Received: 10 December 2019 | Revised: 15 September 2020 | Accepted: 22 September 2020

DOI: 10.1002/qre.2786

SPECIAL ISSUE ARTICLE

WILEY

Function-on-function regression for assessing production quality in industrial manufacturing

Biagio Palumbo | **Fabio Centofanti** | **Francesco Del Re**

University of Naples Federico II, Naples, Italy

Correspondence
Biagio Palumbo, Department of Industrial Engineering, University of Naples Federico II, 80125 Naples, Italy.
Email: biagio.palumbo@unina.it

Abstract

Key responses of manufacturing processes are often represented by spatially or time-ordered data known as *functional data*. In practice, these are usually treated by extracting one or few representative scalar features from them to be used in the following analysis, with the risk of discarding relevant information available in the whole profile and of drawing only partial conclusions. To avoid that, new and more sophisticated methods can be retrieved from the functional data analysis (FDA) literature. In this work, that represents a contribution in the direction of integrating FDA methods into the manufacturing field, the use of function-on-function linear regression modelling is proposed. The approach is based on a finite-dimensional approximation of the regression coefficient function by means of two sets of basis functions, and two roughness penalties to control the degree of smoothness of the final estimator. The potential of the proposed method is demonstrated by applying it to a real-life case study in powder bed fusion additive manufacturing for metals to predict the mechanical properties of an additively manufactured artefact, given the particle size distribution of the powder used for its production.

KEYWORDS

additive manufacturing, functional data analysis, functional linear regression, particle size distribution, stress–strain curve

1 | INTRODUCTION

The fast technological developments of in-line sensing systems and non-contact acquisition architectures allow gathering huge amounts of data ordered by space or time from production processes. Those data are often referred to as *functional data* or *profiles*. The most common practice to handle them, however, mainly consist of extracting and analysing only one or few representative scalar features from each profile observation, based on engineering conjectures and/or using feature-oriented methods.[1] This approach is indeed advantageous and easy to interpret, but it may involve the risk of discarding non-trivial information and to veil relevant knowledge on the characteristics under study, possibly hidden in the profile. Functional data analysis (FDA) methods, which have emerged in the past decade in the statistical field,[2–5] represent a relevant extension to be considered in all those cases in which the researcher suspects complex influence among measured profiles. In particular, *functional regression* deserves attention as a practical tool to model the relationship between a quality characteristic (referred to as *response*) and one or more independent variables (referred to as *covariates*), in which at least

one of them is apt to be modelled as a function over a continuum. Excellent overviews are provided by Cuevas,[6] Morris,[7] Horváth and Kokoszka[3] and Ramsay and Silverman.[2] Functional regression, and in general FDA, under the assumptions that measured profiles in hand are intrinsically smooth, gives the advantage, over conventional multivariate data methods, of using the information in the slopes and curvatures (i.e., derivatives). The use of non-parametric smoothing methods to obtain functional data does not require strong distributional assumption that are in fact rarely fulfilled by real complex data. The main counterpart that hampers the spread of FDA among practitioners is the so-called pre-processing phase, which is mainly needed to obtain functional data objects by smoothing raw measurements, and by warping them into the same compact domain if needed. The present motivation-applied paper aims to contribute to the spread of FDA among practitioners and focuses on the setting in which both the response and covariate are functional and the relationship between them is linear, that is usually referred to as *function-on-function (FoF) linear regression model*. The estimation of FoF linear regression models was originally introduced by Ramsay and Dalzell[8] through piecewise Fourier basis. The first steps in the estimation development of FoF models are mainly due to Besse and Cardot,[9] who used spline-based approaches, and Ramsay and Silverman,[2] that proposed a general estimation method based on regularization. Then, Yao et al[10] presented a method based on the functional principal components decomposition[11] of both the response and covariate and Ivanescu et al[12] extended penalized functional regression (PFR) of Goldsmith et al[13] to the FoF setting. More recent articles on the FoF model are Luo and Qi[14, 15] and Centofanti et al.[16, 17]

In order to show its practical applicability and to highlight the benefits of a functional data approach, the FoF linear regression model is applied to a real-life case study in the additive manufacturing (AM) field. The impact of many characteristics of raw materials on the final properties of the produced parts is in fact still an open research issue and hampers the industrialization of additive technologies,[18–21] which are, on the other hand, also very costly. In particular, the considered AM process refers to a laser powder bed fusion (L-PBF) technique for metals, which enables the layer-wise production of complex-shaped components through spreading and selective laser melting of subsequent metal powder layers onto a substrate plate.[22–24]

As is known, not all the powder is melted to form the final component. Then, the final properties of the produced parts are affected by the reuse of powder, which may have been altered, from previous production runs.[25–27] In particular, the reuse may modify the size, in terms of the so-called PSD, and the shape of powder particles and may affect flowability and packing properties, usable layer thickness and thus, powder-processing behaviour. The particle size is affordably measurable and very critical for the mechanical properties of metal parts produced through laser melting processes.[28–30] The static mechanical properties are usually summarized into stress–strain curves that are very popular in materials science and engineering field.[31, 32]

In this setting, practitioners classically settle the analyses by extracting scalar features from both the sample PSD (e.g., 10th, 50th and 90th percentiles) and the observed stress–strain curve (e.g., yield strength, ultimate tensile strength and elongation at break).[33] The application of the FoF linear regression methods allows instead the estimation and prediction of the full stress–strain curve of produced part (i.e., the functional response) given the powder PSD (i.e., the functional covariate) that incidentally changes with powder reuse from previous runs.

More in general, both suppliers and users can benefit from the complete analysis of PSD for product specifications and manufacturing control, as well as for industrial engineering research and development. Indeed, the laser melting process parameters are usually tuned with respect to the distribution of the powder particle size. In addition, the ability to reuse metal powder as many time as possible, without undermining process stability and final part quality, opens up these manufacturing processes to the possibility of becoming increasingly popular as green technology.

The model estimation is based on a finite-dimensional approximation of the FoF regression by means of two sets of basis functions (one for the response and one for the covariate), and two roughness penalties to control the degree of smoothness of the final estimator. The robustness of the main model extra-parameter choice, such as the number and order of B-splines basis and the order of the differential operators, is then investigated by means of a cross-validation approach.

The model estimation uncertainty is quantified by means of a straightforward extension of the *resampling cases* bootstrap developed in the classical regression setting,[34, 35] which is broadly based on resampling with replacement of both response and covariate functional observations. To date, little work has been done on bootstrap for functional data. In this regard, it is worth mentioning De Castro et al[36] for a bootstrap method to evaluate the range of the forecasts of sulfur dioxide levels near a power plant and Cuevas et al[37] for a comparison of bootstrap confidence bands (obtained with different resampling methods) of several functional estimators. The literature on bootstrap for functional regression is even poorer, with the exception of González-Manteiga and Martnez-Calvo.[38] who obtained pointwise confidence intervals by means of a bootstrap procedure for functional linear model with scalar response and functional covariate.

Lastly, a Monte Carlo simulation study is performed to asses the predictive performance of the proposed functional regression method. This is done by comparing it with multivariate (non-functional) methods commonly used in the literature to deal with high-dimensional regression problems.

In Section 2, the proposed FoF regression approach is described. In particular, the estimation method and the model selection issues are presented in Sections 2.1 and 2.2, respectively. The bootstrap approach proposed is illustrated in Section 2.3, whereas Section 3 presents the Monte Carlo simulation study. The real-life case study in the AM field is presented in Section 4, and conclusions are reported in Section 5. An appendix contains additional details on the estimation method and a graphical residuals analysis for the real-life case study. All computations and plots are obtained through the software environment R.[39]

2 | METHODOLOGY

Let the set of observations (X_i, Y_i), for $i = 1, \dots, n$, of the functional covariate $X(s)$ and the functional response $Y(t)$, that are assumed to belong to the Hilbert space of square integrable functions $L^2(\mathcal{S})$ and $L^2(\mathcal{T})$, be defined for s and t in the compact intervals $\mathcal{S}$ and $\mathcal{T}$, respectively. Without loss of generality, we assume that X_i and Y_i, which refer to $X_i(s)$, for $s \in \mathcal{S}$, and $Y_i(t)$, for $t \in \mathcal{T}$, have zero-mean, for $i = 1, \dots, n$. In practice, this is obtained by subtracting the covariate and response sample means from X_i and Y_i, respectively, for $i = 1, \dots, n$. Then, the FoF linear regression model is defined as

$$Y_i(t) = \int_{\mathcal{S}} X_i(s)\beta(s,t)ds + \varepsilon_i(t) \quad t \in \mathcal{T}, \qquad i = 1, \dots, n, \tag{1}$$

where the regression coefficient β is in $L^2(\mathcal{S} \times \mathcal{T})$, the Hilbert space of bivariate square integrable functions defined on the interval $\mathcal{S} \times \mathcal{T}$; ε_i is a zero mean error process with a covariance function $K(t_1, t_2)$, with t_1 and $t_2 \in \mathcal{T}$ and independent of X_i.

2.1 | Model estimation

The coefficient function β estimator is based on the following penalized least squares estimator:

$$\hat{\beta} = \underset{\beta \in L^2(\mathcal{S}\times\mathcal{T})}{\operatorname{argmin}} \left\{ \sum_{i=1}^{n} ||Y_i - \int_{\mathcal{S}} X_i(s)\beta(s,\cdot)ds||^2 + \lambda_s ||\mathcal{L}_s^{m_s}\beta||^2 + \lambda_t ||\mathcal{L}_t^{m_t}\beta||^2 \right\}, \tag{2}$$

where $\mathcal{L}_s^{m_s}$ and $\mathcal{L}_t^{m_t}$ are the m_s-th and m_t-th order differential operators with respect to variables s and t, respectively, and $||\cdot||$ denotes the L^2-norm corresponding to the inner product $\langle f, g \rangle = \int fg$; λ_s and λ_t are smoothing parameters that control the trade-off between the roughness of the estimator $\hat{\beta}$ and the goodness of fit. As an example, when $\lambda_s = 0$ and $\lambda_t = 0$, $\hat{\beta}$ is the usual least-squares estimator, whereas for $\lambda_s \to \infty$ and $\lambda_t \to \infty$, $\hat{\beta}$ converges to a bivariate polynomial function with degree $\max(m_s, m_t) - 1$. Clearly, for any finite n, the optimization problem in Equation (2) has infinite solutions; and thus, restrictions on β must be placed in order to obtain a unique estimator. In this regard, by following the approach proposed in Chapter 16 by Ramsay and Silverman[2] or Chapter 8 of Horvath and Kokoszka,[3] we consider the approximation of β as double expansion in terms of L basis functions $\{\psi_i^s\}$, defined on $\mathcal{S}$, and M basis functions $\{\psi_j^t\}$, defined on $\mathcal{T}$, that is,

$$\beta(s,t) \approx \sum_{i=1}^{L}\sum_{j=1}^{M} b_{ij}\psi_i^s(s)\psi_j^t(t) = \boldsymbol{\psi}^s(s)^T \boldsymbol{B}\boldsymbol{\psi}^t(t) \quad s \in \mathcal{S}, t \in \mathcal{T}, \tag{3}$$

where $\boldsymbol{B} = \{b_{ij}\} \in \mathbb{R}^{L\times M}$ is the coefficient matrix, $\boldsymbol{\psi}^s = (\psi_1^s, \dots, \psi_L^s)^T$ and $\boldsymbol{\psi}^t = (\psi_1^t, \dots, \psi_M^t)^T$. By plugging Equation (3) in the model of Equation (1), the FoF model becomes

$$Y_i(t) = \boldsymbol{X}_i^T \boldsymbol{B}\boldsymbol{\psi}^t(t) + \varepsilon_i(t) \quad t \in \mathcal{T}, \qquad i = 1, \dots, n, \tag{4}$$

with $\boldsymbol{X}_i = \int_{\mathcal{S}} X_i(s)\boldsymbol{\psi}^s(s)ds$.

Thus, the problem of estimating β is reduced to the estimation of the unknown coefficient matrix $\boldsymbol{B}$. Indeed, the optimization problem in Equation (2) becomes

$$\hat{\boldsymbol{B}} = \underset{\boldsymbol{B}\in\mathbb{R}^{L\times M}}{\operatorname{argmin}} \left\{ \sum_{i=1}^{n} ||Y_i - \boldsymbol{X}_i^T \boldsymbol{B}\boldsymbol{\psi}^t||^2 + \lambda_s ||\mathcal{L}_s^{m_s}(\boldsymbol{\psi}^{sT}\boldsymbol{B}\boldsymbol{\psi}^t)||^2 + \lambda_t ||\mathcal{L}_t^{m_t}(\boldsymbol{\psi}^{sT}\boldsymbol{B}\boldsymbol{\psi}^t)||^2 \right\}, \tag{5}$$

where $\hat{\boldsymbol{B}}$ is the estimator of $\boldsymbol{B}$.

As shown in Appendix A, the estimator of $\boldsymbol{B}$ is

$$\operatorname{vec}(\hat{\boldsymbol{B}}) = \left[\boldsymbol{W}_Y \otimes \boldsymbol{X}^T\boldsymbol{X} + \lambda_s \boldsymbol{W}_Y \otimes \boldsymbol{R}_X + \lambda_t \boldsymbol{R}_Y \otimes \boldsymbol{W}_X\right]^{-1} (\boldsymbol{I} \otimes \boldsymbol{X}^T) \operatorname{vec}(\boldsymbol{Y}), \tag{6}$$

where $\boldsymbol{W}_X = \int_{\mathcal{S}} \boldsymbol{\psi}^s(s)\boldsymbol{\psi}^s(s)^T ds$, $\boldsymbol{W}_Y = \int_{\mathcal{T}} \boldsymbol{\psi}^t(t)\boldsymbol{\psi}^t(t)^T dt$, $\boldsymbol{R}_X = \int_{\mathcal{S}} \mathcal{L}_s^{m_s}[\boldsymbol{\psi}^s(s)]\mathcal{L}_s^{m_s}[\boldsymbol{\psi}^s(s)]^T ds$, $\boldsymbol{R}_Y = \int_{\mathcal{T}} \mathcal{L}_t^{m_t}[\boldsymbol{\psi}^t(t)]\mathcal{L}_t^{m_t}[\boldsymbol{\psi}^t(t)]^T dt$, $\boldsymbol{X} = (\boldsymbol{X}_1, \ldots, \boldsymbol{X}_n)^T$ and $\boldsymbol{Y} = (\boldsymbol{Y}_1, \ldots, \boldsymbol{Y}_n)^T$, with $\boldsymbol{Y}_i = \int_{\mathcal{T}} Y_i(t)\boldsymbol{\psi}^t(t)dt$.

Given the matrices $\boldsymbol{A} \in \mathbb{R}^{j\times k}$ and $\boldsymbol{B} \in \mathbb{R}^{l\times m}$, vec($\boldsymbol{A}$) indicates the vector of length jk obtained by writing the matrix $\boldsymbol{A}$ as a column-wise vector, whereas $\boldsymbol{A} \otimes \boldsymbol{B}$ represents the Kronecker product matrix of dimensions $jl \times km$. Then, an estimator of β is obtained as follows:

$$\hat{\beta}(s,t) = \boldsymbol{\psi}^s(s)^T \hat{\boldsymbol{B}} \boldsymbol{\psi}^t(t) \quad s \in \mathcal{S}, t \in \mathcal{T}. \tag{7}$$

Particularly, the prediction $\hat{Y}_*$ at given new realization X_* of the covariate X is defined as

$$\hat{Y}_* = \boldsymbol{X}_*^T \hat{\boldsymbol{B}} \boldsymbol{\psi}^t, \tag{8}$$

with $\boldsymbol{X}_* = \int_{\mathcal{S}} X_*(s)\boldsymbol{\psi}^s(s)ds$ and can be regarded as the estimated expectation of the response conditional to the covariate $X = X_*$.

2.2 | Model selection

The choice of the number L and M of basis functions is not crucial for penalized methods, because the roughness of the estimator is controlled by the smoothing parameters[40] λ_s and λ_t, which however must be chosen carefully. To this aim, the K-fold cross-validation[41] is the most popular method because of its simplicity and applicability. Broadly speaking, it consists of dividing the set of observations (X_i, Y_i), for $i = 1, \ldots, n$, into K roughly equal-sized parts. For each part $k = 1, \ldots, K$, the model is fit to the remaining $K - 1$ parts and the prediction error is calculated on the observation that belongs to the part k. Then, from the K estimates of the prediction error, the K-fold cross-validation estimated prediction error, denoted by $CV(\lambda_s, \lambda_t)$, is obtained as a function of the smoothing parameters λ_s and λ_t that are usually made to vary in a pre-specified grid of values. The values of λ_s and λ_t are chosen such that $CV(\lambda_s, \lambda_t)$ is minimum. However, we suggest whenever possible to thoroughly inspect the $CV(\lambda_s, \lambda_t)$ function and customarily choose the values of λ_s and λ_t to achieve a good prediction performance (i.e., the corresponding value of K-fold cross-validation estimated prediction error is near the minimum of $CV(\lambda_s, \lambda_t)$) and interpretability (i.e., λ_s and λ_t are large enough so that $\hat{\beta}$ results sufficiently smooth). The most common choice for K is 10.[41] Standard choices for the basis functions $\{\psi_i^s\}$ and $\{\psi_j^t\}$ are B-spline, Fourier and Wavelet basis.[2] In particular, we suggest, at least for the real-life case study presented in Section 4, to use the B-splines of order 4 (i.e. *cubic B-spline*) with evenly spaced knots, due to their good properties and wide applicability.[42] Lastly, the standard choice for the order of the differential operators $\mathcal{L}_s^{m_s}$ and $\mathcal{L}_t^{m_t}$ is $m_s = 2$ and $m_t = 2$. That is, the curvature of the coefficient function β is penalized in both s and t directions[2] when cubic B-splines are used. With regard to the predictive validation, the use of more advanced methods could be envisaged by extending, for example, the 0.632+ bootstrap method[43] to FoF linear regression models. However, to the best of the authors' knowledge, such cross-validation method is still not developed, and further researches will be needed to figure out all the implications in the functional setting.

2.3 | Bootstrap analysis

To assess the uncertainty of the estimator $\hat{\beta}$ of the coefficient function β, we propose a simple bootstrap procedure. Bootstrap is a re-sampling technique that allows assigning a measure of accuracy to a sample estimate.[35,44] It is extremely useful when the uncertainty of an estimate is either impossible or too difficult to be gauged by analytical calculations. The proposed bootstrap procedure is the following:

1. Sample with replacement from the set of observations $(X_1, Y_1), \dots, (X_n, Y_n)$ to obtain B bootstrap samples of size n $(X^*_{1b}, Y^*_{1b}), \dots, (X^*_{nb}, Y^*_{nb})$, for $b = 1, \dots, B$.
2. Fit the model in Equation (1) B times to obtain the estimates $\hat{\beta}^*_1, \dots, \hat{\beta}^*_B$ of the coefficient function β.
3. Obtain pointwise confidence intervals with confidence level $1-\alpha$ by using the $\frac{\alpha}{2}$ and $1-\frac{\alpha}{2}$ quantiles of the empirical distribution of $\hat{\beta}(s,t)$, estimated by using $\hat{\beta}^*_1(s,t), \dots, \hat{\beta}^*_B(s,t)$ for each $s \in \mathcal{S}$ and $t \in \mathcal{T}$.

Along with the uncertainty quantification of $\hat{\beta}$, the reliability of the prediction $\hat{Y}_*$ in Equation (8), given a new realization X_* of the covariate, should be evaluated as well. Let $e_1, \dots, e_n$ be the residuals, obtained as $e_i = Y_i - \boldsymbol{X}_i^T \hat{\boldsymbol{B}} \boldsymbol{\psi}^t$. Then, we propose the following procedure:

1. Sample with replacement from the set of observations $(X_1, Y_1), \dots, (X_n, Y_n)$ to obtain B bootstrap samples of size n $(X^*_{1b}, Y^*_{1b}), \dots, (X^*_{nb}, Y^*_{nb})$, for $b = 1, \dots, B$.
2. Fit the model in Equation (1) B times to obtain the estimates $\hat{\beta}^*_1, \dots, \hat{\beta}^*_B$ of the coefficient function β and thus the estimates $\hat{\boldsymbol{B}}^*_1, \dots, \hat{\boldsymbol{B}}^*_B$ of the coefficient matrix $\boldsymbol{B}$.
3. Obtain $e^*_1, \dots, e^*_B$ by sampling with replacement from $e_1, \dots, e_n$.
4. Compute $\delta^*_1, \dots, \delta^*_B$ realizations of the estimated prediction error δ^* as $\delta^*_i = \boldsymbol{X}_*^T \hat{\boldsymbol{B}}^*_i \boldsymbol{\psi}^t - (\boldsymbol{X}_*^T \hat{\boldsymbol{B}} \boldsymbol{\psi}^t + e^*_i)$, where $\hat{\boldsymbol{B}}$ is as in Equation (5).
5. Obtain the prediction limits at level $1-\alpha$ as $[\boldsymbol{X}_*^T \hat{\boldsymbol{B}} \boldsymbol{\psi}^t(t) - \Delta_{1-\frac{\alpha}{2}}(t), \boldsymbol{X}_*^T \hat{\boldsymbol{B}} \boldsymbol{\psi}^t(t) - \Delta_{\frac{\alpha}{2}}(t)]$ for $t \in \mathcal{T}$, where $\Delta_{\frac{\alpha}{2}}(t)$ and $\Delta_{1-\frac{\alpha}{2}}(t)$ are the $\frac{\alpha}{2}$ and $1-\frac{\alpha}{2}$ quantiles of the empirical distribution of $\delta^*(t)$, estimated by using $\delta^*_1(t), \dots, \delta^*_B(t)$ for each $t \in \mathcal{T}$.

In this case, the quantity to be predicted is $Y_* = \boldsymbol{X}_*^T \boldsymbol{B} \boldsymbol{\psi}^t + \varepsilon_*$, where the random error ε_* is assumed independent of $\varepsilon_1, \dots, \varepsilon_n$, and the point estimator is $\hat{Y}_* = \boldsymbol{X}_*^T \hat{\boldsymbol{B}} \boldsymbol{\psi}^t$, as shown in Equation (8). Thus, to assess the accuracy of $\hat{Y}_*$, we estimate the pointwise distribution of the prediction error $\delta = \hat{Y}_* - Y_* = \boldsymbol{X}_*^T \hat{\boldsymbol{B}} \boldsymbol{\psi}^t - (\boldsymbol{X}_*^T \boldsymbol{B} \boldsymbol{\psi}^t + \varepsilon^*)$ with the pointwise empirical distribution of $\delta^*_i = \boldsymbol{X}_*^T \hat{\boldsymbol{B}}^*_i \boldsymbol{\psi}^t - (\boldsymbol{X}_*^T \hat{\boldsymbol{B}} \boldsymbol{\psi}^t + e^*_i)$. The proposed prediction limits are inspired by the bootstrap prediction limits in the classical regression setting.[35]

3 | SIMULATION STUDY

In this section, the performance of the estimator described in Section 2.1 for the FoF linear regression model in Equation (1) is investigated. In particular, we consider the scenario where the coefficient function β is defined as follows:

$$\beta(s,t) = s + 10t + 10s^2 + t^2 \quad s,t \in [0,1] \times [0,1]. \tag{9}$$

It is worth noting that the shape of the coefficient function is inspired by that in the real-life case study of Section 4, and, without loss of generality, we consider $\mathcal{S} \times \mathcal{T} = [0,1] \times [0,1]$.

The set of observations (X_i, Y_i) are generated as follows. The covariates X_i and the error ε_i are obtained as $X_i = \sum_{j=1}^{16} c_{ij} \Psi^x_i + 2\Phi + 1$ and $\varepsilon_i = r \sum_{j=1}^{20} d_{ij} \Psi^\varepsilon_i$. The constants c_{ij} and v_{ij} are obtained as independent realizations of the standard normal random variable; the basis $\Psi^x_i(s)$ and $\Psi^\epsilon_i(s)$ are cubic B-splines with an evenly spaced knot sequence and their number (i.e., 16 and 20) has been randomly chosen between 10 and 50; Φ is the density function of a normal random variable with mean 0.5 and standard deviation 0.1. The constant r is chosen such that the signal-to-noise ratio function $SN \doteq \int_{\mathcal{T}} \mathrm{Var}[E(Y_i|X_i)] / \int_{\mathcal{T}} \mathrm{Var}(\varepsilon_i)$ is equal to 4. Then the realizations of the response Y_i are readily obtained, given the coefficient function β. For each run, a dataset composed of a training and test sets of n and $N = 4000$ observations, respectively, is generated. The former is used to estimate the coefficient function, while the latter to assess the predictive

performance. We consider 100 runs and sample sizes $n = 50, 100, 200$. The predictive performance is measured through the prediction mean-squared error (PMSE), defined as

$$\text{PMSE} = \frac{1}{N} \sum_{(X,Y)\in T} \int_0^1 \left(Y(t) - \hat{Y}(t)\right)^2 dt, \tag{10}$$

where $\hat{Y}$ is a prediction of Y in the test set T obtained through the observations in the training set.

The proposed estimator is hereinafter referred to also as SMOOTH and is compared with four competitors which are representative of non-functional methods commonly used in the literature to deal with high-dimensional regression problems. In particular, the first two methods provide a response prediction via the principal component regression[45] on either the coefficients of a spline decomposition (hereinafter referred to as BPCR) or a regular grid of discrete values (referred to as DPCR) of both the covariate and response. Moreover, the partial least squares (PLS) regressions[46] on either the coefficients of a spline decomposition (hereinafter referred to as BPLS) or a regular grid of the discrete values (referred to as DPLS) of both the covariate and response are also considered. The SMOOTH estimator is computed by using $L = M = 50$ cubic B-splines with evenly spaced knots, $m_s = m_t = 2$, and, with smoothing parameters λ_s and λ_t chosen by means of a 10-fold cross-validation. For the BPCR and BPLS, 50 cubic B-splines with evenly spaced knots are also considered, whereas the DPCR and DPLS are computed through a discrete grid of 50 values. The the percentage of total variation retained by the principal components used in BPCR and DPCR as well as the number of retained components in BPLS and DPLS are chosen through 10-fold cross-validation. Note that, in order to ensure a fair comparison, all the methods considered are based on the same cross-validation based model selection criterion.

The performance of the considered estimators in terms of PMSE for $n = 50, 100, 200$ is displayed in Figure 1. It is readily clear that the functional SMOOTH method outperforms all the competitors for all training set sample size n. In particular, the performance differences among the SMOOTH method and the competitors decrease as the number of observations in the training set increases. This is expected because all the considered estimators shall converge in a large number. Moreover, the performance of the SMOOTH estimator is less affected by n than the competitors. This means the functional method is able to better account for the information in the available observations than the competitors, especially for small training set sample size n. Finally, by looking at PMSE variability, the SMOOTH method is also providing the more stable predictions. Among the competitors those based on PLS are the better, as expected.[47]

4 | A REAL-LIFE CASE STUDY IN LASER POWDER BED FUSION ADDITIVE MANUFACTURING

As stated before in the Introduction, a real dataset in L-PBF for metals manufacturing is analysed.[33] The effect of the number of powder reuses on the tensile properties of the produced AlSi10Mg parts is estimated through scalar one-way analysis of variance (ANOVA). Nine consecutive runs were carried out, starting from as-received powders only, in the first, and from a mix of as-received and reused powder (coming from the previous run), in the following eight runs. In particular, at the end of each run, all the powder left in the whole system (from the build chamber and from both the powder feeding and the overflow compartments) was removed, sieved together, and loaded again in the system for the subsequent run. The work flow used to test the effect of powder reuse is shown in Figure 2, in which the variables considered also in the present work are highlighted by a bolded frame.

As shown in Figure 3, each AM-build hosted six cylindrical bars for tensile test, six near net shape specimens for fatigue test, six cuboids for metallographic analysis and two lattice spheres. Both tensile bars and fatigue specimens were tested after machining, that was necessary to achieve the standard geometry required for mechanical testing. All builds were produced by a EOS EOSINT M280 3D printer with optimal parameters provided by the producer through the *EOS Part Property Profile AlSi10Mg Speed 30 μm*. In this work, only the cylindrical tensile specimens, indicated by the white arrows in Figure 3, are considered (six in each run, except in the seventh and the last ones for which only five specimens were available), to evaluate the mechanical properties achieved under the different processing conditions due to the different number of reuses. At each run, a metal powder sample was collected and examined. In particular, a Malvern MS2000 equipment for laser diffraction was used to measure the PSD of each sample according to ASTM B822-17.[48] Standard tensile tests were conducted by an Instron 1185 TSTM at room temperature and with the displacement rate of cross head equal to 0.0075 mm/s, according to ISO 6892:2016[49] and ASTM E8/E8M-16a[50] standards.

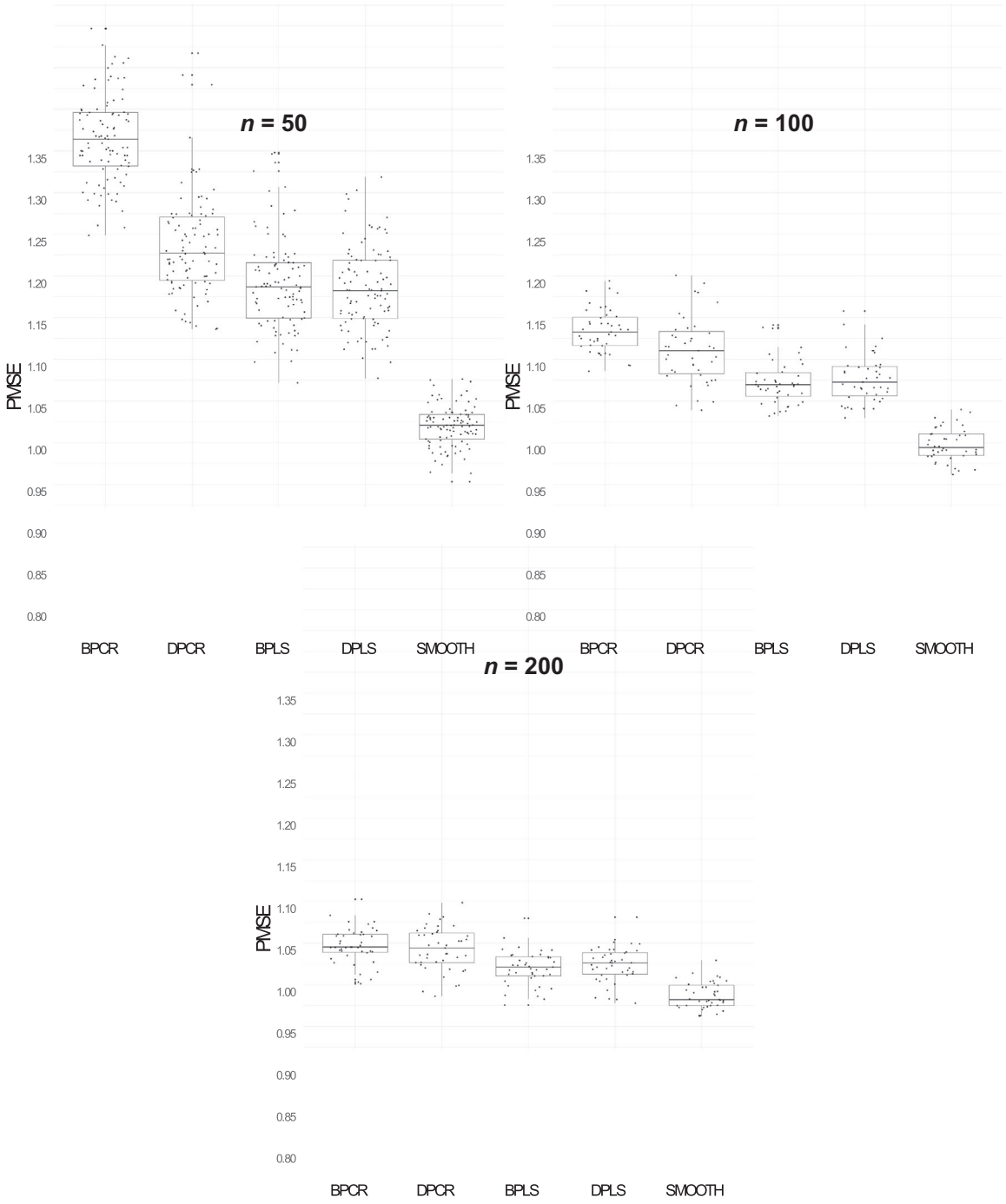

FIGURE 1 The PMSE for the BPCR, DPCR, BPLS, DPLS and SMOOTH estimators for $n = 50, 100, 200$

Detailed size information on the powder particle size is usually provided through a histogram, even if, as stated before, more common approach refers only to the 10th, 50th and 90th sample percentiles, denoted by *D10*, *D50*, and *D90*, respectively. An extensive explanation of tensile strength testing and stress–strain curves is beyond the objectives of this work so just brief hints about them will be given below. The presented description is mainly referred to metals. Other materials, such as plastics or ceramics, can exhibit different behaviours and kind of curves. Detailed information can be found in Davis.[32] As it is known, stress–strain curves are widely used in materials science and engineering to describe, for a given material, the relationship between the applied stress, denoted by σ, and the resulting strain (or elongation) exhibited by the material itself, denoted by ϵ. The former is calculated as the ratio of the applied load (orthogonal to the cross section) and the cross section of the specimen and is measured in pascal. The latter is the length of elongation exhibited by the

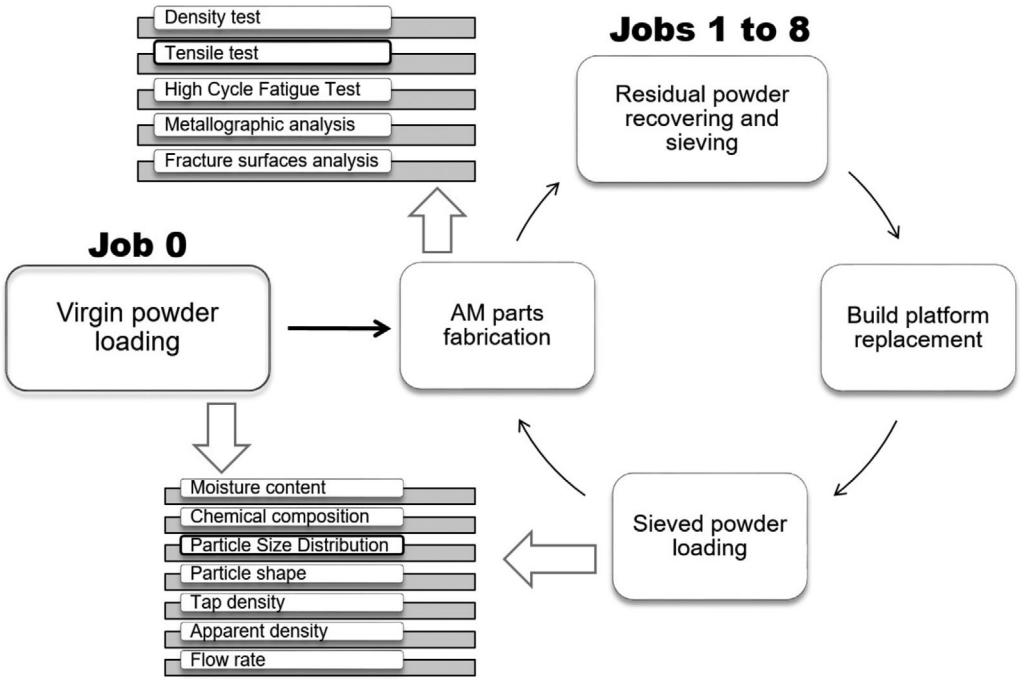

FIGURE 2 Work flow to test the effect of powder reuse in a L-PBF process[33]

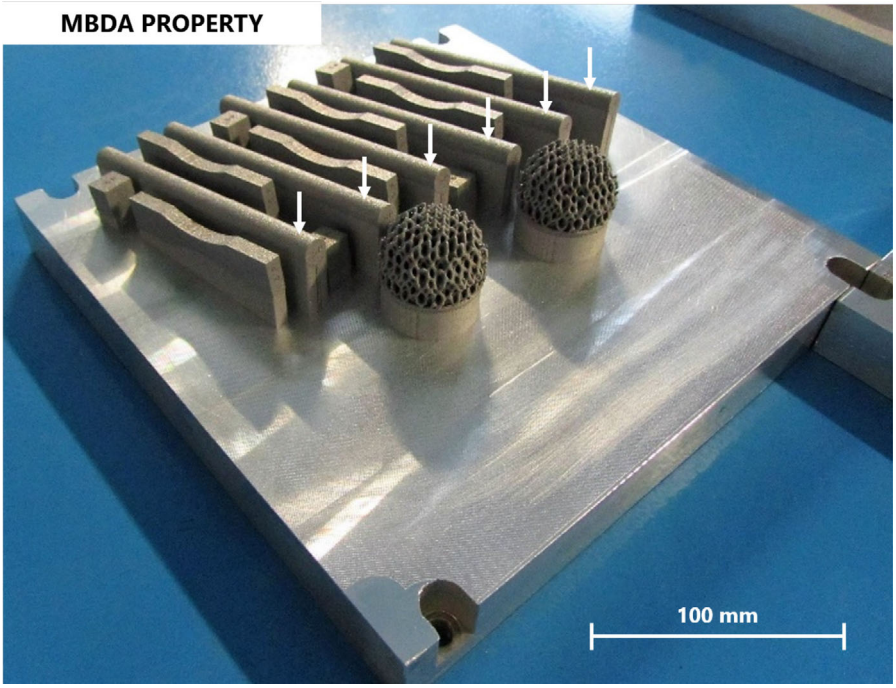

FIGURE 3 Example of an AM build for studying the effects of powder reuse times (the white arrows indicate the cylindrical tensile specimens considered in this work)[33]

material under the applied load divided by the original length of the material. Since both these lengths are usually measured in the same unit, strain can be considered unitless and is generally expressed as a percentage. A $\sigma - \epsilon$ curve provides information on the degree of ductility, or brittleness, of a material and on the maximum loads that it can withstand before it breaks. Some scalar parameters are usually derived from $\sigma - \epsilon$ curves to define the mechanical properties of the specific material.[33] The most commonly used are the *Young's modulus* (or *elastic modulus* or *modulus of elasticity*), E, the *yield strength,* Y_S, the *ultimate tensile strength, UTS*, and the *elongation at break, %A*.

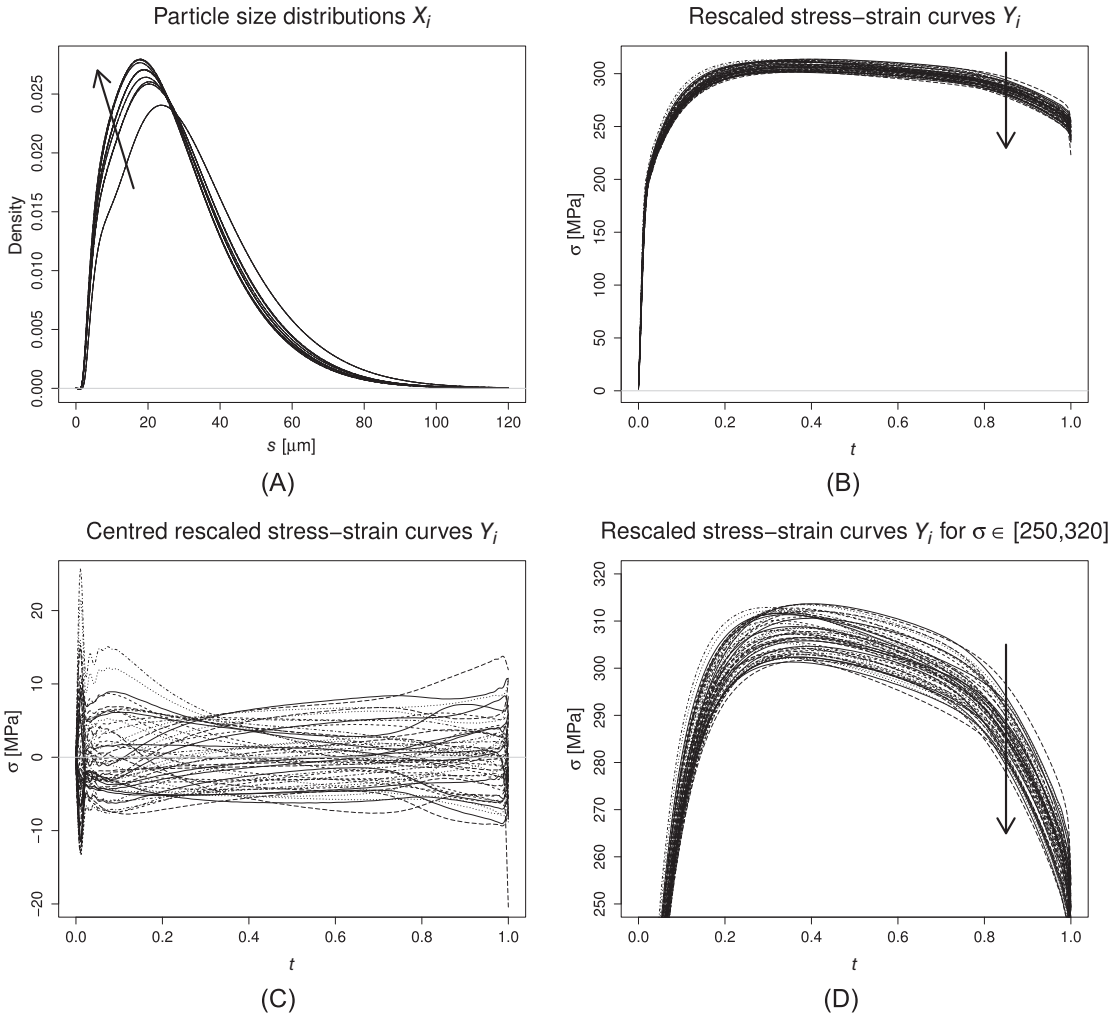

FIGURE 4 (A) Estimated pdfs of powder particle size; (B) rescaled stress–strain curves; (C) centred rescaled stress–strain curves and (D) rescaled stress–strain curves for σ values between 250 and 320 MPa. The arrows indicate how the observed profiles change with the powder reuse times

4.1 | Implementation details and results

The dataset available for the analysis, consisting of 52 $\sigma - \epsilon$ and 9 PSD curves, can be rearranged by coupling the estimated probability density functions (pdf) of the particle size, denoted by X_i, and the corresponding stress–strain curves, denoted by Y_i, for $i = 1, \dots, n$. The X_is are obtained through the kernel density estimation method[51] proposed by Charpentier and Flachaire,[52] that ensures positiveness of domain points $s \in \mathcal{S} = [0, 120]$ of the estimated pdfs. In contrast, the stress–strain curves Y_i are not necessarily defined on the same domain and thus they need to be opportunely rescaled on a common domain $\mathcal{T} = [0, 1]$ with respect to the elongation. Covariate and rescaled response observations are displayed in Figure 4 for illustrative purposes.

The density functions are strictly linked to the number of metal powder reuses. We have more than one observation with the same covariate function. Strictly speaking, we have in total $n = 52$ observations instead of 54 due to two missing specimens.

The basis functions $\{\psi_i^s\}$ and $\{\psi_j^t\}$ used to approximate β in Equation (3) are cubic B-splines with evenly spaced knots defined on the domains $\mathcal{S}$ and $\mathcal{T}$, respectively. By means of 10-fold cross-validation, the optimal smoothing parameters λ_s and λ_t that trade-off predictive performance and interpretability of the resulting estimator are found (as explained in Section 2.2) and both are set equal to 10^3. Accordingly, the truncation parameters, L and M, in Equation (3) are both set equal to 30. The estimator $\hat{\beta}$ is plotted in Figure 5 as a function of $s \in \mathcal{S}$ at different $t = 0.1, 0.2, \dots, 0.9$ along with level 1-α pointwise bootstrap confidence intervals obtained as described in Section 2.3, with $\alpha = 0.05$ and $B = 100$.

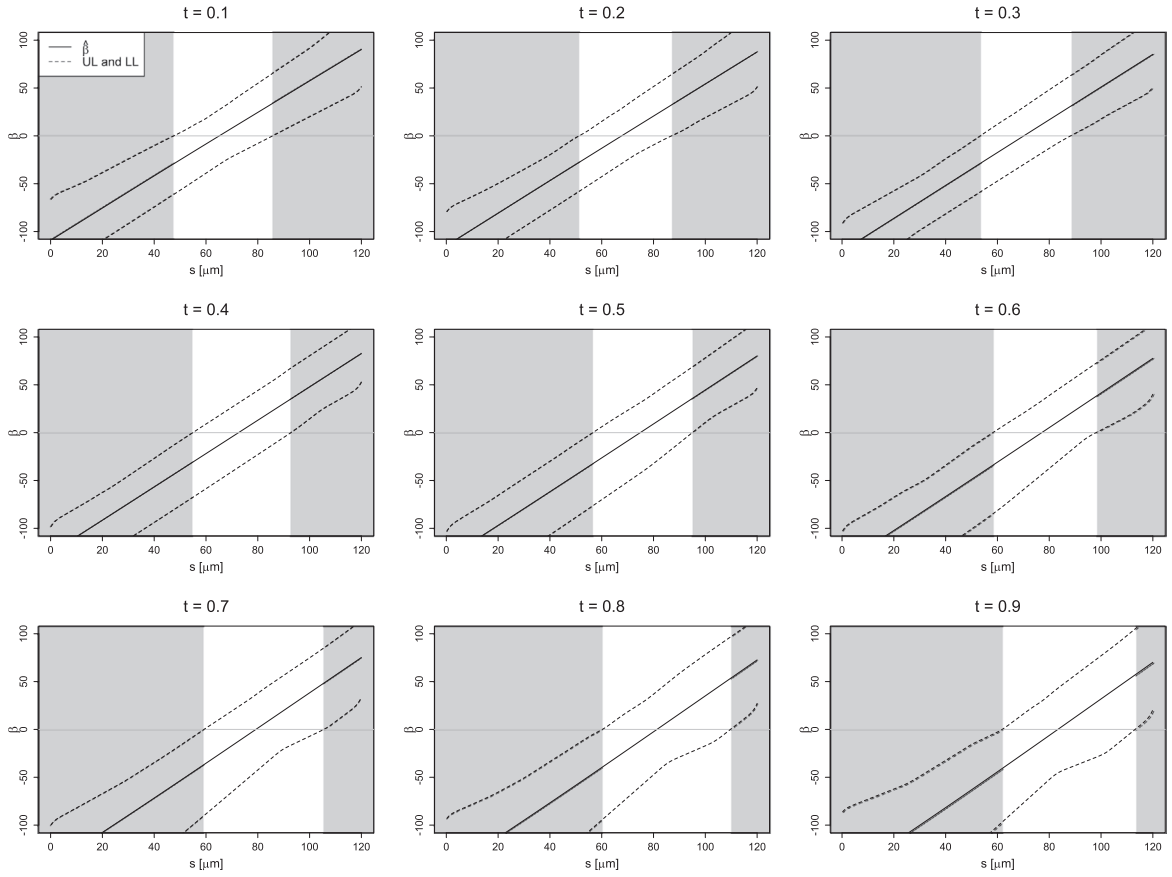

FIGURE 5 Coefficient function $\beta(s,t)$ estimate evaluated at $t \in \{0.1, 0.2, \dots, 0.8, 0.9\}$, with upper and lower limits (dotted lines) of the level 1-α pointwise bootstrap confidence intervals, for $\alpha = 0.05$. The shadowed regions indicate where the bootstrap bands are both above or below zero

From Figure 5, it appears that $\hat{\beta}$ is significantly different from zero in the shadowed regions for all $t \in \{0.1, 0.2, \dots, 0.8, 0.9\}$. Few (resp., many) small particles produce higher (resp., lower) stress values in the stress–strain curve. This behaviour is related to the effect of small particles (and thus big ones) particles on the powder packing factor, that affects how particles bind together during laser melting and therefore the resulting mechanical properties of the produced parts.[53]

As t increases, that is moving along the stress–strain curve, the effect of the particle size distribution on the stress–strain changes after the ultimate tensile strength is reached and the necking occurs, corresponding to $t \simeq 0.3 - 0.4$. From this point, the width of the shadowed region on the left (resp., right) progressively increases (resp., reduces).

For instance, to appreciate the prediction performance of the FoF linear model, the centred and the non-centred predictions of one observation, randomly selected, of the $\sigma - \epsilon$ curve are shown in Figure 6, along with the true centred and non-centred realizations of the response. Moreover, level 1-α pointwise bootstrap prediction limits, calculated as described is Section 2.3 with $\alpha = 0.05$ and $B = 100$, are displayed as well.

It is clear that the FoF linear regression model estimated by the procedure illustrated in Section 2 provides good predictions, and thus could be used to predict the mechanical properties of an artefact produced by the powder bed fusion technique, given the particle size distribution of the powder used for its production. The validity of the bootstrap procedures as well as the cross-validation model selection criterion rely on the hypothesis of independent and identically distributed (i.i.d.) functional errors. We tested this hypothesis by means of the functional portmanteau independence test,[54] where a p-value equal to 0.56 does not allow rejecting the null hypothesis of i.i.d. errors. The latter hypothesis is furthermore supported by the graphical residual analysis reported in Appendix B.

The proposed method performance is substantially better than that of the competitor methods, as shown in the simulation study of Section 3. However, as a proof of concept, we applied also the BPLS method as it resulted as the best competitor. The cross-validation mean squared error (MSE) achieved by the BPLS method is 7.12 and is higher than that

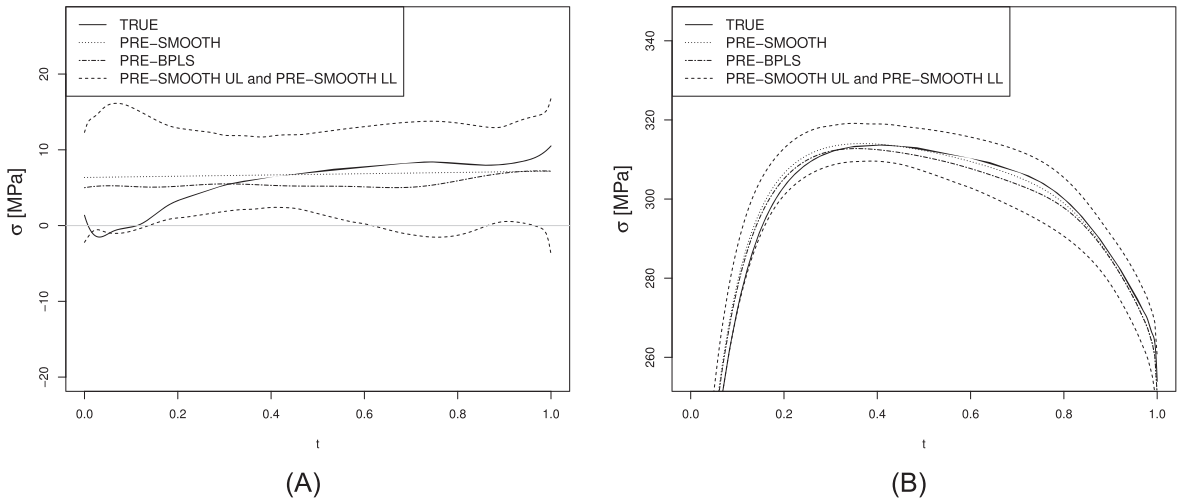

FIGURE 6 (A) Centred prediction for the proposed (PRE-SMOOTH) method of a randomly chosen response observation, its actual centred realization (TRUE) and the level $1-\alpha = 0.95$ pointwise bootstrap prediction upper (PRE-SMOOTH UL) and lower (PRE-SMOOTH LL) limits; (B) the corresponding non-centred $\sigma - \epsilon$ curve for σ values between 250 and 350 MPa. The centred prediction and the corresponding non-centred $\sigma - \epsilon$ curve for the BPLS method (PRE-BPLS) is also depicted

of the proposed method, which is equal to 6.86. The centred and the non-centred predictions for the BPLS method of a randomly chosen response observation of the $\sigma - \epsilon$ curve are superimposed on Figure 6.

The prediction curve obtained through the BPLS method is very similar to that obtained through the proposed method, even though the latter provides a better estimate the true function in a larger portion of domain, that is for $t \in [0.4, 1]$. Although the competing methods may achieve comparable prediction performance, due to their non-functional nature, they are not able to estimate the functional relation between the covariate and the response. This drawback reduces the interpretability of competing methods with respect to the proposed one and thus seriously affects the impact of the analysis.

Note that the application of the proposed method relies on the choice of the type of splines, that is, the number L and M of basis and the their orders or_s and or_t, and the order m_s and m_t of the differential operators used in Equation (2). In what follows, we investigate the robustness (in terms of predictive performance) of the results in the real-life case study of Section 4 with respect to these choices. In particular, we consider without loss of generality the case $L = M = nb$, $or_s = or_t = or$ and $m_s = m_t = m$. Figure 7 shows the cross-validation mean squared error (MSE) of the proposed method as a function of the number of basis, the B-spline and the differential operator order. Note that the allowed spline order must be larger than or equal to $m + 2$. The predictive performance is shown not to be significantly affected by the number of basis as well as by the spline and differential operator orders. Indeed, the cross-validation MSE behaviours, as a function of nb, or and m, are practically overlapping one each other. In particular, the behaviour of the cross-validation MSE as function of the number of basis is expected because the two penalties in Equation (2) regularize the final estimator by avoiding the over-fitting problem. Moreover, under-fitting is also prevented because $nb = 10$ is sufficient to capture the overall behaviour of the coefficient function. In summary, the spline type and differential operator orders do not substantially impact the predictive performance of the proposed method, and thus, from a practical point of view, the analysis results.

5 | CONCLUSIONS

Assessing production quality of AM products represents a great challenge due to the complexity of the data generated by this kind of processes which are subject, for example, to the use of different raw materials, processing parameters, environmental conditions and post-processing operations. To embrace the fully functional nature of modern AM process data, we propose the use of a function-on-function regression model based on non-parametric smoothing methods, which does not require strong distributional assumptions rarely fulfilled by real complex data, and two roughness penalties to avoid over-fitting of the coefficient function estimator. By applying it to a real-life case study in the laser powder bed

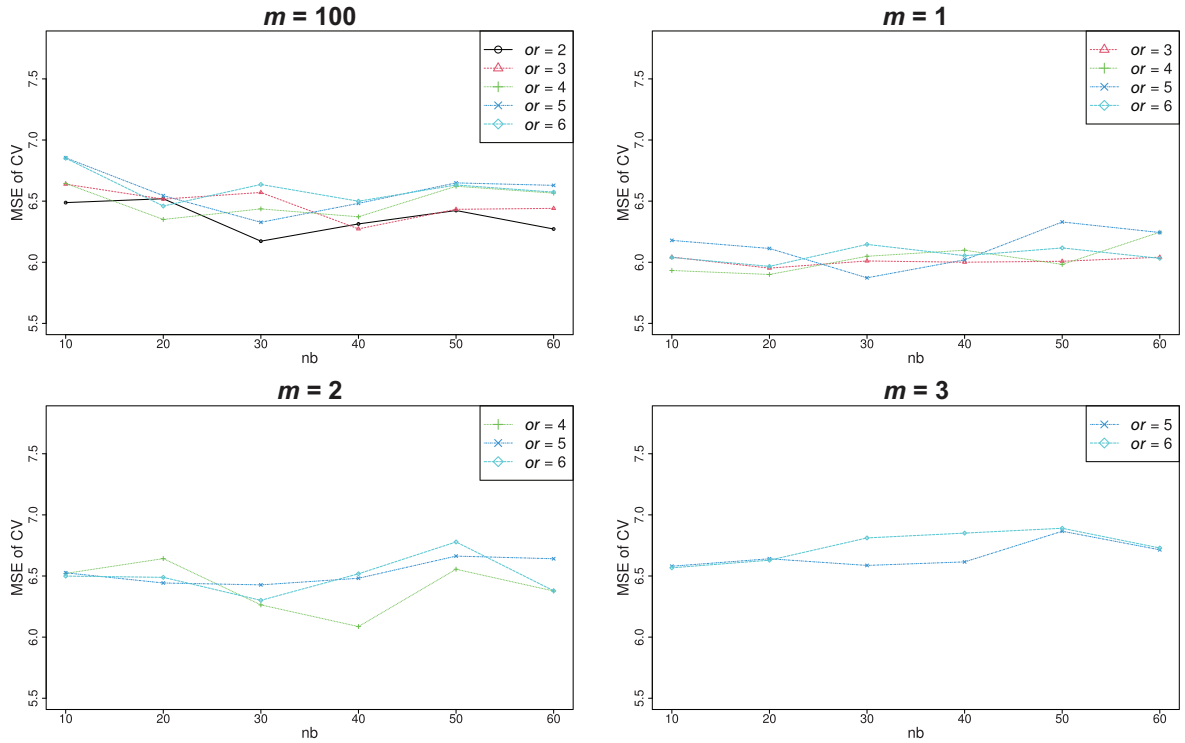

FIGURE 7 Cross-validation MSE at $m = \{0, 1, 2, 3\}$, $or = \{2, 3, 4, 5, 6\}$ and $nb = \{10, 20, 30, 40, 50, 60\}$ in the real-life case study

fusion additive manufacturing for metals, we show the practical advantages of the proposed approach with respect to the methods commonly used in the manufacturing field based on the analysis of scalar features (e.g., sample mean, variance, percentiles). The developed regression method is shown to well predict the $\sigma - \epsilon$ curve (i.e. the response function) given the particle size distribution of the raw powdered material, which can be affordably obtained before starting a new AM production. By means of a Monte Carlo simulation study, the proposed approach showed to outperform, in terms of prediction mean squared error, non-functional competitors from the existing literature. Irrespective of the powder reuse times, the mechanical properties of the produced parts exhibited a small variability, with a potential acceptable decay up to the last printing run. The possibility of reusing metal powder, at least for a certain number of times, allows companies to minimize the material waste, to save money and to reduce the overall environmental impact.

A step forward could refer to the possibility of adding other types of materials and inspection tests, as well as other covariates, not necessarily functional, which may be able to describe the powder properties (e.g., chemical composition, moisture content or tap and apparent density) to increase the predictive power of the method. Indeed, other materials, not limited to metals, or other tests can return more complex curves, for which the functional approach may result indispensable.

ACKNOWLEDGEMENTS

The authors are deeply grateful to the editor and the anonymous referees for their thorough and insightful reviews, which led to significant improvements of the article.

The authors are grateful to MBDA company and in particular to the Fusaro site staff of both IVL and RADOME departments.

The experimental activities have been co-funded by Contratto di Programma Regionale per lo Sviluppo Innovativo delle Filere Manifatturiere Strategiche in Campania POR Campania FESR 2007-2013 Obiettivo Operativo 2.2 (Filiera Aerospazio, Iniziativa Wisch, Work Into Shaping Campania's Home) and were performed in collaboration with T2STAR (Tecnologie dei Sistemi per la Sicurezza Territoriale e AeRea) Consortium led by MBDA. TECNEVA (TECNologie EVolutive per sistemi Avionici) Consortium was the proponent subject, 3F&EDIN S.p.A. the lead partner.

ORCID

Biagio Palumbo https://orcid.org/0000-0003-1036-8127

Fabio Centofanti https://orcid.org/0000-0002-5007-525X
Francesco Del Re https://orcid.org/0000-0001-8499-3422

REFERENCES

1. Reis MS, Rendall R, Palumbo B, et al. Predicting ships' CO_2 emissions using feature-oriented methods. *Appl Stochastic Models Bus Ind*. 2020;36(1):110–123.
2. Ramsay J, Silverman B. *Functional Data Analysis*. Springer Series in Statistics. Berlin: Springer; 2005.
3. Horváth L, Kokoszka P. *Inference for Functional Data with Applications*. Springer Series in Statistics, vol. 200. Berlin: Springer Science & Business Media; 2012.
4. Ferraty F, Vieu P. *Nonparametric Functional Data Analysis: Theory and Practice*. Berlin: Springer Science & Business Media; 2006.
5. Kokoszka P, Reimherr M. *Introduction to Functional Data Analysis*. Boca Raton, FL: CRC Press; 2017.
6. Cuevas A. A partial overview of the theory of statistics with functional data. *J Stat Plan Inference*. 2014;147:1–23.
7. Morris JS. Functional regression. *Annu Rev Stat Its Appl*. 2015;2:321–359.
8. Ramsay JO, Dalzell C. Some tools for functional data analysis. *J R Stat Soc. Series B (Method)*. 1991;53(3):539–572.
9. Besse PC, Cardot H. Approximation spline de la prévision d'un processus fonctionnel autorégressif d'ordre 1. *Can J Stat*. 1996;24(4):467–487.
10. Yao F, Müller HG, Wang JL. Functional linear regression analysis for longitudinal data. *Ann Stat*. 2005;33(6):2873–2903.
11. Yao F, Müller HG, Wang JL. Functional data analysis for sparse longitudinal data. *J Am Statist Assoc*. 2005;100(470):577–590.
12. Ivanescu AE, Staicu AM, Scheipl F, Greven S. Penalized function-on-function regression. *Comput Stat*. 2015;30(2):539–568.
13. Goldsmith J, Bobb J, Crainiceanu CM, et al. Penalized functional regression. *J Comput Graph Stat*. 2011;20(4):830–851.
14. Luo R, Qi X. Function-on-function linear regression by signal compression. *J Am Stat Assoc*. 2017;112(518):690–705.
15. Luo R, Qi X. Interaction model and model selection for function-on-function regression. *J Comput Graph Stat*. 2019:1–14.
16. Centofanti F, Lepore A, Menafoglio A, et al. Functional regression control chart. *Technometrics*. 2020:1–14.
17. Centofanti F, Fontana M, Lepore A, Vantini S. Smooth lasso estimator for the function-on-function linear regression model. *arXiv preprint. arXiv:2007.00529*; 2020.
18. Khajavi SH, Deng G, Holmström J, et al. Selective laser melting raw material commoditization: impact on comparative competitiveness of additive manufacturing. *Int J Prod Res*. 2018;56(14):1–23.
19. Mani M, Lane BM, Donmez MA, et al. A review on measurement science needs for real-time control of additive manufacturing metal powder bed fusion processes. *Int J Prod Res*. 2017;55(5):1400–1418.
20. Slotwinski JA, Garboczi EJ. Metrology needs for metal additive manufacturing powders. *JOM*. 2015;67(3):538–543.
21. Hague R, Mansour S, Saleh N. Material and design considerations for rapid manufacturing. *Int J Prod Res*. 2004;42(22):4691–4708.
22. Gardan J. Additive manufacturing technologies: state of the art and trends. *Int J Prod Res*. 2016;54(10):3118–3132.
23. Petrovic V, Vicente Haro Gonzalez J, Jordá Ferrando O, et al. Additive layered manufacturing: sectors of industrial application shown through case studies. *Int J Prod Res*. 2011;49(4):1061–1079.
24. Gu D, Meiners W, Wissenbach K, Poprawe R. Laser additive manufacturing of metallic components: materials, processes and mechanisms. *Int Mater Rev*. 2012;57(3):133–164.
25. Barclift M, Joshi S, Simpson T, Dickman C. Cost modeling and depreciation for reused powder feedstock in powder bed fusion additive manufacturing. In *Proceedings of the 27th Annual International Solid Freeform Fabrication Symposium—An Additive Manufacturing Conference. Austin, TX*; 2016.
26. Cordova L, Campos M, Tinga T. Revealing the effects of powder reuse for selective laser melting by powder characterization. *JOM*. 2019;71(3):1062–1072.
27. Jacob G, Jacob G, Brown CU, et al. *Effects of powder recycling on stainless steel powder and built material properties in metal powder bed fusion processes*. Gaithersburg, MD: US Department of Commerce, National Institute of Standards and Technology; 2017.
28. Herzog D, Seyda V, Wycisk E, Emmelmann C. Additive manufacturing of metals. *Acta Mater*. 2016;117:371–392.
29. Khairallah SA, Anderson AT, Rubenchik A, King WE. Laser powder-bed fusion additive manufacturing: physics of complex melt flow and formation mechanisms of pores, spatter, and denudation zones. *Acta Mater*. 2016;108(15):36–45.
30. Sutton AT, Kriewall CS, Leu MC, Newkirk JW. Powder characterisation techniques and effects of powder characteristics on part properties in powder-bed fusion processes. *Virtual Phys Prototyping*. 2017;12(1):3–29.
31. Ramberg W, Osgood WR. *Description of Stress-Strain Curves by Three Parameters*. Washington, DC: National Advisory Committee for Aeronautics; 1943.
32. Davis JR. *Metals Handbook, Desk Edition 2nd Edition I*, vol. 464. Materials Park, OH: ASM International; 2003.
33. Del Re F, Contaldi V, Astarita A, et al. Statistical approach for assessing the effect of powder reuse on the final quality of alsi10mg parts produced by laser powder bed fusion additive manufacturing. *Int J Adv Manuf Technol*. 2018;97(5–8):1–10.
34. Efron B, et al. Bootstrap methods: another look at the jackknife. *Ann Stat*. 1979;7(1):1–26.
35. Davison AC, Hinkley DV. *Bootstrap Methods and Their Application*, vol. 1. Cambridge, UK: Cambridge University Press; 1997.
36. De Castro BF, Guillas S, Manteiga WG. Functional samples and bootstrap for predicting sulfur dioxide levels. *Technometrics*. 2005;47(2):212–222.
37. Cuevas A, Febrero M, Fraiman R. On the use of the bootstrap for estimating functions with functional data. *Comput Stat Data Anal*. 2006;51(2):1063–1074.
38. González-Manteiga W, Martínez-Calvo A. Bootstrap in functional linear regression. *J Stat Plan Inference*. 2011;141(1):453–461.

39. R Core Team. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing; 2019.
40. Cardot H, Ferraty F, Sarda P. Spline estimators for the functional linear model. *Stat Sin*. 2003;13(3):571–591.
41. Hastie T, Tibshirani R, Friedman J. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. New York, NY, USA: Springer series in statistics; 2009.
42. De Boor C, Mathématicien EU, et al. *A Practical Guide to Splines*. Applied Mathematical Sciences, vol. 27. New York, NY: Springer-Verlag; 1978.
43. Efron B, Tibshirani R. Improvements on cross-validation: the 632+ bootstrap method. *J Am Stat Assoc*. 1997;92(438):548–560.
44. Efron B, Tibshirani R. Bootstrap methods for standard errors, confidence intervals, and other measures of statistical accuracy. *Stat Sc*. 1986;1(1):54–75.
45. Jolliffe I. *Principal Component Analysis*. Berlin, Germany: Springer; 2011.
46. Martens H, Naes T. *Multivariate Calibration*. New York, NY: John Wiley & Sons; 1992.
47. Naes T, Martens H. Comparison of prediction methods for multicollinear data. *Commun Stat–Simul Comput*. 1985;14(3):545–576.
48. ASTM B822–17. *Standard test method for particle size distribution of metal powders and related compounds by light scattering*. West Conshohocken, PA: American Society for Testing and Materials; 2017.
49. ISO 6892:2016. *Metallic Materials - Tensile Testing. Part 1: Method of Test at Room Temperature*. ISO, International Organization for Standardization; 2016.
50. ASTM E8/E8M–16a. *Standard Test Methods for Tension Testing of Metallic Materials*. West Conshohocken, PA: American Society for Testing and Materials; 2016.
51. Silverman BW. *Density Estimation for Statistics and Data Analysis*. London: Routledge; 2018.
52. Charpentier A, Flachaire E. Log-transform kernel density estimation of income distribution. *L'Actualité écon*. 2015;91(1-2):141–159.
53. Sames WJ, List FA, Pannala S, et al. The metallurgy and processing science of metal additive manufacturing. *Int Mater Rev*. 2016;61(5):315–360. https://doi.org/10.1080/09506608.2015.1116649.
54. Gabrys R, Kokoszka P. Portmanteau test of independence for functional observations. *J Am Stat Assoc*. 2007;102(480):1338–1348.

AUTHOR BIOGRAPHIES

Biagio Palumbo is an associate professor of "statistics for experimental and technological research" at the Department of Industrial Engineering of the University of Naples Federico II. His major research interests include reliability, design and analysis of experiments, statistical methods for process monitoring and optimization and data science for technology.

Fabio Centofanti is a PhD student at the Department of Industrial Engineering of the University of Naples Federico II. His research interests include functional data analysis and statistical process monitoring.

Francesco Del Re is research assistant at the Department of Industrial Engineering of the University of Naples Federico II. His major research interests refer to the use of statistical methods for the characterization and the optimization of complex manufacturing processes.

How to cite this article: Palumbo B, Centofanti F, Del Re F. Function-on-function regression for assessing production quality in industrial manufacturing. *Qual Reliab Engng Int*. 2020;36:2738–2753. https://doi.org/10.1002/qre.2786

APPENDIX A: DERIVATION OF THE ESTIMATOR $\hat{B}$

The penalty $||\mathcal{L}_s^{m_s}(\boldsymbol{\psi}^{sT}\boldsymbol{B}\boldsymbol{\psi}^t)||^2$ on the right-hand side of Equation (5) may be written as

$$\begin{aligned}||\mathcal{L}_s^{m_s}\left(\boldsymbol{\psi}^{sT}\boldsymbol{B}\boldsymbol{\psi}^t\right)||^2 &= \int_{\mathcal{S}}\int_{\mathcal{T}} \mathcal{L}_s^{m_s}\left[\boldsymbol{\psi}^s(s)^T\boldsymbol{B}\boldsymbol{\psi}^t(t)\right]\mathcal{L}_s^{m_s}\left[\boldsymbol{\psi}^s(s)^T\boldsymbol{B}\boldsymbol{\psi}^t(t)\right]dsdt\\ &= \int_{\mathcal{S}}\int_{\mathcal{T}} \mathcal{L}_s^{m_s}\left[\boldsymbol{\psi}^s(s)^T\right]\boldsymbol{B}\boldsymbol{\psi}^t(t)\boldsymbol{\psi}^t(t)^T\boldsymbol{B}^T\mathcal{L}_s^{m_s}[\boldsymbol{\psi}^s(s)]dsdt\\ &= \mathrm{Tr}\left[\left(\int_{\mathcal{S}}\mathcal{L}_s^{m_s}[\boldsymbol{\psi}^s(s)]\mathcal{L}_s^{m_s}\left[\boldsymbol{\psi}^s(s)^T\right]ds\right)\boldsymbol{B}\left(\int_{\mathcal{T}}\boldsymbol{\psi}^t(t)\boldsymbol{\psi}^t(t)^Tdt\right)\boldsymbol{B}^T\right]\\ &= \mathrm{Tr}\left[\boldsymbol{B}^T\boldsymbol{R}_X\boldsymbol{B}\boldsymbol{W}_Y\right],\end{aligned} \tag{A1}$$

where $\boldsymbol{R}_X$ and $\boldsymbol{W}_Y$ are as in Equation (6) and $\mathrm{Tr}(\boldsymbol{A})$ is the trace of a generic matrix $\boldsymbol{A}$. Analogously to Equation (11), the penalty: $||\mathcal{L}_t^{m_t}(\boldsymbol{\psi}^{sT}\boldsymbol{B}\boldsymbol{\psi}^t)||^2$ on the right-hand side of Equation (5) may be written as

$$||\mathcal{L}_t^{m_t}\beta||^2 = \mathrm{Tr}\left[\boldsymbol{B}^T\boldsymbol{W}_X\boldsymbol{B}\boldsymbol{R}_Y\right], \tag{A2}$$

where $\boldsymbol{W}_X$ and $\boldsymbol{R}_Y$ are as in Equation (6). Moreover,

$$\begin{aligned}||Y_i - \boldsymbol{X}_i^T\boldsymbol{B}\boldsymbol{\psi}^t||^2 &= \int_{\mathcal{T}}\left[Y_i(t) - \boldsymbol{X}_i^T\boldsymbol{B}\boldsymbol{\psi}^t(t)\right]^2dt\\ &= \int_{\mathcal{T}}\left[Y_i(t)^2 - 2Y_i(t)\boldsymbol{\psi}^t(t)^T\boldsymbol{B}^T\boldsymbol{X}_i + \boldsymbol{X}_i^T\boldsymbol{B}\boldsymbol{\psi}^t(t)\boldsymbol{\psi}^t(t)^T\boldsymbol{B}^T\boldsymbol{X}_i\right]dt\\ &= \int_{\mathcal{T}}Y_i(t)^2dt - 2\left(\int_{\mathcal{T}}Y_i(t)\boldsymbol{\psi}^t(t)^Tdt\right)\boldsymbol{B}^T\boldsymbol{X}_i + \boldsymbol{X}_i^T\boldsymbol{B}\left(\int_{\mathcal{T}}\boldsymbol{\psi}^t(t)\boldsymbol{\psi}^t(t)^Tdt\right)\boldsymbol{B}^T\boldsymbol{X}_i\\ &= \int_{\mathcal{T}}Y_i(t)^2dt - 2\boldsymbol{Y}_i^T\boldsymbol{B}^T\boldsymbol{X}_i + \boldsymbol{X}_i^T\boldsymbol{B}\boldsymbol{W}_Y\boldsymbol{B}^T\boldsymbol{X}_i,\end{aligned} \tag{A3}$$

and, thus:

$$\sum_{i=1}^n||Y_i - \boldsymbol{X}_i^T\boldsymbol{B}\boldsymbol{\psi}^t||^2 = \sum_{i=1}^n\int_{\mathcal{T}}Y_i(t)^2dt - 2\,\mathrm{Tr}\left[\boldsymbol{X}\boldsymbol{B}\boldsymbol{Y}^T\right] + \mathrm{Tr}\left[\boldsymbol{X}^T\boldsymbol{X}\boldsymbol{B}\boldsymbol{W}_Y\boldsymbol{B}^T\right], \tag{A4}$$

where $\boldsymbol{X}$ and $\boldsymbol{Y}$ are as in Equation (6).

Then, the optimization problem in Equation (5) becomes

$$\begin{aligned}\widehat{\mathbf{B}} = \underset{\mathbf{B}\in\mathbb{R}^{L\times M}}{\mathrm{argmin}}\Bigg\{&\sum_{i=1}^n\int_{\mathcal{T}}Y_i(t)^2dt - 2\,\mathrm{Tr}\left[\mathbf{X}\mathbf{B}\mathbf{Y}^T\right] + \mathrm{Tr}\left[\mathbf{X}^T\mathbf{X}\mathbf{B}\mathbf{W}_Y\mathbf{B}^T\right]\\ &+\lambda_s\,\mathrm{Tr}\left[\mathbf{R}_X\mathbf{B}\mathbf{W}_Y\mathbf{B}^T\right] + \lambda_t\,\mathrm{Tr}\left[\mathbf{W}_X\mathbf{B}\mathbf{R}_Y\mathbf{B}^T\right]\Bigg\}.\end{aligned} \tag{A5}$$

By taking the derivative of the right-hand side of Equation (14) with respect to $\boldsymbol{B}$ and setting it equal to zero, we obtain

$$\boldsymbol{X}^T\boldsymbol{X}\boldsymbol{B}\boldsymbol{W}_Y + \lambda_s\boldsymbol{R}_X\boldsymbol{B}\boldsymbol{W}_Y + \lambda_t\boldsymbol{W}_X\boldsymbol{B}\boldsymbol{R}_Y = \boldsymbol{X}^T\boldsymbol{Y}, \tag{A6}$$

and, thus, by applying the vec operator on both sides of Equation (16), we obtain:

$$\mathrm{vec}\left(\boldsymbol{X}^T\boldsymbol{X}\boldsymbol{B}\boldsymbol{W}_Y\right) + \lambda_s\,\mathrm{vec}\left(\boldsymbol{R}_X\boldsymbol{B}\boldsymbol{W}_Y\right) + \lambda_t\,\mathrm{vec}\left(\boldsymbol{W}_X\boldsymbol{B}\boldsymbol{R}_Y\right) = \mathrm{vec}\left(\boldsymbol{X}^T\boldsymbol{Y}\right). \tag{A7}$$

By using the the fact that, for generic matrices $\boldsymbol{A}$, $\boldsymbol{B}$ and $\boldsymbol{C}$ of appropriate dimensions, $\mathrm{vec}(\boldsymbol{ABC}) = (\boldsymbol{C}^T \otimes \boldsymbol{A})\,\mathrm{vec}(\boldsymbol{B})$ and $\mathrm{vec}(\boldsymbol{AB}) = (\boldsymbol{I} \otimes \boldsymbol{A})\,\mathrm{vec}(\boldsymbol{B})$, where $\boldsymbol{I}$ is an identity matrix of appropriate dimensions, the estimator of $\boldsymbol{B}$ is:

$$\mathrm{vec}\left(\hat{\boldsymbol{B}}\right) = \left[\boldsymbol{W}_Y \otimes \boldsymbol{X}^T\boldsymbol{X} + \lambda_s \boldsymbol{W}_Y \otimes \boldsymbol{R}_X + \lambda_t \boldsymbol{R}_Y \otimes \boldsymbol{W}_X\right]^{-1}\left(\boldsymbol{I} \otimes \boldsymbol{X}^T\right)\mathrm{vec}\left(\boldsymbol{Y}\right). \tag{A8}$$

APPENDIX B: GRAPHICAL RESIDUALS ANALYSIS FOR THE REAL-LIFE CASE STUDY

With respect to the real-life case study of Section 4, the residuals from the FoF linear regression model are calculated for each of the six cylindrical tensile specimens and reported in Figure B.1A. Alongside, in Figure B.1B their L^2 norms are plotted as a function of the powder reuse. Both plots do not exhibit any strong unusual pattern and confirm the result of the portmanteau independence test of Section 4, which does not reject the null hypothesis of i.i.d. errors.

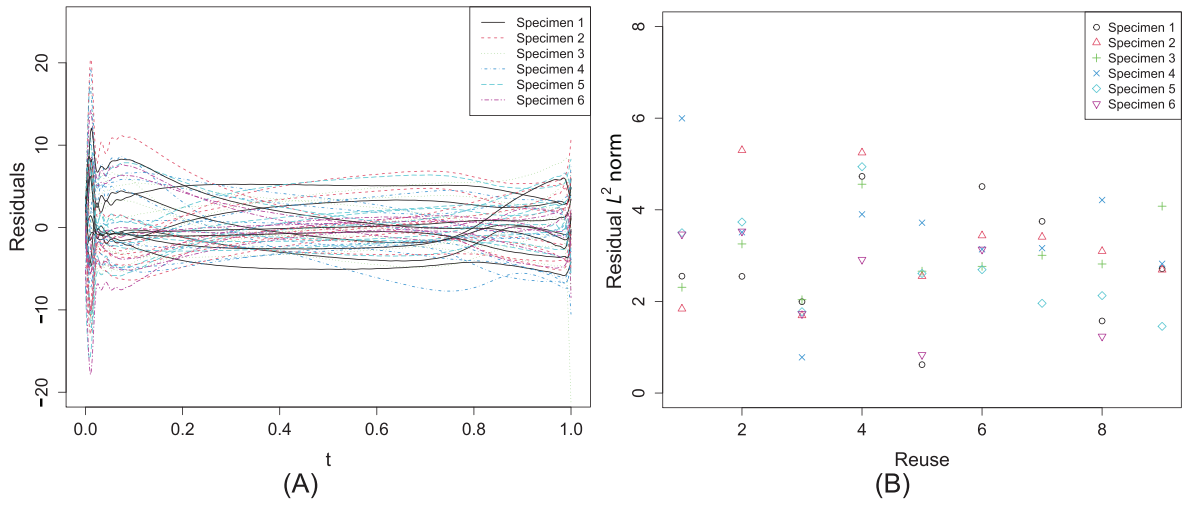

FIGURE B.1 (A) Residuals from the FoF linear regression model in the real-life case study for each of the six cylindrical tensile specimens and (B) the corresponding L^2 norm as a function of the powder reuse