausdorff distance between each pair of aligned shapes was used to compare the effects of different landmark configurations. Figure 25 shows the 95% confidence intervals of the average Hausdorff distance from the prototype by varying the number and location of landmarks. All tested landmark configurations yielded no statistically significant difference on this average distance. In addition, a mixed effects regression model was fitted to the measured average Hausdorff distances, where the landmark configuration was included as a fixed factor, whereas the different shapes were included as random factor. Table 2 summarizes some results. In particular, Table 2 shows that the between-shape variability accounts for about 88.4% of the overall variability, whereas the different landmark configurations accounts for the remaining 11.6%. This confirms that the contribution of the alignment step to the natural variability of the monitored statistics is small with respect to actual between shape variability.

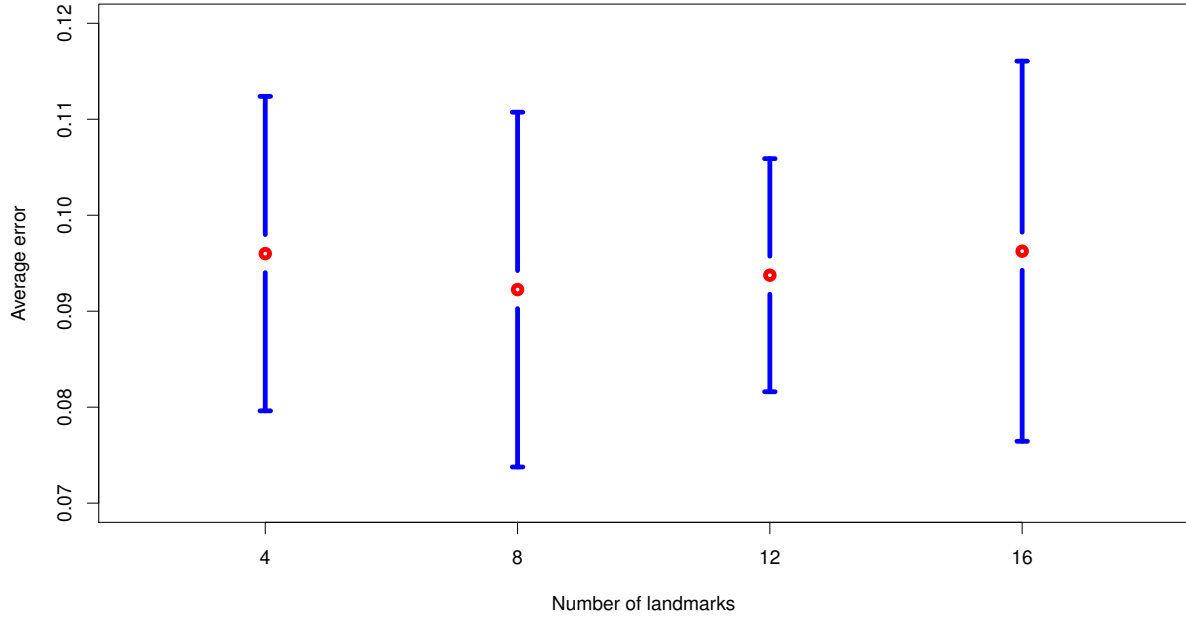


Figure 25: 95% confidence intervals for the average Hausdorff distance between reconstructed shapes and the nominal prototype for different registration landmark selections

Source of variability	Factor type	Variance	% of Total Variance	R^2_{adj}
Egg-shape	Random effect	0.00010	88.44%	89.39%
Landmarks	Fixed effect	0.000013	11.56%	

Table 2: Summary of results from the mixed effects regression model for the average Hausdorff distance response.

B Code for Simulation studies

Code for the reproducibility of the simulation studies is available at <https://github.com/RiccardoScimone/Stat-mod-compl-shapes.git>

References

Aitchison, J. (1986). *The Statistical Analysis of Compositional Data*. Chapman & Hall, Ltd. London.

- Alt, H., Bra, P., Godau, M., Knauer, C., and Wenk, C. (2003). Computing the hausdorff distance of geometric patterns and shapes. *Discrete and Computational Geometry*, 25:65–76.
- Aspert, N., Santa-cruz, D., and Ebrahimi, T. (2002). Mesh: Measuring errors between surfaces using the Hausdorff distance. In *Proceedings of the IEEE International Conference in Multimedia and Expo (ICME)*, volume 1, pages 705 – 708.
- Babu, G., Canty, A., and Chaubey, Y. (2002). Application of bernstein polynomials for smooth estimation of a distribution and density function. *Journal of Statistical Planning and Inference*, 105:377–392.
- Barceló-Vidal, C. (2001). Mathematical foundations of compositional data analysis. In *Proceedings of IAMG’01 - The sixth annual conference of the International Association for Mathematical Geology*, page 20.
- Besl, P. and McKay, H. (1992). A method for registration of 3D shapes. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 14:239–256.
- Chen, Q., Wynne, R., Goulding, P., and Sandoz, D. (2000). The application of principal component analysis and kernel density estimation to enhance process monitoring. *Control Engineering Practice*, 8:531–543.
- Cignoni, P., Callieri, M., Corsini, M., Dellepiane, M., Ganovelli, F., and Ranzuglia, G. (2008). MeshLab: an Open-Source Mesh Processing Tool. In Scarano, V., Chiara, R. D., and Erra, U., editors, *Eurographics Italian Chapter Conference*. The Eurographics Association.
- Colosimo, B., Cicorella, P., Pacella, M., and Blaco, M. (2014). From profile to surface monitoring: Spc for cylindrical surfaces via Gaussian processes. *Journal of Quality Technology*, 46:95–113.
- Colosimo, B., Huang, Q., Dasgupta, T., and Tsung, F. (2018). Opportunities and challenges of quality engineering for additive manufacturing. *Journal of Quality Technology*, 50:233–252.

- Colosimo, B. M. and Pacella, M. (2010). A comparison study of control charts for statistical monitoring of functional data. *International Journal of Production Research*, 48(6):1575–1601.
- Conci, A. and Kubrusly, C. S. (2018). Distance between sets - a survey.
- del Castillo, E., Colosimo, B., and Tajbakhsh, S. (2014). Geodesic Gaussian processes for the parametric reconstruction of a free-form surface. *Technometrics*, 57:87–99.
- Egozcue, J. J., L. Díaz-Barrero, J., and Pawłowsky-Glahn, V. (2006). Hilbert space of probability density functions based on Aitchison geometry. *Acta Mathematica Sinica, English Series*, 22:1175–1182.
- Eiter, T. and Mannila, H. (1999). Distance measures for point sets and their computation. *Acta Informatica*, 34:109–133.
- Filzmoser, P., Hron, K., and Templ, M. (2018). *Applied Compositional Data Analysis*. Springer Series in Statistics. Springer.
- Galimberti, G. (2018). *From metallic powder to the object. Digital aesthetic of new products obtained by Selective Laser Melting process*. PhD thesis, Politecnico di Milano.
- Gardner, A., Kanno, J., Duncan, C., and Selmic, R. (2014). Measuring distance between unordered sets of different sizes. pages 137–143.
- Gerard van den Boogaart, K., Egozcue, J. J., and Pawłowsky-Glahn, V. (2010). Bayes linear spaces. *SORT*, 34:201–222.
- Gibson, I., Rosen, D. W., Stucker, B., et al. (2014). *Additive manufacturing technologies*, volume 17. Springer.
- Górski, F., Kuczko, W., and Wichniarek, R. (2013). Influence of process parameters on dimensional accuracy of parts manufactured using fused deposition modelling technology. *Advances in Science and Technology Research Journal*, 7:27–35.
- Hausdorff, F. (1914). *Grundzüge der mengenlehre*. Leipzig, Von Veit.

- Henrikson, J. (1999). Completeness and total boundedness of the Hausdorff metric. *MIT Undergraduate Journal of Mathematics*, 1:69:80.
- Horváth, L. and Kokoszka, P. (2012). *Inference for Functional Data with Applications*. Springer Series in Statistics. Springer New York.
- Hron, K., Menafoglio, A., Templ, M., Hrušová, K., and Filzmoser, P. (2014). Simplicial principal component analysis for density functions in bayes spaces. *Computational Statistics & Data Analysis*, 94:330–350.
- Jensen, D. R. and Solomon, H. (1972). A gaussian approximation to the distribution of a definite quadratic form. *Journal of the American Statistical Association*, 67(340):898–902.
- Jorge, V. C. (2018). AM technological development in airbus defence & space system. Bilbao-BEC, Spain. Addit 3D.
- Kang, Y. and Qiu, P. (2014). Jump detection in blurred regression surfaces. *Technometrics*, 56:539–550.
- Kazhdan, M. and Hoppe, H. (2013). Screened poisson surface reconstruction. *ACM Trans. Graph.*, 32(3).
- Kwan-Ho Lin, Baofeng Guo, Kin-Man Lam, and Wan-Chi Siu (2001). Human face recognition using a spatially weighted modified Hausdorff distance. In *Proceedings of 2001 International Symposium on Intelligent Multimedia, Video and Speech Processing. ISIMP 2001 (IEEE Cat. No.01EX489)*, pages 477–480.
- Lila, E., Aston, J., and Sangalli, L. (2016). Smooth principal component analysis over two-dimensional manifolds with an application to neuroimaging. *The Annals of Applied Statistics*, 10:1854–1879.
- Menafoglio, A., Grasso, M., Secchi, P., and Colosimo, B. (2018). Profile monitoring of probability density functions via simplicial functional pca with application to image data. *Technometrics*, 60:497–510.

- Mémoli, F. and Sapiro, G. (2004). Comparing point clouds. In *SGP '04: Proceedings of the 2004 Eurographics/ACM SIGGRAPH Symposium on Geometry Processing*, volume 71, pages 33–42.
- Montgomery, D. C. (2009). *Introduction to Statistical Quality Control. 6th Edition*. John Wiley & Sons, New York.
- Nomikos, P. and MacGregor, J. F. (1995). Multivariate spc charts for monitoring batch processes. *Technometrics*, 37(1):41–59.
- Pawlowsky-Glahn, V., Egozcue, J. J., and Van den Boogaart, K. (2014). Bayes Hilbert spaces. *Australia and New Zealand Journal of Statistics*, 56:171–194.
- Ramsay, J. and Silverman, B. W. (2005). *Functional Data Analysis*. Springer.
- Sangalli, L. M., Secchi, P., and Vantini, S. (2014). AneuRisk65: A dataset of three-dimensional cerebral vascular geometries. *Electron. J. Statist.*, 8(2):1879–1890.
- Seal, A. and Bhowmick, A. (2017). Performance analysis of iterative closest point (icp) algorithm using modified hausdorff distance. *International Research Journal of Engineering and Technology*, 04(07):2794–2800.
- Sikaroudi, A. E., Welch, D. A., Woehl, T. J., Faller, R., Evans, J. E., Browning, N. D., and Park, C. (2018). Directional statistics of preferential orientations of two shapes in their aggregate and its application to nanoparticle aggregation. *Technometrics*, 60(3):332–344.
- Silverman, B. (2018). *Density Estimation for Statistics and Data Analysis*.
- Stankus, S. and Castillo-Villar, K. (2018). An improved multivariate generalised likelihood ratio control chart for the monitoring of point clouds from 3D laser scanners. *International Journal of Production Research*, 57:1–12.
- Syam, W. P., Jianwei, W., Zhao, B., Maskery, I., Elmadih, W., and Leach, R. (2018). Design and analysis of strut-based lattice structures for vibration isolation. *Precision Engineering*, 52:494–506.

- Valerga, A., Batista Ponce, M., Salguero, J., and Girot, F. (2018). Influence of pla filament conditions on characteristics of fdm parts. *Materials*, 11:1322.
- Wang, A., Wang, K., and Tsung, F. (2014). Statistical surface monitoring by spatial-structure modeling. *Journal of Quality Technology*, 46(4):359–376.
- Wells, L. J., Megahed, F. M., Niziolek, C. B., Camelio, J. A., and Woodall, W. H. (2013). Statistical process monitoring approach for high-density point clouds. *Journal of Intelligent Manufacturing*, 24(6):1267–1279.
- Zang, Y. and Qiu, P. (2017). Phase I monitoring of spatial surface data from 3D printing. *Technometrics*, 60:169–180.
- Zang, Y. and Qiu, P. (2018). Phase II monitoring of free-form surfaces: An application to 3D printing. *Journal of Quality Technology*, 50:379–390.
- Zhao, C., Shi, W., and Deng, Y. (2005). A new Hausdorff distance for image matching. *Pattern Recognition Letters*, 26:581–586.
- Zhou, Z.-Q. and Wang, B. (2009). A modified Hausdorff distance using edge gradient for robust object matching. In *Proceedings of 2009 International Conference on Image Analysis and Signal Processing, IASP 2009*, pages 250 – 254.
- Zhu, L., Brereton, R. G., Thompson, D. R., Hopkins, P. L., and Escott, R. E. (2007). On-line hplc combined with multivariate statistical process control for the monitoring of reactions. *Analytica Chimica Acta*, 584(2):370 – 378.