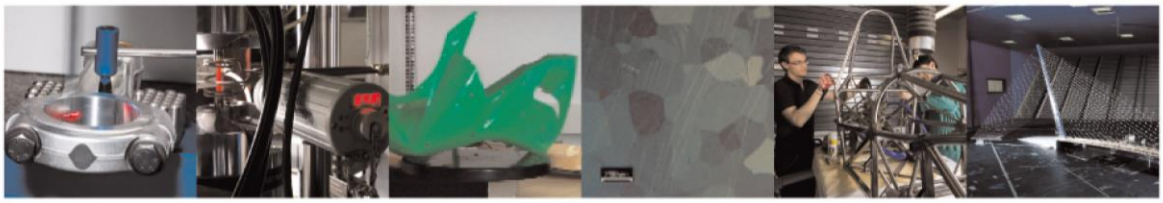


Multisensor data fusion via Gaussian process models for dimensional and geometric verification

B.M. Colosimo, M. Pacella, N. Senin

This is a post-peer-review, pre-copyedit version of an article published in Precision Engineering. The final authenticated version is available online at:
https://doi.org/10.1016/j.precisioneng.2014.11.011

DIPARTIMENTO DI MECCANICA ■ POLITECNICO DI MILANO
via G. La Masa, 1 ■ 20156 Milano ■ EMAIL (PEC): pecmecc@cert.polimi.it
http://www.mecc.polimi.it
Rev. 0

Multisensor Data Fusion via Gaussian Process Models for Dimensional and Geometric Verification

B.M. Colosimo[1]*, M. Pacella[2], N. Senin[3]

[1]Politecnico di Milano, Dip. di Meccanica, via G. La Masa 34, 20156 Milano, Italy, biancamaria.colosimo@polimi.it

[2]Università del Salento, Dip. di Ingegneria dell'Innovazione, via per Monteroni, 73100 Lecce, Italy, massimo.pacella@unisalento.it

[3]Università di Perugia, Dip. di Ingegneria, via G. Duranti 67, 06125 Perugia, Italy, nicola.senin@unipg.it

Abstract

An increasing amount of commercial measurement instruments implementing a wide range of measurement technologies is rapidly becoming available for dimensional and geometric verification. Multiple solutions are often acquired within the shop-floor with the aim of providing alternatives to cover a wider array of measurement needs, thus overcoming the limitations of individual instruments and technologies.

In such scenarios, multisensor data fusion aims at going one step further by seeking original and different ways to analyze and combine multiple measurement datasets taken from the same measurand, in order to produce synergistic effects and ultimately obtain overall better measurement results.

In this work an original approach to multisensor data fusion is presented, based on the development of Gaussian process models (the technique also known as kriging), starting from point sets acquired from multiple instruments. The approach is illustrated and validated through the application to a simulated test case and two real-life industrial metrology scenarios involving structured light scanners and coordinate measurement machines.

The results show that not only the proposed approach allows for obtaining final measurement results whose metrological quality transcends that of the original single-sensor datasets, but also it allows to better characterize metrological performance and potential sources of measurement error originated from within each individual sensor.

Keywords: Multisensor data fusion, coordinate metrology, CMM, structured light scanner, Gaussian process, kriging

*corresponding author

1. Introduction

1.1. Multisensor instruments for dimensional metrology

The combined use of multiple measurement sensors is becoming commonplace in dimensional metrology and an increasingly wider array of instruments equipped with multiple probes is becoming available. Popular commercial solutions for the measurement of parts include touch-probe CMMs equipped with additional optical and/or vision sensors [1-4], and measuring arms equipped with touch-probes and laser point or line scanners [5, 6]. Even in surface metrology, where the aim is the characterization of surface texture at micro and sub-micro scales, 3D microscopes have recently become available equipped with multiple measurement heads implementing different measurement technologies (e.g. vertical scanning interferometry + focus variation [7]).

All such commercial offerings are based on the same conceptual approach: “one fixture, multiple sensors”, i.e. all these instruments are designed to provide multiple measurement options within a single measurement setup, essentially letting the user select the proper sensor for each task, thus overcoming the limitations of each single measurement technology. Once the workpiece is mounted onto the instrument, depending on the type of characterization, part accessibility, time and accuracy requirements, the user is free to select the probe / sensor technology that is better suited to accomplish the inspection/verification task.

1.2. Multisensor data fusion

Multisensor data fusion tries to go one step further [8-13], and refers to the process of combining multiple sensor data sets with the goal of obtaining a result which either marks an improvement with respect to what obtainable from each data set taken singularly, or constitutes an entirely new piece of information, which could not be obtained by simply analyzing any of the individual datasets.

In multisensor data fusion, it is not necessary that all the datasets come from the same instrument, and data may have been acquired at different places and times. Combining multiple data sets may refer to combining data coming from different sensors and/or sensor types, but may also refer to combining data coming from the same sensor, used with different setups, or even used multiple times with identical operating conditions (i.e. combining replicate data sets).

Data types, which can be integrated in dimensional metrology, include:

- conventional digital images (RGB, greyscale): as acquired by digital cameras;
- range images (images whose pixels contain distance information): as acquired by structured light scanners, 3D microscopes and photogrammetry systems;
- point clouds (i.e. set of points in 3D space): as acquired by CMM, measuring arms, single-point laser trackers, laser radar, etc., and

- volume data (i.e. 3D matrices): as acquired by X-ray Computerized Tomography.

In general, the term *homogeneous integration* is used when combining the same type of data (e.g. 3D point clouds), while *inhomogeneous integration* is used in all the other cases.
Depending on the type of data to be integrated and on the overall characterization goal, many multisensor data fusion scenarios may be imagined; these can be categorized into three main classes according to a popular classification scheme [14]. In *competitive data fusion*, redundancy originated by replicate data sets acquired with the same sensor and in the exact same operating conditions is used to improve the metrological quality of the result. For example, multiple identical images of the measurand can be combined in order to extract an average image, more robust to noise. In *complementary data fusion*, homogeneous data sets, taken by the same sensor but in slightly different operating conditions, provide information so that each set is meant to complement the others. Fusion in this case is meant to take advantage of such complementarity. For example, digital images with the same magnification but slightly different localization may be stitched to obtain increased spatial coverage, or images taken at different magnification could be fused to obtain a result that covers a wider array of spatial resolutions (scales). Finally, *cooperative data fusion* gathers all the types of integration involving homogeneous/inhomogeneous data sets, which cannot be classified under competitive or complementary integration. A few scenarios of cooperative data fusion have already gained some popularity [9]: in dimensional and geometric verification, vision can be used to acquire global shape information needed to automatically produce an inspection path for the touch probe; in defect identification vision can be used to identify and localize a defect, then localization information can be used to drive a laser line scanner which performs the actual shape measurement of the defect; in reverse engineering, high-density point clouds obtained by an optical sensor can be stitched together with the help of a few reference points obtained by a touch probe to reconstruct the full-3D shape of an object. Some sensor technologies are intrinsically based on some form of cooperative integration [9]: for example, depth from focus, shape from shading and photogrammetry are 3D imaging technologies, which are based on fusing data obtained from conventional 2D images in order to obtain 3D information.

1.3. An overview of some notable approaches to multisensor data fusion

Some of the cooperative scenarios cited above can also be classified as sequential data fusion, an additional category where the first dataset is used to obtain the second, and then it is discarded. For example, in [12] high-density, low-quality information acquired by means of a vision system is used to guide the acquisition of a low-density, high quality dataset via a touch probe CMM. The dataset obtained by vision is discarded after the CMM dataset is available. An approach where both datasets are kept can be found in [15], where a laser

scanner is used to acquire free-form surface patches, while a CMM is used to acquire patch boundaries only. Fusion is achieved by simply adding the two datasets together.
Fusion is also meant as a way to define the appropriate compensation (i.e., a roto-translation matrix) to be applied to a sensor to achieve information provided by the other one [16]. As in many applications of data fusion, this approach assumes that all the sensors acquire data at the same locations, an assumption that usually does not hold when measurement systems based on high-density optical scanning are considered.
In [17] a method is proposed for fusing high-resolution and a low-resolution data: after registration and elimination of redundant points, merging is achieved by remeshing all the acquired data points. No statistical models to represent measurement errors are considered in the merging procedure. On the contrary, in [18] the datasets acquired by different sensors are considered as different responses of a multivariate linear (or non-linear) model, and Bayesian estimates of the unknown coefficients are carried out. The statistical model is used specifically to correct laser trackers responses; The same locations are measured with all the available sensors (or multiple times by the same tracker) in order to compute the fusion step.

Most of the aforementioned approaches for data fusion either combine information by simply adding data points originated from different observations (after appropriate elimination of redundant data) or, assuming points are taken at the exact same locations, use one dataset to correct the other. Furthermore, most of the methods assume deterministic data fusion or statistics as a way to estimate unknown coefficients. A notable exception is the method presented in [19], aimed at multisensor data alignment. In this case, the main idea is to reconstruct the information provided by all the different sensors before performing the fusion step. This approach has the main advantages of i) include statistical modelling while reconstructing the information provided by different data sets, providing prediction intervals for the local discrepancies between different data sets as well as on the final prediction of the shape at any given location; ii) relaxing the assumption of acquiring all the data at the same location set. This approach is considered as starting reference for the fusion procedure presented in this work.

1.4. Multisensor scenarios involving 3D point sets with different densities and metrological performance

The specific data fusion scenarios investigated by this work involve multiple 3D point sets (point clouds, i.e. homogeneous data) acquired from the same measurand surface as part of an inspection / verification process [20].
In these scenarios, the datasets are supposed to belong to one of the two following main categories:

- *Points coming from touch-probe CMMs*:
  In a conventional CMM, point acquisition is generally very slow (acquisition in single point mode), or slightly less so (acquisition in profile mode); the localization of the points on the measurand can be accurately controlled by the operator, and variable point spatial density on the measurand surface is usually explicitly designed into the verification process with the aim of collecting an increased number of points in high-curvature or otherwise critical regions. Point density is generally low anyways, given the cost of slow measurement. Due to the optimal traceability of CMMs, measurement can be considered as fairly accurate, hence in general high trueness and high precision are to be expected. In summary, these points sets will be referred to as characterized by low-density (LD) and high (metrological) performance, but will tend to be scarcely populated, given the high acquisition costs.

- *Points coming from line scanners, structured light scanners, and photogrammetry:*
  In these techniques, point acquisition is generally extremely fast and data sets are highly populated. However, there is usually little to no control on the exact localization of each acquired point over the measurand surface. In imaging techniques such as structured light scanning and photogrammetry, points are usually spread along a regular, rectangular grid in the *x,y* projection plane, and are differentiated from each other by height (*z*) information (i.e. they form a range image). Spatial density (in the *x,y* plane) is very high. However, trueness and precision tend to be low. These instruments are poorly characterized from a metrological standpoint and scarcely traceable. Moreover, the likelihood of local bias (usually due to some form of optical distortion) is potentially very high. In summary, these points sets will be referred to as characterized by high-density (HD) and low metrological performance, and will tend to be highly populated, given the low acquisition costs.

It is clear that measurement with line scanners, structured light scanners and photogrammetry possesses some very desirable properties, in particular in terms of acquisition speed and low cost for obtaining a very large amount of points. However, the appeal of these techniques is somehow hampered by their poor metrological performance. This is why currently the CMM touch-probe is still considered the standard de-facto for metrological characterization, while these high-speed techniques are relegated to reverse engineering and in general to any type of shape reconstruction where metrological accuracy is not as relevant.

1.5. Multisensor data fusion scenario investigated in this work

This work investigates the possibility of improving the metrological performance of a point set acquired with a structured light scanner or analogous high-speed technique (cheap, many

points, high-density, low metrological performance). The main goal is to improve local trueness (reduce local bias) of the dataset by fusing it with a smaller set of "truer" points acquired with a touch-probe CMM, or analogous technique with high metrological performance but low acquisition speed – high measurement cost). The idea is to implement a fusion process where the CMM points act as local attractors, essentially introducing corrections to local bias into the high density point set, or at least highlighting those regions in the set where the bias problem is more significant.
Therefore, two datasets are assumed available in this scenario:

- a high-density, highly-populated, low metrological performance one, generated from a structured light scanner or analogous technique. This will be simply referred to as the HD (high density) set;
- a low-density, scarcely-populated, high metrological performance one, generated from a touch probe CMM. This will be simply referred to as the LD (low density) set.

It is assumed that both the LD and HD sets have already been correctly localized within the same coordinate system, i.e. that the *registration problem* has been solved. This is already a non-trivial task, and a fundamental one in determining the successfulness of fusion. Details on how registration can be successfully accomplished in multisensor data fusion scenarios can be found in recent work by the authors [21].

Since the main goal of the proposed approach is to reduce the systematic measurement error, which can greatly contribute to the overall uncertainty budget, this paper will mainly focus on this aspect (from Section 2 to Section 4).
Additional benefits of the proposed approach are illustrated in Section 5: the model for multisensor data fusion can be used as a valid tool to assess the precision (i.e. the random error component) that can be associated to each dataset involved in fusion, even without the need to resort to replicate measurements. Moreover it allows for investigating how the random error components associated to the individual datasets propagate through fusion and ultimately affect the prediction error associated to the fusion result. Conclusions and potential directions for future research are reported in Section 6.

2. The Multisensor Data Fusion (MsDF) model

2.1. Formal representation of the point sets

It is assumed that both point sets can be represented by discrete functions of the type: $z\left(x_i, y_i\right)$, i.e. where the *z*-coordinate of the *i*-th point is expressed as a function of its position on the *x,y* plane. This formal encoding is ideally suited to the HD set, since structured light scanners and similar imaging devices produce range image data (i.e., only one *z* value can

exist at any given *x,y* position) and can be adapted to most CMM point sets (LD sets) as long as the low density allows for identifying a transform that maps the points to the same *x,y* plane of the HD set. Then, by expressing point position as a vector of two components: $\mathbf{v}_i = \left(x_i, y_i\right)$ the HD set becomes: $z_{HD}\left(\mathbf{v}_i\right)$ with $\mathbf{v}_i \in V_{HD} \subset \mathbb{R}^2$, $i = 1,2,\cdots,n_{HD}$ and similarly, the LD set becomes: $z_{LD}\left(\mathbf{v}_i\right)$ with $\mathbf{v}_i \in V_{LD} \subset \mathbb{R}^2$, $i = 1,2,\cdots,n_{LD}$. As stated previously, the HD set is larger than the LD set, i.e. $n_{HD} > n_{LD}$ and in general the points of the two sets will not share the same locations on the *x,y* plane, i.e. $V_{HD} \cap V_{LD} = \varnothing$, because it is very hard to acquire a point with a touch probe CMM at the exact same location of a point belonging to a range image obtained with a structured light scanner or similar.

2.2. First stage of the data fusion model

Data fusion is achieved using a two-stage *Multisensor Data Fusion* (MsDF) model [19, 22]. At the first stage, a representation of the geometry is obtained by using the HD data only and can be defined as follows:

$$z_{HD}\left(\mathbf{v}_i\right) = f_{HD}\left(\mathbf{v}_i\right) + \varepsilon_{HD} \tag{1}$$

where the term ε_{HD} accounts for the error, and is assumed to follow a normal distribution with 0 mean and $\sigma^2_{\varepsilon HD}$ variance, i.e. $\varepsilon_{HD} \sim N\left(0, \sigma^2_{\varepsilon HD}\right)$. In the ideal case of a perfect fitting, that is when the model $f_{HD}\left(\mathbf{v}_i\right)$ of the underlying surface is appropriate and the related coefficients are correctly estimated, ε_{HD} should represent the HD measurement error. Therefore, it is clear that the core part of equation (1) is the term $f_{HD}\left(\mathbf{v}_i\right)$, representing the pattern of the underlying feature observed with noise. Among the different models that can be assumed for this term, a Gaussian Process (GP) model [23-25] is considered in this work. A GP model (the technique also known as kriging) is a particular type of random process where the *pdf* associated to any process observation (i.e. the *z* value at a given location in the *x,y* plane) is normal, and the joint probability distributions associated to any finite subset of process observations are normal as well [26-28]. Formally, a GP model is defined by:

$$f_{HD}\left(\mathbf{v}_i\right) = GP\left(m_{zHD}\left(\mathbf{v}_i\right), k_{zHD}\left(\mathbf{v}_i, \mathbf{v}_j\right)\right) \tag{2}$$

where $m_{zHD}\left(\mathbf{v}_i\right) = E\left[z_{HD}\left(\mathbf{v}_i\right)\right]$ is the mean function, which is used to describe the expected *z* value at $\mathbf{v}_i$. $k_{zHD}\left(\mathbf{v}_i, \mathbf{v}_j\right) = E\left[\left(z_{HD}\left(\mathbf{v}_i\right) - m_{zHD}\left(\mathbf{v}_i\right)\right)\left(z_{HD}\left(\mathbf{v}_j\right) - m_{zHD}\left(\mathbf{v}_j\right)\right)\right]$ is the covariance function, which is used to describe the variance of the *z* value at location $\mathbf{v}_i = \mathbf{v}_j$ as well as

the covariance between the *z* values located at $\mathbf{v}_i \neq \mathbf{v}_j$ (a measure of how much such *z* values change together).

In this work, we used a simple linear model (geometrically a plane in space) to represent the mean function of the GP model: $m_{zHD}\left(\mathbf{v}_i\right) = \beta_{0,HD} + \beta_{1,HD} x_i + \beta_{2,HD} y_i$. In addition, we used the *squared exponential* function to represent the covariance of the GP model, i.e.:

$$k_{zHD}\left(\mathbf{v}_i, \mathbf{v}_j\right) = \sigma_{zHD}^2 \exp\left(-\frac{\left\|\mathbf{v}_i - \mathbf{v}_j\right\|^2}{2l_{HD}^2} \right) \tag{3}$$

where $\left\|\mathbf{v}_i - \mathbf{v}_j\right\|$ is the Euclidean distance between locations $\mathbf{v}_i$ and $\mathbf{v}_j$ in the *x,y* plane, σ_{zHD}^2 is the constant variance of the GP model (note that the overall variance of $z_{HD}\left(\mathbf{v}_i\right)$ is constant too and is equal to $\sigma_{zHD}^2 + \sigma_{\varepsilon HD}^2$) and l_{HD} is the characteristic length-scale (i.e., the length over which there is no significant relationship between two *z* values). In practice, according to equation (3), *z* values which lie closely together on the *x,y* plane (no matter where they are located) are likely to be more similar. The squared exponential is one of the most popular choices for GP models, because it yields positive definite correlation matrices, enables proper convergence of the statistical estimation algorithms and can model smooth and infinitely differentiable functions [27].

Parameters $\left\{\beta_{0,HD}, \beta_{1,HD}, \beta_{2,HD}, \sigma_{\varepsilon HD}^2, l_{HD}, \sigma_{zHD}^2\right\}$ of the geometry model described by equations (1)-(3) are all unknown and must be estimated from the actual measurement data $z_{HD}\left(\mathbf{v}_i\right)$ with $\mathbf{v}_i \in V_{HD} \subset \mathbb{R}^2$, $i = 1,2,\cdots,n_{HD}$.

Details on the statistical estimation of the GP parameters can be found the appendix of this paper. Fitting of GP models was implemented in this paper based on the code developed in [29].

Once parameter estimation is complete, the knowledge of the mean $m_{zHD}\left(\cdot\right)$ and covariance $k\left(\cdot,\cdot\right)$ functions makes it possible to estimate the function value $z_{HD}\left(\mathbf{v}\right)$ at any new location $\mathbf{v}$ in the *x,y* plane, given the measured points $z_{HD}\left(\mathbf{v}_i\right)$ with $\mathbf{v}_i \in V_{HD} \subset \mathbb{R}^2$, $i = 1,2,\cdots,n_{HD}$ (refer to the appendix for details). Both a punctual ($\hat{z}_{HD}\left(\mathbf{v}\right)$) and a prediction interval for the function value $z_{HD}\left(\mathbf{v}\right)$ can be computed at any location $\mathbf{v}$. Note that prediction intervals are particularly useful in providing information concerning the level of uncertainty corresponding to each new prediction computed at a new location.

2.3. First stage of the data fusion model

In order to illustrate how the first stage of the MsDF works, a simple simulated example is shown. Without loss of generality, a profile is used as reference instead of a surface, to ease the graphical representation of the approach. Figure 1 shows the profile to be measured.

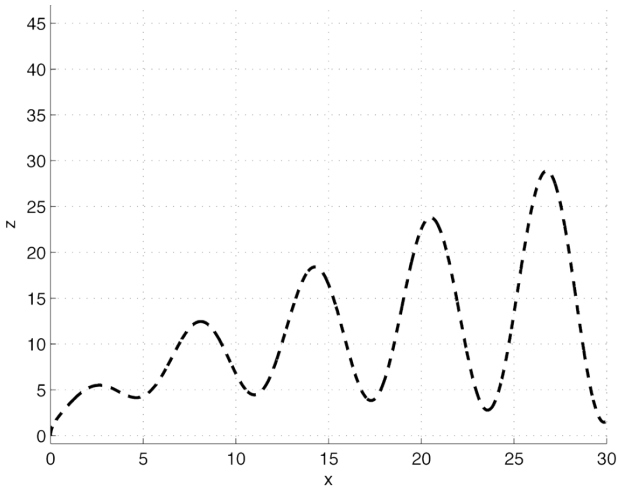
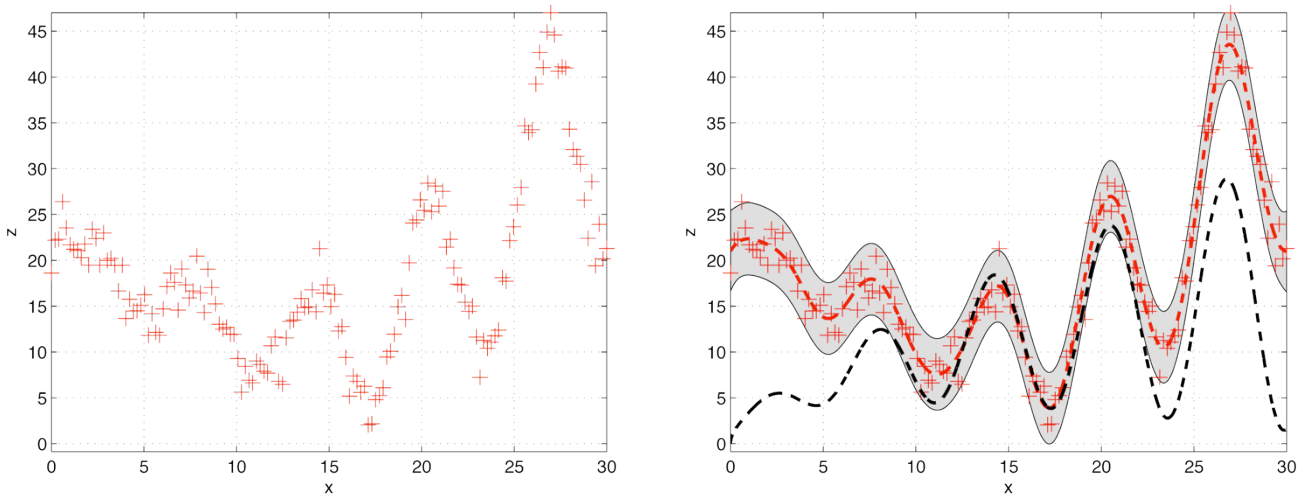


Figure 1: a simple example of profile to be measured

In Figure 2.a the set of points resulting from low-quality, high density, simulated measurement (HD set) is shown. Consistent with the type of optical measurement being investigated, precision is low over the entire profile. The simulation replicates also typical calibration conditions of optical systems and in particular of structured light scanners: measurement is unbiased at the center of the field of view, but not so at the boundaries, where effects such as optical aberration introduce an increasing amount of bias.



(a) (b)

Figure 2: Fitting a GP model to the HD set; a) original HD set; b) results of fitting: original profile (black), original HD set (red dots), prediction interval (gray band) and mean line (red, dashed)

In Figure 2.b, the results of fitting to the model in equation (1), i.e. a GP model with a simple line as mean function and the squared exponential as covariance function, are shown overimposed to the original profile (black, continuous line) and the original HD set (red dots). The red, dashed line represents the local estimate $\hat{z}_{HD}(\mathbf{v})$, while the grey band represents the prediction interval at 95% confidence level. The prediction interval appears quite narrow about the mean line thanks to the high density of the HD points.
As it is clear from the figure, the predicted model follows the HD set, drifting away from the real profile in presence of measurement bias.

2.4. Second stage of the data fusion model

As seen in the previous section, the GP model built over the HD data set allows for having a continuous function that can be interrogated at any location, and that provides surface information in statistical terms (estimate of *z*, and associated prediction interval at any location). The downside is that the model is incapable of correcting local bias which may be present in the original HD set. This is where data fusion really comes to help.
Recall the availability of the second set of points, a low-density, high quality set set: $z_{LD}(\mathbf{v}_i)$.
It is assumed that high quality comes in the form of high precision (i.e. small variance for the measurement error) and unbiasedness (i.e. zero mean measurement error) over the entire measurement range. Unbiasedness in particular means that even in those locations where the HD set drifts away from the surface being measured, the LD set would still provide a reliable reference. Thus it makes sense to find a way to use the information contained in the LD set to "correct" the GP model produced starting from the HD set.
To understand how the correction is made, it is necessary to introduce a "linkage" model first, i.e. a model that describes the difference between the *z*-values estimated by the GP model obtained from HD data (where the punctual prediction is represented by $\hat{z}_{HD}(\mathbf{v})$) and the *z*-values of the LD set $z_{LD}(\mathbf{v}_i)$, in correspondence to those locations $\mathbf{v}_i$ where the LD data area available, i.e $\mathbf{v}_i \in V_{LD}$, $i = 1,2,\cdots,n_{LD}$. Following [19, 22], the following form of the linkage model is assumed:

$$z_{LD}(\mathbf{v}_i) = \rho(\mathbf{v}_i)\hat{z}_{HD}(\mathbf{v}_i) + \delta(\mathbf{v}_i) + \varepsilon_{MsDF}, \qquad \mathbf{v}_i \in V_{LD}, \quad i = 1,2,\cdots,n_{LD} \tag{4}$$

The terms $\rho(\mathbf{v}_i)$ and $\delta(\mathbf{v}_i)$ represent a scaling and a shifting factor, respectively. The linkage model is saying that, in correspondence to a location containing a LD point, the LD *z*-value can be obtained by scaling and shifting the *z*-value obtained by the GP model fitted to the HD data $\hat{z}_{HD}(\mathbf{v}_i)$, plus/minus an error associated to the LD data. As the model in

equation (4) is statistical, it also includes an error term ε_{MsDF} which models the residual differences between the LD set $z_{LD}(\mathbf{v}_i)$ and $\rho(\mathbf{v}_i)\hat{z}_{HD}(\mathbf{v}_i)+\delta(\mathbf{v}_i)$. These residuals are assumed to be normally distributed with zero mean and constant variance $\sigma^2_{\varepsilon MsDF}$, i.e., $\varepsilon_{MsDF} \sim N(0,\sigma^2_{\varepsilon MsDF})$.

The scaling function can be assumed as linear: $\rho(\mathbf{v}_i)=\rho_0+\rho_1 x_i+\rho_2 y_i$ (again, a plane in space). On the authors knowledge, more complex models for the scaling function are unnecessary in common multi-resolution scenarios. The shifting function can be modeled by another GP of constant mean δ_0 and covariance function $k_\delta(\mathbf{v}_i,\mathbf{v}_j)$, i.e.:

$$\delta(\mathbf{v}_i) \sim GP\left(\delta_0, k_\delta(\mathbf{v}_i,\mathbf{v}_j)\right), \tag{5}$$

As in the previous equation (3), the covariance function can be described by the squared exponential, i.e.: $k_\delta(\mathbf{v}_i,\mathbf{v}_j)=\sigma^2_\delta \exp\left(-\left\|\mathbf{v}_i-\mathbf{v}_j\right\|^2 / 2l^2_\delta\right)$.

Parameters $\left\{\rho_0,\rho_1,\rho_2,\delta_0,\sigma^2_\delta,l_\delta,\sigma^2_{\varepsilon MsdF}\right\}$ in equations (4) and (5) are all unknown and must be estimated from the measurements $z_{LD}(\mathbf{v}_i)$ with $\mathbf{v}_i \in V_{LD} \subset \mathbb{R}^2$, $i=1,2,\cdots,n_{LD}$. Again, this is mathematically handled as an optimization problem where the parameters that define the GP in equation (4) are varied while searching for the best solution measured in terms of the likelihood of such parameters for the given LD point set.

The successful estimation of all the unknown parameters leads to the full reconstruction of the linkage model. The model can then be used to extract the correction factors to be applied to the original GP model defined from the HD set. In fact, at any generic location $\mathbf{v}$, the original prediction $\hat{z}_{HD}(\mathbf{v}_i)$ can be corrected into: $\hat{z}_{MsDF}(\mathbf{v})=\rho(\mathbf{v})\hat{z}_{HD}(\mathbf{v})+\hat{\delta}(\mathbf{v})$. This has the effect of bending the original GP model (biased) towards the LD set (unbiased) where discrepancies exist.

2.5. Second stage of the data fusion model: example

To better illustrate the second stage of the data fusion model, the same example introduced before can be used. In Figure 3, the profile being measured is shown as a black, dashed line. The original HD points are shown as red dots, and the estimated *z* value $\hat{z}_{HD}(\mathbf{v}_i)$ of the original GP model fitted to the HD points is shown as a red, dashed line. The LD points are shown as blue dots, while the final model (after the correction by the linkage model) is represented by a dashed blue line (corrected z estimates) and associated prediction interval (grey band).

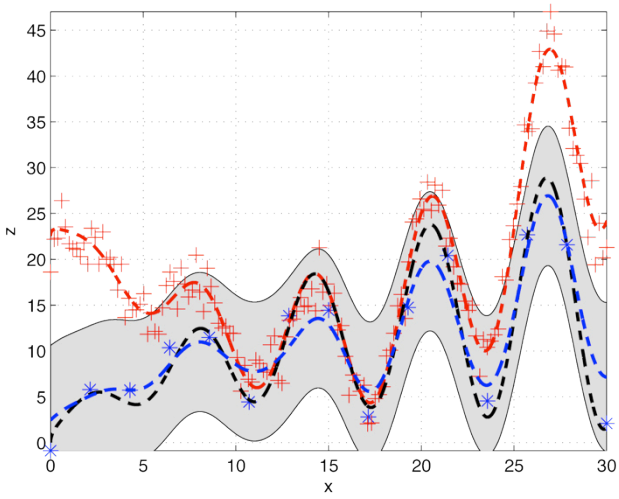


Figure 3: Correction of the GP model fitted to HD data by means of the linkage model; profile being measured (black, dashed); HD points (red dots), z estimates of the original GP model fitted to the HD set (red, dashed); LD points (blue dots), z estimates of the corrected GP model (blue, dashed); prediction interval of the corrected GP model (grey band)

By comparing Figure 3 with the uncorrected results in Figure 2.b, the final effect of data fusion (reduced bias in the prediction) is evident.

2.6. Fusion model vs. using only the LD set

Essentially, in the two-stage fusion process, a GP model is first built starting from HD data, and then it is corrected (scaling and shifting) by means of another GP (linkage) model based on LD data. At this point, one may wonder if it would be more convenient to build a continuous function by directly fitting a GP model to the LD set, discarding the HD data altogether and thus eliminating the bias problem.

Unfortunately, for the specific application scenario discussed in this work, this would be inadvisable, due to the low-density of the LD dataset. This is immediately visible with the help of Figure 4, where a GP model has been fitted directly to the LD points acquired from the test profile. The original LD points are shown in Figure 4.a; the fitted GP model $z_{LD}\left(\mathbf{v}_i\right) = f_{LD}\left(\mathbf{v}_i\right) + \varepsilon_{LD}$, where $f_{LD}\left(\mathbf{v}_i\right)$ is a GP model in all similar to that used in section 2.2 (i.e., a simple line as mean function and the squared exponential as covariance function). The result is shown in Figure 4.b in terms of its z-estimates $\hat{z}_{LD}\left(\mathbf{v}\right)$ (light blue) and prediction interval (gray band). In Figure 4.c, the illustration is completed with the addition of the original profile being measured (black, dashed line). It is evident that, tough at least the line

of the estimates adequately follows the general orientation of the profile, it does not accurately reproduce its shape; also, the prediction interval is too large to be of any practical use. On the contrary, the use of the HD data set, albeit introducing bias, is invaluable because, given its high density, it is capable of providing a better starting point of the GP model which can be later corrected through fusion.

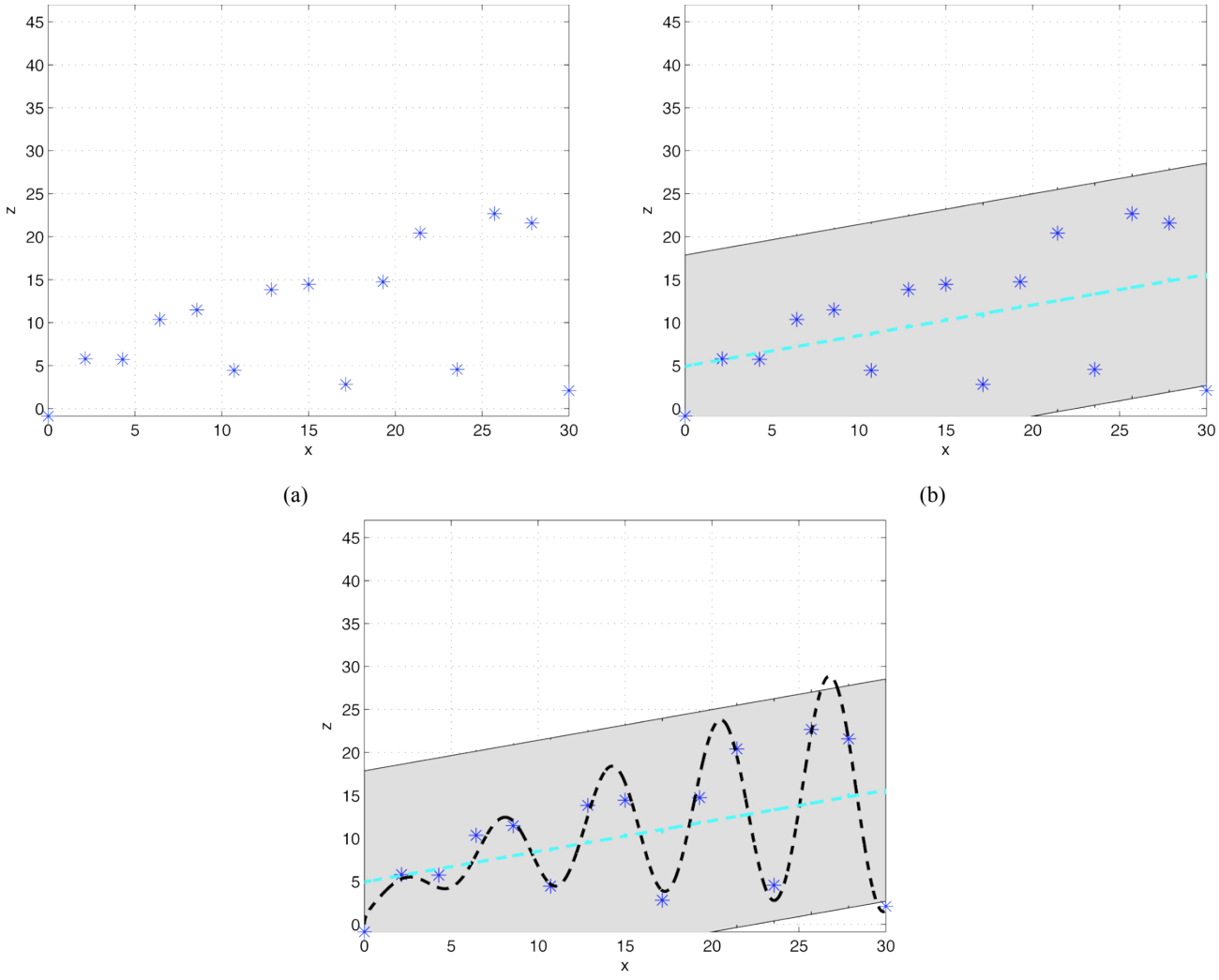


Figure 4: Directly fitting the LD dataset to a GP model: a) LD dataset; b) line of the z estimates (light blue) and prediction interval (gray band) of the fitted GP model; c) same as the previous, with added the profile being measured (black dashed line)

2.7. Other uses of the MsDF model

Even without applying the correction, the simple observation of the fitted terms $\rho(\mathbf{v})$ and $\delta(\mathbf{v})$ functions of the linkage model provides hints at regions where discrepancies between

the LD and HD sets are significant. This may be an indication of local bias, as in the simple profile example, but it may also indicate the presence of other problems in the measurement set-up. Therefore, by assuming consistency among sensors as an indication of higher reliability, the information may be used to decide what parts of both point sets to retain, and what to discard.

3. Experimental validation of the MsDF model

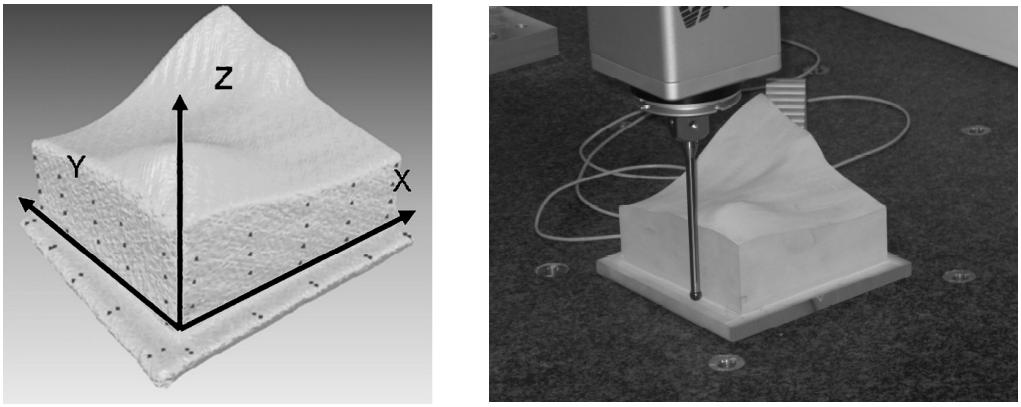
The freeform surface presented in [30] is used as a first case study (Figure 5). It was machined from a 100x100x100 mm workpiece. Three orthogonal planes (where dots are shown) define a reference system for the artifact.

The freeform surface was measured by means of two metrological devices, namely a structured light (SL) scanner and a Coordinate Measuring Machine (CMM) Zeiss “Prismo 5 HTG VAST” equipped with a analogue probe head with maximum probing error MPEP = 2 μm (according to ISO 10360-2). The free-form surface was first measured via SL using the calibration procedure presented in the literature [31]. The resulting point cloud was referred to the coordinate system defined by three orthogonal planes as shown in Figure 5.a, in order to allow the CMM to replicate measurements at the same locations where SL data were available.

The SL point cloud represents the HD dataset and consists of a total of $n_{HD} = 9635$ points (Figure 6.b). The LD dataset consists of $n_{LD} = 100$ CMM data points (Figure 6.a).

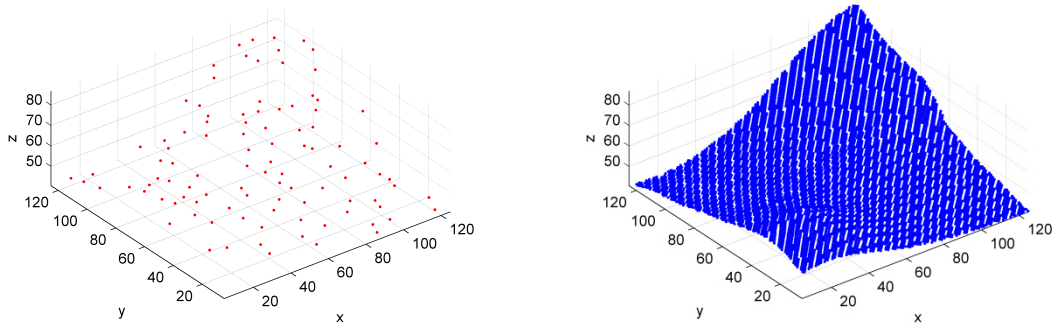
An additional set of $n_{test} = n_{HD} - n_{LD}$ CMM data points were acquired at the same locations where HD data were available acting as testing dataset V_{test}, i.e., the accurate measurements used to evaluate the prediction ability of all the competing methods.

The LD and HD data set were registered to the same Cartesian coordinate system (Figure 5.a) by aligning the points taken on the three orthogonal, reference surfaces of the object to be inspected (reference-based alignment - Figure 5.a). Since calibration error in one or both instruments may lead to an erroneous alignment result and thus influence fusion, an additional scenario was investigated where the dataset were subjected to further registration refinement via the Iterative Closest Point (ICP) procedure [32].



(a) (b)

Figure 5: The free-form surface used as case study. Three orthogonal reference surfaces (a). CMM sampling of the artifact (b)



(a) (b)

Figure 6: The free-form surface used as case study. 100 LD data set obtained via a CMM (a), 9635 HD data measured by SL scanner (b) (all values in [mm])

In order to evaluate the effectiveness of the proposed procedures, the Mean Squared Prediction Error (MSPE) was considered as performance indicator:

$$MSPE = \frac{1}{n_{test}} \sum_{\mathbf{v}_i \in V_{test}} \left(z_{test}\left(\mathbf{v}_i\right) - \hat{z}_{method}\left(\mathbf{v}_i\right) \right)^2 \qquad i = 1,2,\ldots,n_{test}\,, \qquad (6)$$

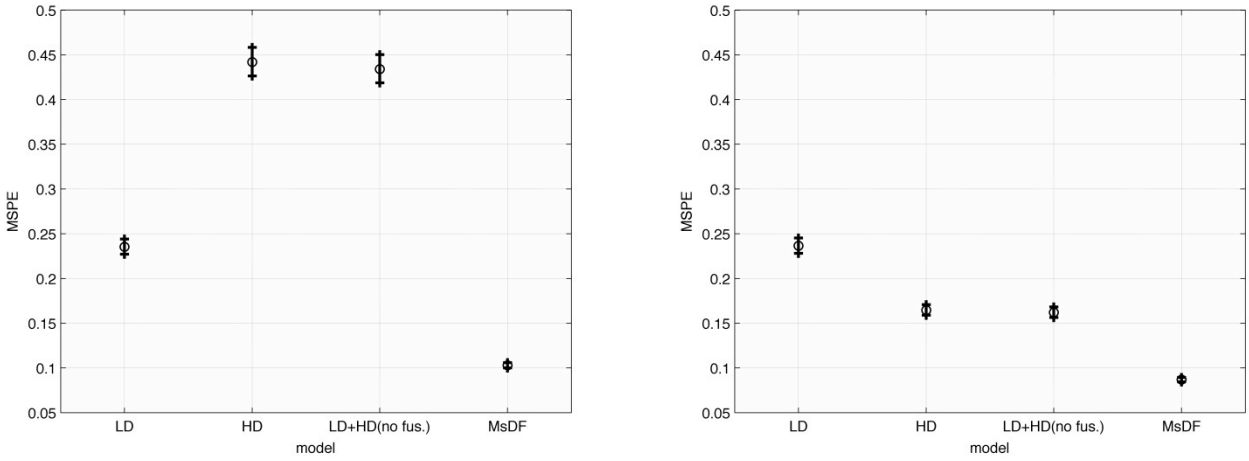
where $z_{test}\left(\mathbf{v}_i\right)$ is the CMM measurement at $\mathbf{v}_i \in V_{test}$ and $\hat{z}_{method}\left(\mathbf{v}_i\right)$ represents the predicted value for the surface at location $\mathbf{v}_i$, obtained using one of the following four approaches:

1. a GP model estimated using the LD data only ($n_{LD} = 100$);
2. a GP model estimated using the HD data only ($n_{HD} = 9635$);

3. a GP model estimated using the LD and the HD sets merged into a single dataset, as if they came from the same measurement system ($n_{LD} + n_{HD}$);
4. the MsDF approach illustrated in the previous section (GP model estimated using the LD data only and corrected by means of the HD data via the linkage model ($n_{LD} + n_{HD}$).

Figure 7 summarizes the results in terms of the MSPE (showing also the 95% confidence intervals on the MSPE, computed as the confidence interval on the mean value of the squared prediction errors observed at all the locations). Two scenarios are represented, depending on whether a fine ICP alignment is performed or not after the standard reference-based registration procedure. The MsDF method performs best, leading to predictions which are closer to the ideal result represented by $z_{test}\left(\mathbf{v}_i\right)$, thus confirming that the proposed procedure for integrating multi-resolution data is worth. From Figure 7, it appears that merging the LD and HD data without considering data fusion (method 3) does not have significant advantage over the GP model based on the HD data alone. This is probably due to the large size difference between the LD and the HD samples ($n_{LD} << n_{HD}$), which causes the LD sample to have a negligible effect on the performance of the GP model. This result can be observed in both the experimental scenarios (reference-based alignment and reference-based + ICP alignment) as summarized in Figure 7.

The performance of the GP model based on the HD data alone (method 2) is worse than what achievable using a GP model on the LD data (method 1) when reference-based alignment is assumed. This most likely indicates that the GP model alone was not able to compensate for the alignment error between the original HD dataset and the CMM verification set. When the additional ICP alignment is considered, the alignment error is greatly reduced, and the performance of the model based only on the HD data, improves significantly. Nevertheless, this is not enough to beat the MsDF performance, which still results in almost half the prediction error.



(a) (b)

Figure 7: Performance of the different approaches. MSPE and the 95% confidence intervals of the MSPE for each model (values in [mm^2]). Reference-based alignment of the data (a). Reference-based+ICP alignment of the data (b)

Figure 8 depicts the color-plot of the local prediction error ($z_{true}(\mathbf{v}_i) - \hat{z}_{method}(\mathbf{v}_i)$) of the four competing methods, when $n_{LD} = 100$ and only the reference-based alignment is performed. Warm colors represent underestimates of the true value ($z_{true}(\mathbf{v}_i) > \hat{z}_{method}(\mathbf{v}_i)$), while cold colors highlight overestimates of the prediction ($z_{true}(\mathbf{v}_i) < \hat{z}_{method}(\mathbf{v}_i)$). The green color indicates the correctness of the predictions. From Figure 8, it can be observed that the model based on LD data produces severe underestimates/overestimates in small portions of the surfaces. This is probably due to the lack of sampled data in those zones, which makes the GP model unable to estimate correctly the true values. On the other hand, the GP model based on HD and LD+HD (simple merging, no data fusion) produces severe underestimates in large portions of the surface (e.g. for abscissa ranging between 20 and 60).

The MsDF approach produces good predictions because it is able to appropriately take care of the bias through the linkage model, as illustrated in Figure 9, which depicts the color-plot of both the scaling function ($\rho(\mathbf{v})$) and the shifting function ($\delta(\mathbf{v})$). On the one hand, $\rho(\mathbf{v})$ produces a general compensation of bias in the zone with abscissa ranging between 20 and 60. On the other hand, some localized corrections are introduced by $\delta(\mathbf{v})$ especially in those zones where CMM measurement data are placed.

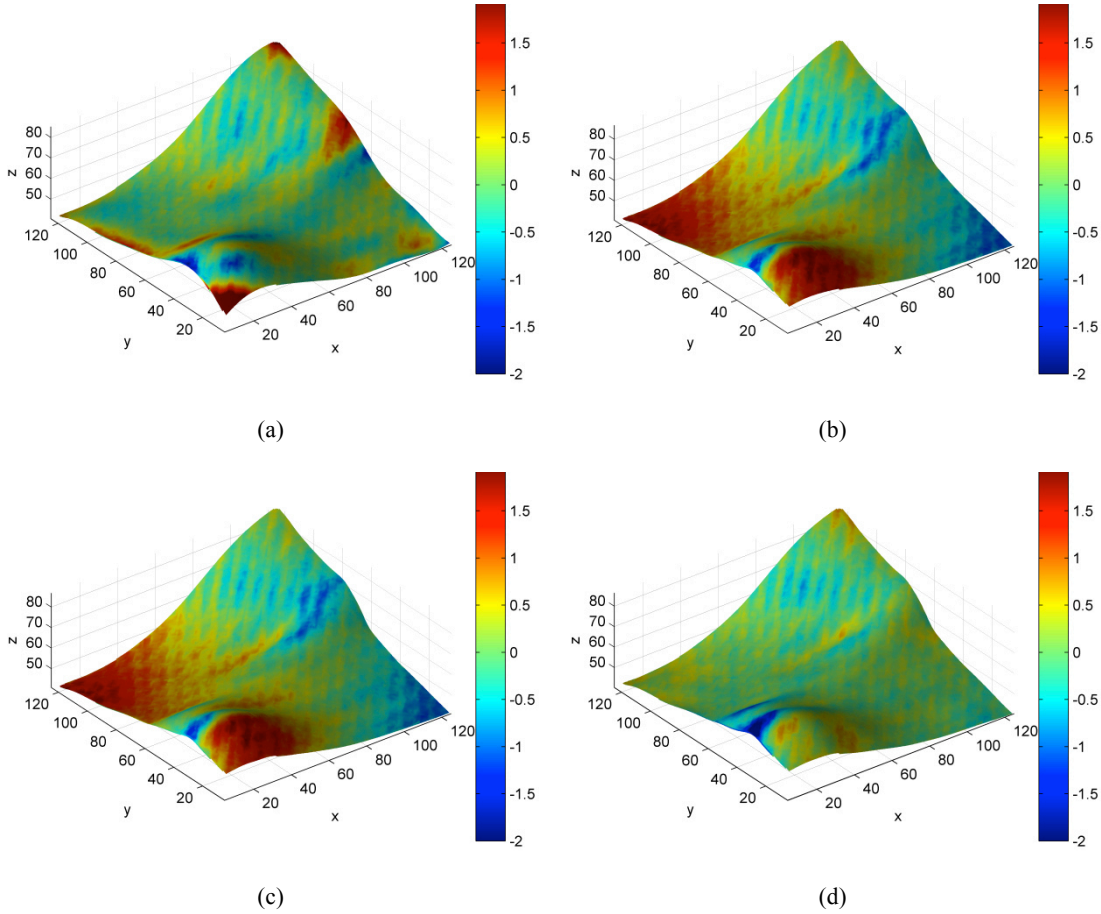


Figure 8: Color plot of the local prediction error (true minus predicted) for each method. 100 LD data points and feature-based alignment of data (all values in [mm]). LD model (a). HD model (b). LD+HD model (no data fusion) (c). MsDF model (d)

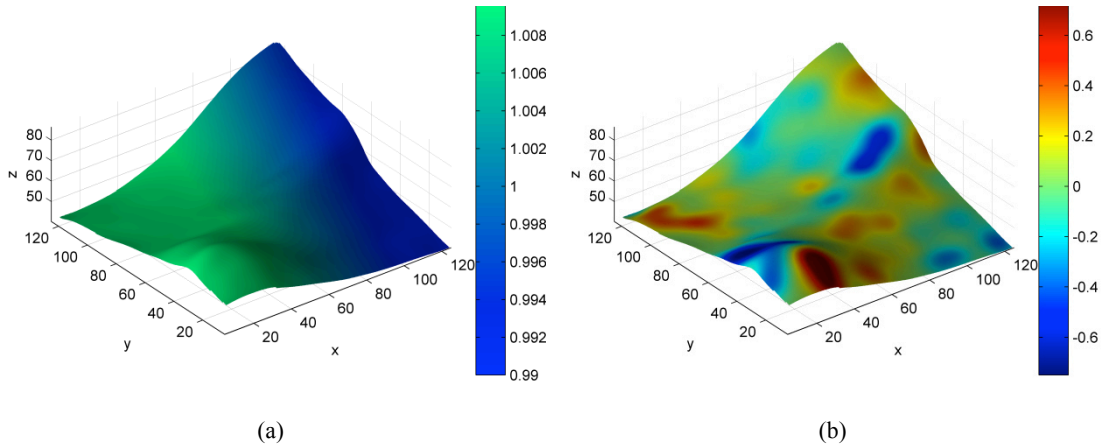


Figure 9: MsDF model. Color plot of $\rho(\mathbf{v})$ (a) and $\delta(\mathbf{v})$ (b) on the predicted surface. 100 LD data and feature-based alignment of data (all values in [mm])

Similar graphs are shown in Figure 10 in the case of 100 LD data points when ICP alignment is further performed to better align the two LD and HD data sets. By comparing the scaling function ($\rho(\mathbf{v})$) of the MsDF model in Figure 9 and Figure 10, it can be observed that the ICP alignment actually compensates for the bias and hence no scale adjustment of the low-resolution GP model is necessary (this is clear also by looking at the color-plots of the prediction error in Figure 10). Despite of this, some localized corrections by the shifting function $\delta(\mathbf{v})$ in those zones where CMM data is located are still required. In fact, the shifting function $\delta(\mathbf{v})$ based on 100 LD points in Figure 11 (after ICP alignment) looks similar to that in Figure 9 (before ICP alignment) because it mainly depends on the LD dataset obtained via CMM.

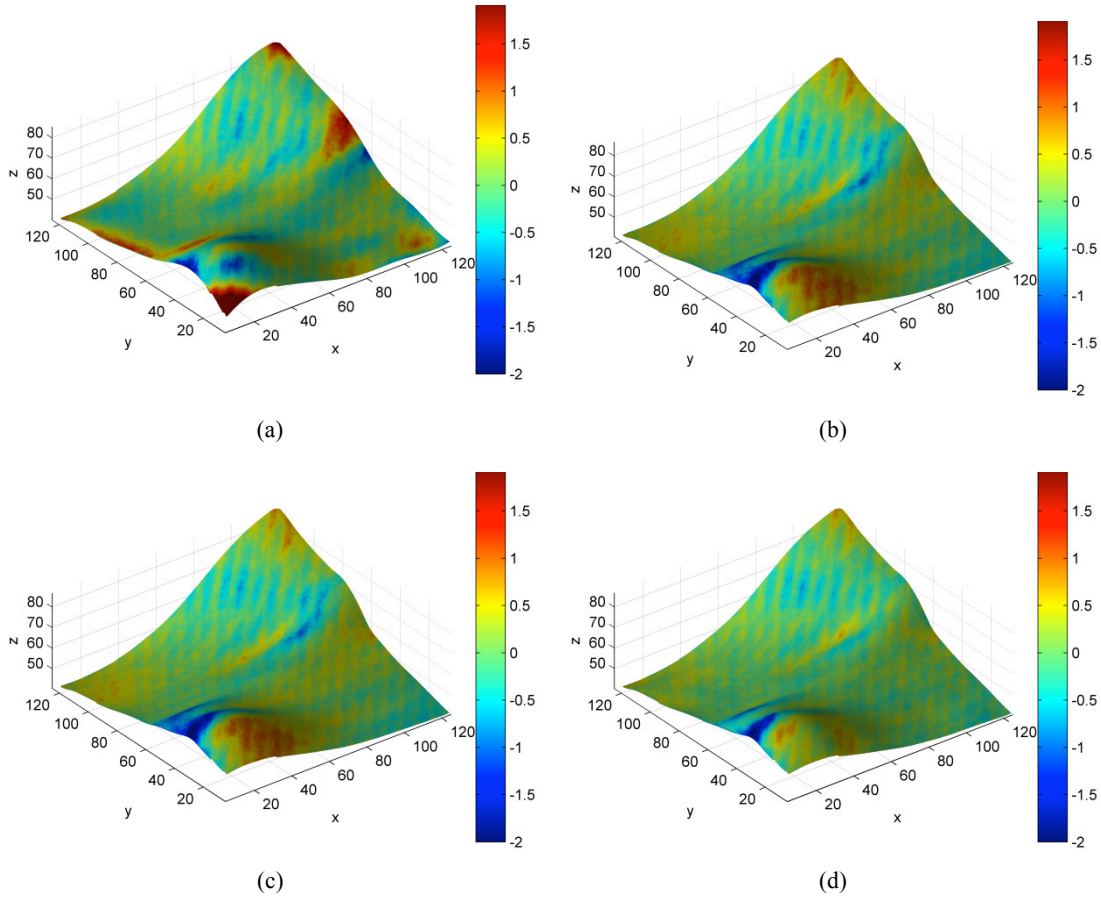
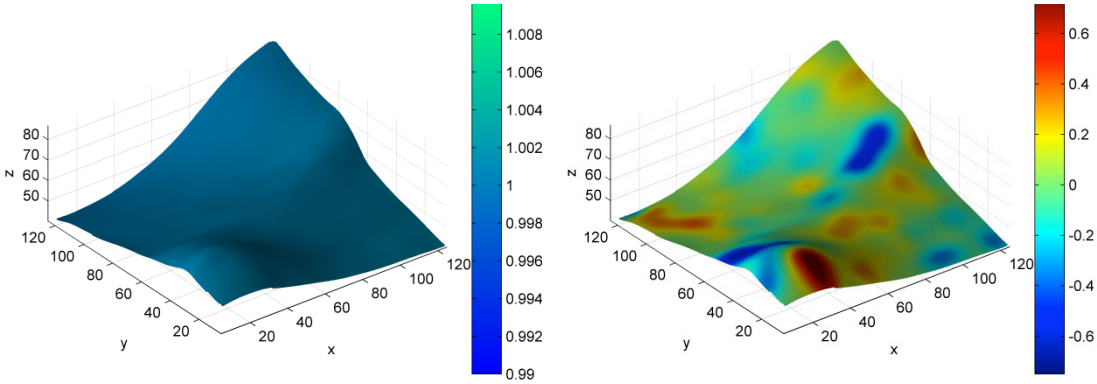


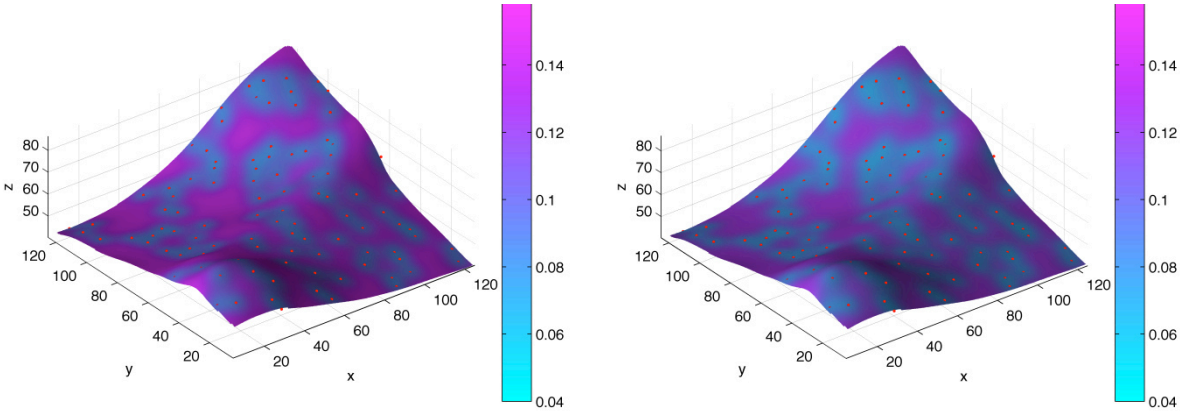
Figure 10: Color plot of the local prediction error (true minus predicted) for each model. 100 LD data points and feature-based +ICP alignment of data (all values in [mm]). LD model (a). HD model (b). LD+HD model (no data fusion) (c). MsDF model (d)



(a) (b)

Figure 11: MsDF model. Color plot of $\rho(\mathbf{v})$ (a) and $\delta(\mathbf{v})$ (b) on the predicted surface. 100 LD data points and feature-based +ICP alignment of data (all values in [mm])

Finally, it is worth noting that the scaling and shifting functions of the MsDF model provide useful information to the analyst concerning measurement uncertainty of the metrology devices considered. In particular, at each location, information about variance of the scaling and shifting functions can be useful to evaluate how reliable the discrepancy map is. As an example, Figure 12 shows the standard deviation of the shifting function. From Figure 12 it can be observed that measurement uncertainty is greater in zones not covered by the CMM samples, regardless of the specific algorithm used to align the point clouds.



(a) (b)

Figure 12: Color plot of the standard deviation of the shifting function of the MsDF model (all values in [mm]). Superimposed in each graph, the LD point cloud (red points). Reference-based alignment of the data (a). Reference-based+ICP alignment of the data (b)

4. Performance of MsDF on multiple connected freeform surfaces

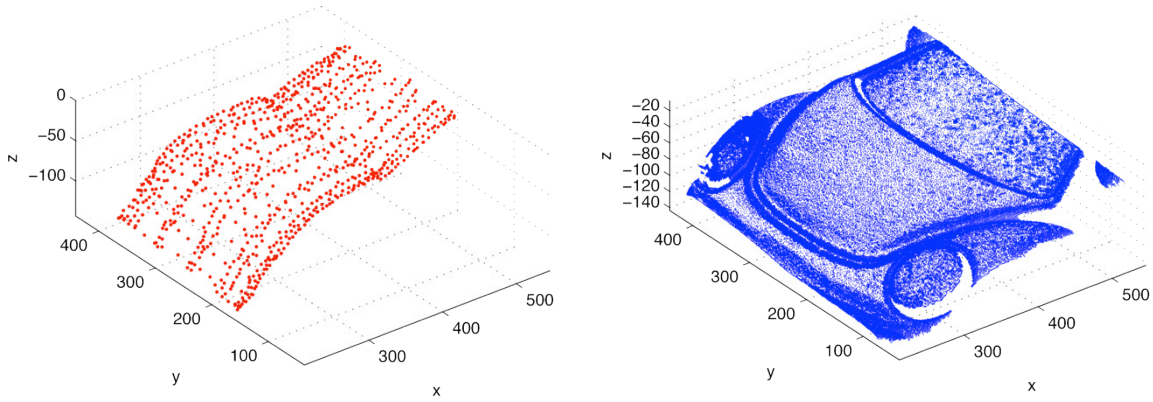
The second case study involves the measurement of multiple connected freeform surfaces of a toy car. The prototype Mobile Spatial coordinate Measuring System – MScMS-II for non-contact, indoors large-scale metrology (presented in [33]) is selected as the main measurement instrument. The MScMS-II is a touch-probe CMM manually operated in single-point mode. The peculiarity with respect to a conventional CMM consists of not having a solid frame to whom the machine coordinate system is referred to; instead, a wireless network of cooperating infra-red sensors, freely placed anywhere in the measurement space is used to localize the touch probe as it is manually operated over the surfaces to be measured [33]. While extremely flexible and capable of covering a wide range of dimensions and complex shapes (in-measurement repositioning of the networked sensors allows for overcoming any problem related to undercuts and can virtually extend the working area at will), the MScMS-II suffers from poor metrological performance if compared to the direct competitors in large-scale metrology (laser trackers and large-scale CMMs). This is mostly due to the difficulty of calibrating the constellation of networked sensors, which -before being able to accurately localize the touch probe- must be able to localize accurately each other within the measurement volume. Furthermore, a new calibration problem must be faced any time a sensor belonging to the constellation is moved in order to overcome the limitations of a fixed-frame CMM.

In this case study, it was interesting to investigate how the proposed approach to data fusion may help improving the metrological performance of the MScMS-II measurement. The fusion involves a high-density point set acquired with a structured light (SL) scanner, covering the same region where the fewer, sparse MScMS-II points are located.

The toy car is shown in Figure 13 (overall envelope dimensions: 507x350x912 mm). The specific surfaces subjected to measurements are shown in Figure 13.b, and include the front hood, windshield, and portions of the front bumper, fenders and lights. The HD point set (SL scanner) is comprised of $n_{HD} = 81790$ points (Figure 14.b); the LD set (MScMS-II) is comprised of $n_{LD} = 853$ points (Figure 14.a).

(a) (b)

Figure 13: Multiple connected freeform surfaces used as case study. Toy car (a); measured surfaces (b)



(a) (b)

Figure 14: Multiple connected freeform surfaces used as case study. 853 LD data set obtained via the MScMS-II measurement system (a), 81790 HD data measured by SL scanner (b) (all values in [mm])

The application of the MsDF model begins by processing the HD data. Computational concerns due to the very high density of the point set, suggested the application of a simpler variation of the procedure previously described in section 2. Thus, instead of fitting the HD points to a GP model, conventional interpolation was used to obtain a simpler triangulated model.

At the second stage of the MsDF, the linkage model connecting the HD to the LD data was computed: as usual with the goal of correcting the prediction of the first-stage model. Similarly to the previous case study, a linear model was used for the scaling function $\rho(\mathbf{v})$ while a GP model was used for the shifting function $\delta(\mathbf{v})$.

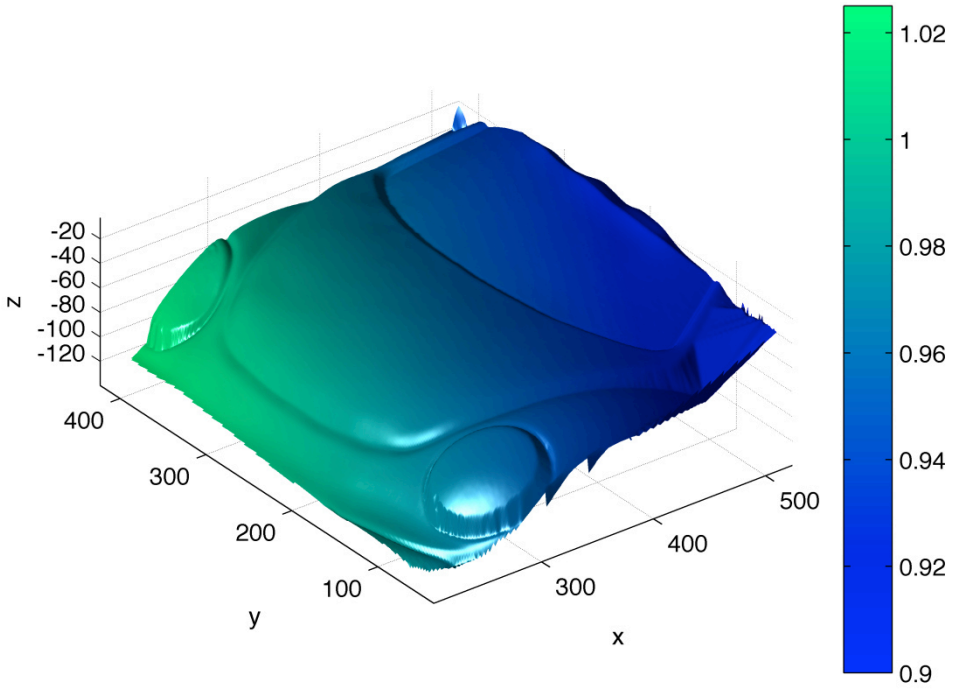
Figure 15 depicts the color-plot of the scaling and shifting functions $\rho(\mathbf{v})$ and $\delta(\mathbf{v})$ computed by the MsDF model. The scaling function produces a (small) compensation of the possible bias between the point cloud data available from the different sensors. Although the

point clouds provided by the two measurement devices (SL and MScMS-II) were first appropriately registered, a scale adjustment of the HD model was actually introduced by the scaling function, whose values range between 0.9 and about 1.02 (dimensionless scaling coefficients for the HD model).

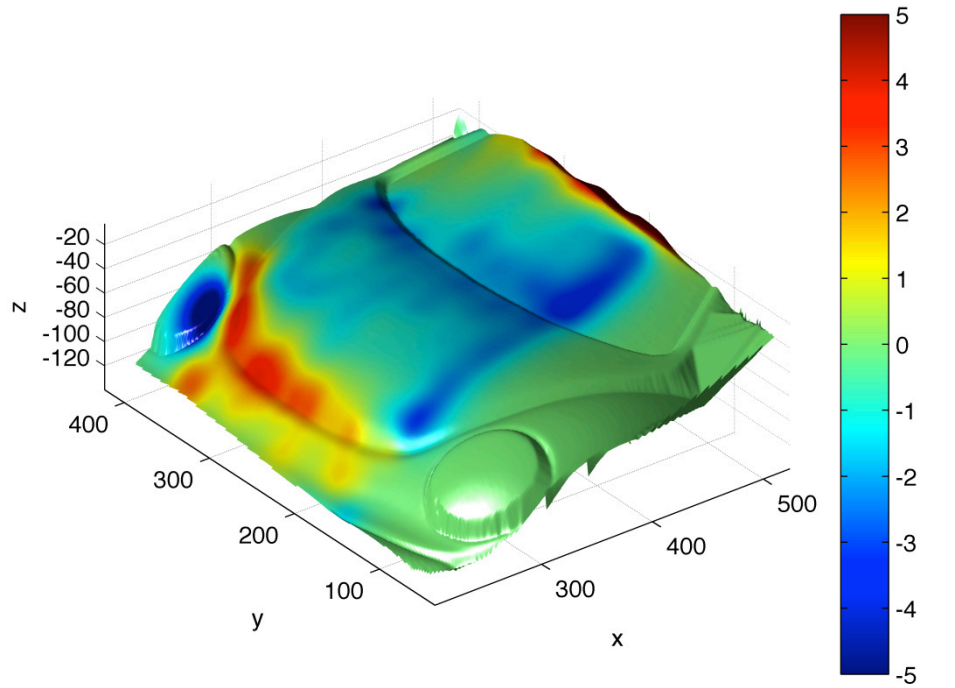
On the contrary, large local corrections are introduced by the shifting function, whose values range approximately between -5 and 5 millimeters (corrections in z-coordinate for the HD model). In particular, the shifting model results in important corrections in zones such as in the right border of the MScMS-II sampling surface as well as in the zone between the hood and bumper of the toy car model.

Two observations can be drawn from this case study. The first is a confirmation that the proposed MsDF approach can be effectively used to enhance reconstruction of surfaces in large-scale metrology. The second is that the scaling and the shifting adjustment functions of the MsDF model provide useful information to the analyst concerning measurement uncertainty of the metrology devices considered. In particular, at each location, information about variance of the scaling and shifting functions can be useful to evaluate how reliable the discrepancy map is. As an example, Figure 16.a shows the standard deviation of the shifting function.

The discrepancy map obtained by MsDF can be usefully exploited in order to decide where it is necessary to have additional measurement data at locations sampled with the LD device (the MScMS-II system in the case study). To this aim, Figure 16.b highlights zones where the shifting function (color axis in Figure 15.b) results statistically different from zero (type-I error rate 10%), given the standard deviation at the same location (color axis in Figure 16.a). Additional measurements of the LD system should be collected in these highlighted zones in order to investigate the causes of the statistically significant discrepancies between the LD and HD models.

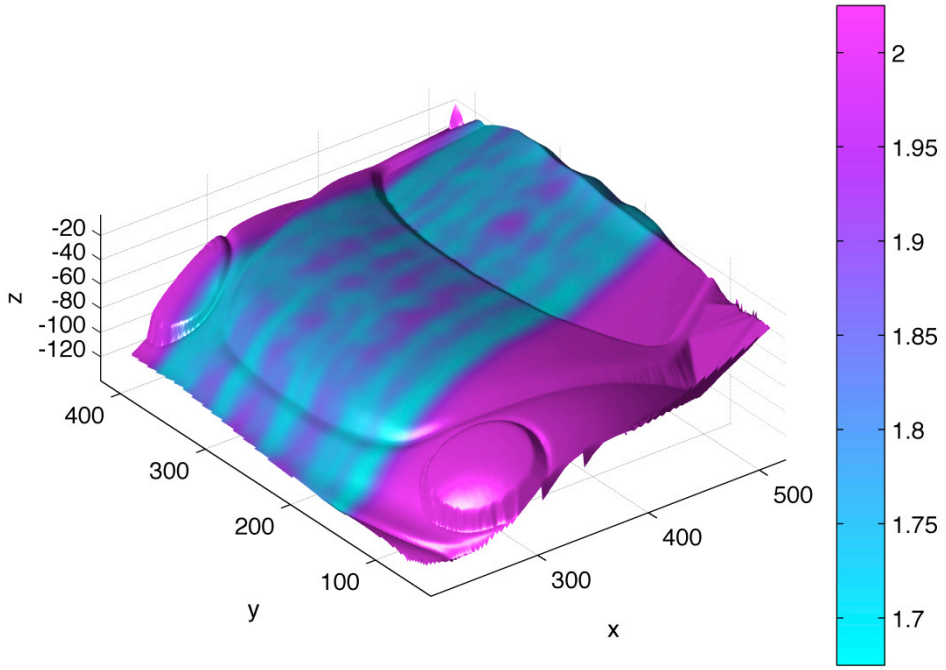


(a)

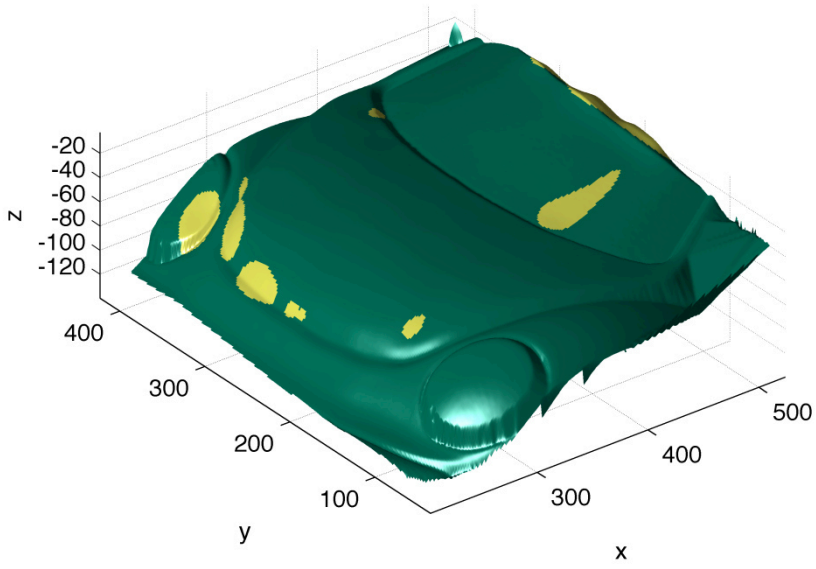


(b)

Figure 15: MsDF model. Color plot of the expected value of the scaling (a) and the shifting (b) functions shown on the reconstruction of the underlying surface obtained by means of the proposed MsDF model (all values in [mm])



(a)



(b)

Figure 16: Color plot of the standard deviation of the shifting function (a) and zones where the shifting function is statistically different from zero, 10% type-I error (b). (All values in [mm])

5. Investigating measurement uncertainty with the proposed MsDF model

Used as a means for identifying and reducing local bias in measurement, the MsDF approach is already serving as a powerful tool for operating on the uncertainty budget, acting in particular on some notable components of systematic error typically affecting most optical systems (e.g. local bias due to optical aberrations). In addition to that, as stated in the introduction, the MsDF approach provides further benefits; notably the estimation of the precision associated to the individual datasets before they are fused (random error sources in the uncertainty budget), and the possibility of observing how random error components associated to the individual datasets propagate through fusion, finally affecting the prediction interval associated to the fusion result.

To illustrate these additional benefits of the MsDF model, the same simple profile example previously introduced in Section 2 is used.

5.1 Estimating the random measurement error associated to the HD dataset

The random measurement error associated to an individual dataset can be assumed as an indication of precision of the instrument that was involved in the generation of the dataset (or better stated, the precision arising from the interaction of the instrument with the specific measurand). In the mathematical framework introduced in section 2.2, where the GP model in equation (1) is used to fit HD data generated by a single high-density sensor (first stage of the MsDF model), such random error is captured by the term $\varepsilon_{HD} \sim N\left(0,\sigma^2_{\varepsilon HD}\right)$. Obviously, ε_{HD} must be considered as part of the uncertainty budget, and the parameter $\sigma^2_{\varepsilon HD}$ (or $\sigma_{\varepsilon HD}$) is a quantitative indication of precision. The estimation of $\sigma_{\varepsilon HD}$ is a consequence of solving the model fitting problem onto the given dataset, therefore the first stage of the MsDF is already providing an indication of precision to be associated to the dataset, useful for uncertainty analysis. It should be noted that this estimate has been obtained without the need of replicate datasets, which is the conventional way one could follow to estimate precision of this type of profile data without resorting to GP model fitting. It remains to be determined how accurate is the estimate of $\sigma_{\varepsilon HD}$ obtained by simply fitting the dataset to the GP model.

In order to investigate this issue, as experimental procedure was set up consisting in simulating HD datasets where at each position $\mathbf{v}_i$ multiple z values were randomly generated, acting as measurement replicates in repeatability conditions. Three cases were

considered, consisting of either n=1, 2 or 5 z-value replicates at each position $\mathbf{v}_i$. A known standard deviation ($\sigma_{\varepsilon HD} = 2$) was used to generate the z-value replicates. On each generated dataset, the GP model described in equations (1-3) was used to estimate all the unknown parameters. In this case, the focus was on the estimate of the standard deviation of the noise term, i.e., $\hat{\sigma}_{\varepsilon HD}$. For each sample size n (n=1, 2 or 5), the entire procedure (random generation of the HD data set and estimation of the unknown parameters) was repeated 35 times. The resulting confidence intervals for $\hat{\sigma}_{\varepsilon HD}$ are shown in Figure 17. Clearly, the estimation is improved as the number of replicates increases (the width of the confidence interval decreases); nevertheless, in all cases the actual standard deviation (=2) is correctly located within the confidence interval, demonstrating that the estimation of dataset precision by GP model fitting is successful even with only one available dataset. In a more rigorous way, it can seen that the null hypothesis $H_0 : \sigma_{\varepsilon HD} = 2$ (vs. $H_1 : \sigma_{\varepsilon HD} \neq 2$) cannot be rejected in any scenario (p-values 0.3295, 0.7134 and 0.7518 for n= 1, 2 and 5, respectively)

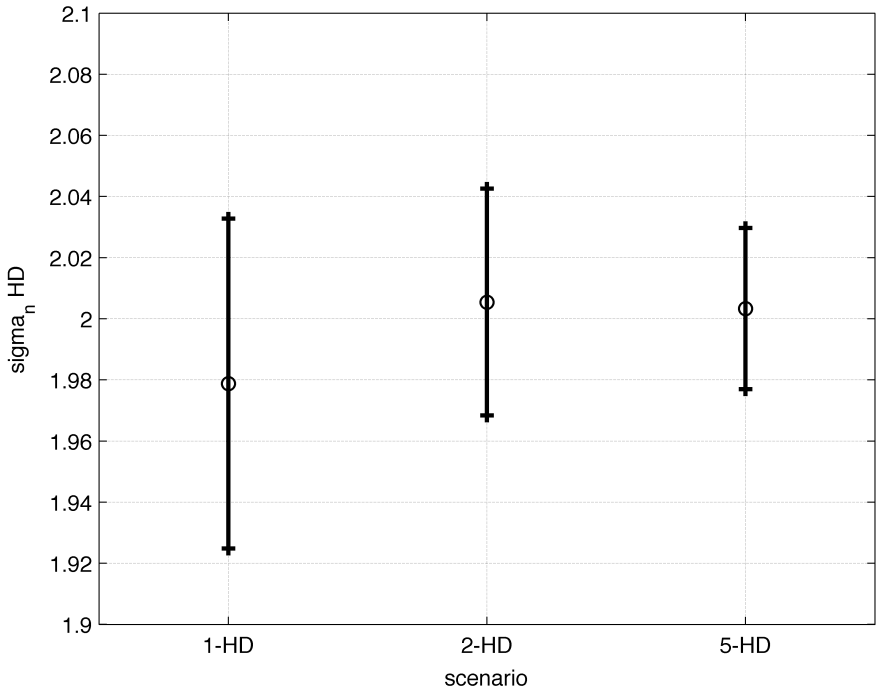


Figure 17: 95% Confidence intervals the estimated standard deviation of the random measurement error for the HD dataset, representative of the HD measurement precision (true value = 2). On the abscissa, the first number (1, 2 and 5) represents the number of replicated HD measurements at each location.

5.2 Estimating the random measurement error associated to the LD dataset

As shown previously in Section 2.6 (Figure 4), the low density of the LD dataset does not lead to a good fitting. Therefore it is not possible to use the same procedure described in Section 5.1, i.e., estimate the precision of the LD dataset by simply fitting a GP model using the LD dataset only. On the other side, this limitation intrinsic to the LD dataset is the main reason for the data fusion approach proposed in this paper, where information provided by an additional (HD) dataset is appropriately combined to the LD one to improve the overall measurement result.

The data fusion model can be used to estimate the LD system random error component by using the second stage of the proposed model. In this second stage, the linkage equation (4), contains two terms:

i) $\rho(\mathbf{v}_i)\hat{z}_{HD}(\mathbf{v}_i) + \delta(\mathbf{v}_i)$ which is supposed to consider the z-values fitted at the first stage on HD data and correct them for systematic deviations observed with respect to the LD data. Note that the random component of the HD measurement system should not be included in the linkage model, since $\hat{z}_{HD}(\mathbf{v}_i)$ is supposed to be a fitted value, i.e., after the noise has been canceled out.

ii) $\varepsilon_{MsDF} \sim N(0, \sigma^2_{\varepsilon MsDF})$ represents the noise which remains unexplained by the linkage model. In principle, this should represent the random part of the LD measurement error, possibly contaminated by an additional component due to under/overfitting. However, usually the presence of replicated response data in the linkage model (i.e., replicated LD data) reduces the over/underfitting problem, thus making the noise term an appropriate model of the randomness due to replicated masurement at the same location (i.e., random term of the measurment error).

In order to check whether that error term in the linkage model can correctly estimate the random part of the LD measurement error, a simulation study similar to the one described in section 5.1 was carried out. Different scenarios were considered, assuming to have n=1, 2, 5 and 10 replicated measurements available. In all these scenarios, the same number of replicates was assumed for both the (LD and HD) systems. The standard deviations of the random error components were assumed equal to 2 and 1 for the HD and LD system respectively. This assumption is consistent with the increased precision of touch probe CMM measurement (standard deviation set to 1) if compared to structured light scanning (standard deviation set to 2). For each scenario (characterized by a different value of the sample size n), the two-stage model was used to estimate all the unknown parameters. The whole procedure (random generation of the replicates and estimates of the unknown parameters) was repeated 35 times.

The resulting confidence intervals for $\hat{\sigma}_{\varepsilon MsDF}$, are shown in Figure 18. The correct value for the standard deviation is included in the confidence interval only when at least 2 z-value replicates per point are available in the LD dataset.

In summary, the estimation of the random error component to be included in the uncertainty budget of the LD dataset is possible with the method; however, the low density of the LD dataset requires the linkage stage, and at least 2 z-value replicates per dataset (i.e. both for the HD and the LD datasets).

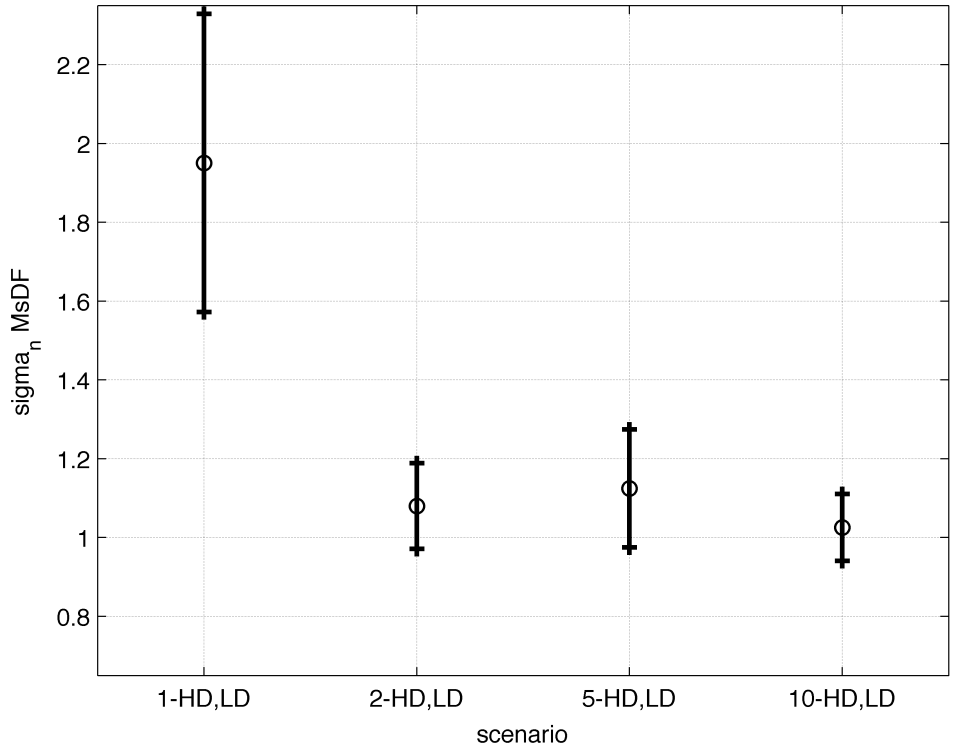


Figure 18: 95% confidence intervals for the estimated standard deviation of the random term of the linkage model, representative of the measurement error for the LD dataset, i.e. its measurement precision (true value 1).

3.3 Estimating how the random error components associated to the HD and LD datasets propagate through fusion and ultimately affect the prediction error of the MsDF model

In the previous sections, it was shown that the fusion method can be effectively used to estimate the random components of the measurement error for the two systems, provided

that i) a dense acquisition is available for the HD system and ii) (at least 2) replicated LD measurements are available at each location.

In this last section, the objective is to show how knowledge of the random component of the LD system error can be easily included in the fusion model. To this aim, the mean square prediction error MSPE (equation (6)) is used as performance index.

In a first scenario, consistent with real-life industrial practice, it can be assumed that $\sigma_{\varepsilon LD}$ is known (i.e. the touch probe CMM is metrologically characterized); in a second scenario it can be imagined that $\sigma_{\varepsilon LD}$ is estimated by the MsDF approach itself. Of course in both cases the precision of the LD dataset is actually represented by the standard deviation of the error term in the linkage model: $\sigma_{\varepsilon MsDF}$, but in the first scenario the value is simply plugged-in, while in the second scenario it needs to be estimated.

In Figure 19, confidence intervals for MSPE are shown. The first scenario is represented by the suffix “FIX” (fixed, i.e. plugged-in) and the second by the suffix “EST” (estimated). Analogously to Sections 5.1 and 5.2, datasets are simulated in 35 replicates, and replicate z-values are provided at each $\mathbf{v}_i$ location. However, for the first scenario (FIX), only z-value replicates of the HD set are needed (1,2 or 5 replicates), while for the second scenario (EST), z-value replicates of both the HD and LD sets are needed in equal numbers (also 1,2 or 5 replicates).

From Figure 19 it is clear that the MSPE is reduced as the number of replicate z-values increases, both in the EST and FIX scenarios. However, confidence intervals on the MSPE obtained in the FIX or EST scenarios are almost indistinguishible (i.e., show large overlap) if at least 2 z-value replicates are available.In other words, the prediction ability of the data fusion model is similar when random error associated to LD system is either known in advance or estimated by the fusion procedure itself, expecially if at least 2 z-value replicates are available for each location in both the LD and HD sets.

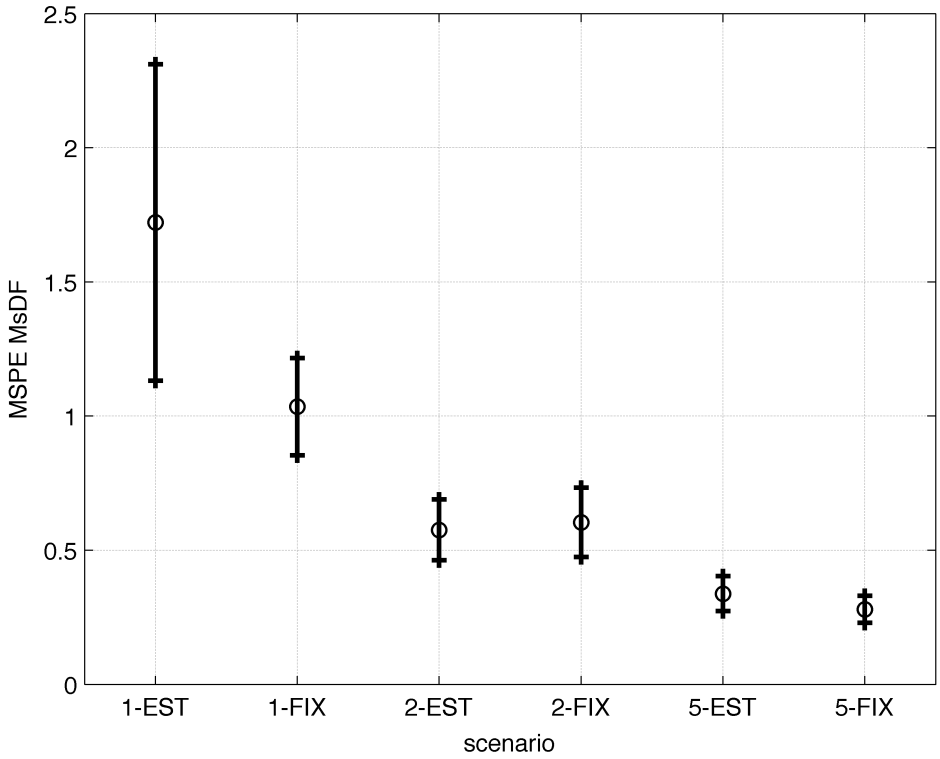


Figure 19: 95% confidence intervals of the Mean Square Prediction Error (MSPE) for the MsDF model. Simulated scenarios: 1,2,5-EST,FIX; where the number refers to the measurement replications at each point, FIX: known (EST: not known) value for the standard deviation to be plugged in the linkage model.

6. Conclusions and directions of further research

A statistical approach to combine high-density, low-quality data (e.g. from structured light scanners) with low-density, high-quality data (e.g. from touch probe CMMs) was presented. Advantages of the proposed method are explored considering both simulated and real case studies. In particular the real case studies focus on reconstruction of free-form shapes involving instruments dedicated to normal- and large-scale metrology. The main characteristics of the proposed approach can be summarized as follows: i) fusion can be performed even when sensors acquire data points at different locations; ii) fusion is performed by (locally) correcting the high-density, low-quality dataset (i.e. reducing bias) by means of the low-density, high-quality dataset; iii) uncertainty of the local correction can be estimated; iv) the proposed approach provides a means to estimate precision (random error component) associated to each dataset ahead of fusion if the dataset is sufficiently dense, otherwise it can still predict the random error component of a not-so-dense dataset at the fusion stage, through the linkage model; v) regardless of whether random error components of the datasets are known in advance or need to be computed by the method, the data

fusion approach provides means to follow the error components as they propagate through fusion and ultimately influence the final result. In fact, the data fusion approach provides as final results both the expected value and the uncertainty of the final geometry reconstructed at any location, in form of a prediction interval.

The proposed method involves the use of Gaussian processes (also known as kriging) as modeling tool to correct the high-density data and to guide the final reconstruction. Different parametric or nonparametric tools (based on spline or local regression) can be considered as well.

Different directions for further research can be outlined. First, the initial alignment of the point sets has shown to play a relevant role in the predictive ability of the different models. This is why more attention will be devoted to this preliminary step. In particular, comparison between the ICP algorithm used throughout the paper and different procedures will be performed.

Second, an additional model able to include the *a priori* knowledge of the uncertainty associated to each measurement system could be included. Third, a multi-stage extension of the two-stage model presented in this paper can be foreseen, where more than two sensors are used to measure the target geometry.

Eventually, the method proposed in this paper for reconstructing a geometry starting from multiple sensors could be appropriately included in the more general framework of inspection (i.e., comparison with the nominal shape) or statistical process monitoring (detecting out-of-control states of the machining or measuring processes).

ACKNOWLEDGMENTS

This work was partially funded by M.I.U.R. Grant: PRIN 2008 – “Large-scale coordinate metrology: study and realization of an innovative system based on a network of distributed and cooperative wireless sensors”, and it was partially supported by the European Union's Seventh Framework Programme (FP7/2007-2013) under grant agreement number 285075 – MuProD.

REFERENCES

[1] Hexagon Metrology Website USA. ”Multisensor & Optical Systems”. Hexagon Metrology Website USA. Accessed November 14, 2014. http://www.hexagonmetrology.us/products/multisensor-optical-systems/

[2] Nikon Metrology website EU. “Multi-sensor metrology systems”. Accessed November 14, 2014. http://www.nikonmetrology.com/en_EU/Products.

[3] Christoph R., Neumann H. J. Multi-Sensor Coordinate Metrology, Measurement of Form, Size, and Location in Production and Quality Control. Werth Messtechnik

GmbH 2013, ISBN: 3-478-93290-4. Available at: http://www.werth.de/us/navigation/presse/werth-newsworthy/book-presentation.html

[4] Carl Zeiss website USA. “Multisensor CMMs”. Accessed November 14, 2014. http://www.zeiss.com/industrial-metrology/en_us/products/systems/multisensor-cmms.html.

[5] Faro website USA. “Metrology” Accessed November 14, 2014. http://www.faro.com/en-us/products/metrology..

[6] Hexagon Metrology website USA. “Portable Measuring Arms.” Hexagon Metrology website USA. Accessed November 14, 2014. http://hexagonmetrology.us/products/portable-measuring-arms

[7] Leica Microsystems website. “3D Optical Surface Metrology System Leica DCM 3D.” Accessed November 14, 2014. http://www.leica-microsystems.com/products/light-microscopes/industrial-materials/upright-microscopes/details/product/dcm-3d/

[8] Liggins M, Hall D, Llinas J. Handbook of Multisensor Data Fusion: Theory and Practice, Second Edition: Taylor & Francis, 2008.

[9] Weckenmann A, Jiang X, Sommer KD, Neuschaefer-Rube U, Seewig J, Shaw L, et al. Multisensor data fusion in dimensional metrology. CIRP Annals - Manufacturing Technology. 2009:701–21.

[10] Campatelli G. Geometric Tolerance Evaluation Using Combined Vision–Contact Techniques and Other Data Fusion Approaches. In: Colosimo BM, Senin N, editors. Geometric Tolerances: Impact on Design, Inspection and Process Monitoring: Springer, 2010.

[11] Huang Y, Qian X, Chena S. Multi-sensor calibration through iterative registration and fusion. Computer-Aided Design. 2009;41 240-55.

[12] Carbone V, Carocci M, Savio E, Sansoni G, De Chiffre L. Combination of a vision system and a coordinate measuring machine for the reverse engineering of freeform surfaces. International Journal of Advanced Manufacturing Technology. 2001;17:263-71.

[13] Shen TS, Huang JB, Menq CH. Multiple-sensor integration for rapid and high-precision coordinate metrology. IEEE-ASME Trans Mechatron. 2000;5:110-21.

[14] Durrant-Whyte HF. Sensor models and multisensor integration. The International Journal of Robotics Research. 1988;7:97-113.

[15] Bradley C, Chan V. A complementary sensor approach to reverse engineering. Journal of Manufacturing Science and Engineering-Transactions of the Asme. 2001;123:74-82.

[16] Wang L, Ding H, Wang S. - Measurement Error Compensation Using Data Fusion Technique for Laser Scanner on AACMMs. 2010;- 6425.

[17] Jamshidi J, Owen GW, Mileham AR. A new data fusion method for scanned models. Journal of Computing and Information Science in Engineering. 2006;6:340-8.

[18] Forbes AB. Weighting observations from multi-sensor coordinate measuring systems. Measurement Science & Technology. 2012;23.

[19] Xia HF, Ding Y, Mallick BK. Bayesian hierarchical model for combining misaligned two-resolution metrology data. IIE Trans. 2011;43:242-58.

[20] B.M. Colosimo, Senin N. Geometric Tolerances. Impact on product design, quality inspection and statistical process monitoring. 2010.

[21] Senin N, Colosimo BM, Pacella M. Point set augmentation through fitting for enhanced ICP registration of point clouds in multisensor coordinate metrology. Robotics and Computer Integrated Manufacturing. 2012;29:39-52.

[22] Qian ZG, Seepersad CC, Joseph VR, Allen JK, Wu CFJ. Building surrogate models based on detailed and approximate simulations. J Mech Des. 2006;128:668-77.

[23] Xia HF, Ding Y, Wang J. Gaussian process method for form error assessment using coordinate measurements. IIE Trans. 2008;40:931-46.

[24] Castillo ED, Colosimo BM, Tajbakhsh S. Geodesic Gaussian Processes for the Reconstruction of a Free-Form Surface. Technometrics. 2013:accepted for publication.

[25] Colosimo BM, Cicorella P, Pacella M, Blaco M. From profile to surface monitoring: SPC for cylindrical surfaces via Gaussian Processes. Journal of Quality Technology, Vol.46, No. 2, April 2014, pp. 95-113.

[26] Cressie NAC. Statistics for Spatial Data, Revised Edition. New York, NY.: John Wiley & Sons, 1993.

[27] Rasmussen CE, Williams CKI. Gaussian Processes for Machine Learning. Cambridge - MA2006.

[28] Forrester A, Sobester A, Keane A. Engineering design via surrogate modelling: a practical guide. New York2008.

[29] Rasmussen CE, Nickisch H. Gaussian Processes for Machine Learning (GPML) Toolbox. J Mach Learn Res. 2010;11:3011-5.

[30] Cavallaro M, Moroni G, Petrò S. Performance evaluation of non contact measuring systems considering bias. Proceedings of the 4th International Conference on Advanced Research on Rapid Prototyping. Leiria Portugal2009.

[31] A.M. Bronstein, M.M. Bronstein, E. Gordon, Kimmel R. High-resolution structured light range scanner with automatic calibration. Department of Computer Science, Technion, Israel., 2003.

[32] Matlab. ICP Variants. Matlab Central File Exchange, 2013.

[33] Galetto M, Mastrogiacomo L, Pralio B. MScMS-II: an innovative IR-based indoor coordinate measuring system for large-scale metrology applications. The International Journal of Advanced Manufacturing Technology. 2011;52:291-302.

APPENDIX

GPs are very popular in spatial statistics for modeling a set of highly correlated random values, associated with a point set in a space [26]. Assume the point set represented by a discrete function of the type: $z(\mathbf{v}_i)$, $\mathbf{v}_i = (x_i, y_i)$ with $\mathbf{v}_i \in V \subset \mathbb{R}^2$, $i = 1,2,\cdots,n$, i.e. where the *z*-coordinate of the *i*-th point is expressed as a function of its position on the *x,y* plane. A statistical representation of the surface geometry can be defined as follows:

$$z(\mathbf{v}_i) = f(\mathbf{v}_i) + \varepsilon, \tag{a1}$$

where ε is a Gaussian noise of zero mean and constant variance σ_ε^2 and $f(\mathbf{v}_i)$ is a GP model defined as:

$$f(\mathbf{v}_i) = GP\left(m(\mathbf{v}_i), k(\mathbf{v}_i, \mathbf{v}_j)\right) \tag{a2}$$

where $m(\mathbf{v}_i)$ is the mean function and $k(\mathbf{v}_i, \mathbf{v}_j)$ is the covariance function of the GP. The mean function can be arbitrarily defined. In some cases, it can be even set equal to a constant value for every point in the space. Similarly, the covariance function can be arbitrarily defined, but common choices are models that guarantee GP to have some important properties such as *stationarity*, *isotropy* and *smoothness*. In the following, a stationary GP is considered, such that $k(\mathbf{v}_i, \mathbf{v}_j) = k(\mathbf{v}_i + \mathbf{h}, \mathbf{v}_j + \mathbf{h})$ for any displacement vector $\mathbf{h}$ and, hence, with a constant variance $k(\mathbf{v}_i, \mathbf{v}_i) = \sigma_z^2$.

Given the models for $m(\mathbf{v}_i)$ and $k(\mathbf{v}_i, \mathbf{v}_j)$, they are usually dependent on a set of parameters (collectively denoted by θ) that are unknown and that must be estimated from the actual measurement data $z(\mathbf{v}_i)$ with $\mathbf{v}_i \in V \subset \mathbb{R}^2$, $i = 1,2,\cdots,n$. Let vectors $\mathbf{z}$ and $\mathbf{m}$ represent respectively the values of functions $z(\mathbf{v})$ and $m(\mathbf{v})$, in the sampled location $\mathbf{v}_i$, i.e.

$$\begin{aligned}\mathbf{z} &= \left[\begin{array}{cccc} z(\mathbf{v}_1) & z(\mathbf{v}_2) & \cdots & z(\mathbf{v}_n) \end{array}\right] \\ \mathbf{m} &= \left[\begin{array}{cccc} m(\mathbf{v}_1) & m(\mathbf{v}_2) & \cdots & m(\mathbf{v}_n) \end{array}\right]\end{aligned}. \tag{a3}$$

Our maximum *a posteriori* estimate of θ occurs when the probability of the parameters given the dataset, say $p\left(\theta\middle|\mathbf{z},\mathbf{v}_1,\mathbf{v}_2,\cdots,\mathbf{v}_n\right)$, is at its maximum value. Assuming we have little prior knowledge about θ, this corresponds to maximizing the probability of the data given the parameters. The logarithm of this quantity is called the log marginal likelihood. In this paper, a numerical gradient-based optimization routine was used in order to find the values of parameters that optimize the marginal likelihood.

After the set of parameters θ has been estimated, we can predict $z(\mathbf{v})$ in any novel location $\mathbf{v}$. To this aim, we first calculate the covariance function among all of the possible combinations between the sampled locations $\mathbf{v}_i$ (for any $i = 1,2,\cdots,n$) and the unsampled one $\mathbf{v}$, summarizing our findings in vector $\mathbf{k}$:

$$\mathbf{k} = \left[\begin{array}{cccc} k(\mathbf{v},\mathbf{v}_1) & k(\mathbf{v},\mathbf{v}_2) & \cdots & k(\mathbf{v},\mathbf{v}_n) \end{array}\right]. \tag{a4}$$

Our best estimate of $z(\mathbf{v})$ in the unsampled location $\mathbf{v}$ is the mean value of the conditional distribution $z(\mathbf{v})\middle|\mathbf{z}$, i.e., $\hat{z}(\mathbf{v}) = E\left[z(\mathbf{v})\middle|\mathbf{z}\right]$. The uncertainty in our estimate is captured in the variance of the conditional distribution $z(\mathbf{v})\middle|\mathbf{z}$, i.e., $Var\left[z(\mathbf{v})\middle|\mathbf{z}\right]$. It can be possible to demonstrate that the conditional distribution $z(\mathbf{v})\middle|\mathbf{z}$ is a Gaussian distribution of mean and variance respectively equal to:

$$\begin{aligned} E\left[z(\mathbf{v})\middle|\mathbf{z}\right] &= m(\mathbf{v}) + \mathbf{k}\cdot\mathbf{K}^{-1}\cdot(\mathbf{z}-\mathbf{m})^T \\ Var\left[z(\mathbf{v})\middle|\mathbf{z}\right] &= \sigma_z^2 + \sigma_\varepsilon^2 - \mathbf{k}\cdot\mathbf{K}^{-1}\cdot\mathbf{k}^T \end{aligned}. \tag{a5}$$