Advanced Modeling and Simulation in Engineering Sciences

RESEARCH ARTICLE

Open Access

Evaluation of POD based surrogate models of fields resulting from nonlinear FEM simulations

Boukje M. de Gooijer*[iD], Jos Havinga[iD], Hubert J. M. Geijselaers[iD] and Anton H. van den Boogaard[iD]

*Correspondence:
b.m.degooijer@utwente.nl
Chair of Nonlinear Solid
Mechanics, University of Twente,
Drienerlolaan 5, 7522 NB
Enschede, The Netherlands

Abstract

Surrogate modelling is a powerful tool to replace computationally expensive nonlinear numerical simulations, with fast representations thereof, for inverse analysis, model-based control or optimization. For some problems, it is required that the surrogate model describes a complete output field. To construct such surrogate models, proper orthogonal decomposition (POD) can be used to reduce the dimensionality of the output data. The accuracy of the surrogate models strongly depends on the (pre)processing actions that are used to prepare the data for the dimensionality reduction. In this work, POD-based surrogate models with Radial Basis Function interpolation are used to model high-dimensional FE data fields. The effect of (pre)processing methods on the accuracy of the result field is systematically investigated. Different existing methods for surrogate model construction are compared with a novel method. Special attention is given to data fields consisting of several physical meanings, e.g. displacement, strain and stress. A distinction is made between the errors due to truncation and due to interpolation of the data. It is found that scaling the data per physical part substantially increases the accuracy of the surrogate model.

Keywords: Metamodel, Multiphysical field, Preprocessing, Proper Orthogonal Decomposition, Radial Basis Function

Introduction

Computational Fluid Dynamics (CFD) and Finite Element (FE) analyses are powerful tools to solve engineering problems for which no analytical solution can be found. However, for nonlinear problems with many degrees of freedom, these analyses can be computationally costly. Consequently CFD and FE models are impractical for direct use in e.g. inverse analyses, (robust) optimization or control algorithms. For these purposes surrogate models can be constructed [1,2]. Surrogate models are cheap and fast representations that mimic the output of expensive models. Surrogate models are constructed from data sets that contain the full input and output fields for multiple evaluations of the (CFD or FE) model. These data sets are considered to be the observations of a black-box function. Several

mathematical methods exist to approximate the black-box function using data [3]. In this work, specifically the use of non-intrusive Radial Basis Function interpolation models is investigated, for the case that no gradient information is available. With non-intrusive it is meant that no access to the source code of the solver is required.

It is often sufficient to construct a scalar surrogate model, also referred to as metamodel, to optimize a single objective function. For example in the work of Wiebenga et al. [4] an industrial V-bending process is optimized. In the V-bending process the flange shape is defined by the main angle, a scalar, that is optimized towards 90°. To perform such task, a continuous description of the interpolant on the input space is needed. However, sometimes a complete output field needs to be described by the surrogate model. For example in the work of Hamdaoui et al. [5] the displacement field of the stamping of an axisymmetric cup from a blank made of high strength steels needs to be described in order to model the major and minor strain in a Forming Limit Diagram. In the work of Iuliano and Quagliarella [6] an airfoil in different transonic conditions is optimized. The lift, drag and pitching moment coefficient of the airfoil depend on the flow field evaluated using CFD. When such output field is large, it becomes inefficient to use standard scalar techniques for constructing the surrogate model. Furthermore, large data sets may lead to storage issues as well. To overcome these issues, reduction techniques can be applied to construct full field surrogate models. A commonly used method is Proper Orthogonal Decomposition. Proper Orthogonal Decomposition (POD) is an umbrella term that includes Principal Component Analysis (PCA), Karhunen-Loeve Expansion (KLE) and Singular Value Decomposition (SVD) [7,8]. The goal of the decomposition is to obtain a low-dimensional representation of the output field that contains as much as necessary of the variation that can be observed in the data set.

Proper Generalized Decomposition (PGD) can also be used to find such low-dimensional representation of the output field. In PGD the reduced basis and the solution of the problem are solved simultaneously making use of separation of variables [9]. The computational time of expensive numerical models can therefore be drastically reduced with PGD through the reduction of the system of equations. Proper Orthogonal Decomposition can be used in a similar way for reducing the system of equations, although the reduced basis is determined a posteriori, based on the history of previous increments in the calculations of the model. Such applications of POD are intrusive, as they require to be implemented into the nonlinear solver [8,10–13]. In contrast, non-intrusive POD is used to construct a reduced basis of a model based on a set of offline observations. This is often used for constructing a cheap replica of an expensive numerical model. The focus in this paper is on the use of non-intrusive POD with RBF interpolation, as it can be used offline without access to the source code.

The output fields are projected into the reduced basis to obtain the *amplitudes* that will be interpolated to construct a continuous metamodel. These projections can also be referred to as *(POD-)coefficients* [14], *coordinates* [15], *modal coordinates* [16], *factor scores* [17] or *weights*. To construct the projection-based reduced-order surrogate model, the amplitudes can be interpolated using many different metamodelling techniques, such as Kriging, Response Surface Method and Radial Basis Function interpolation. For an extensive overview of metamodelling techniques see [1,2,18,19]. In this work, Radial Basis Function (RBF) interpolation is used to interpolate the amplitudes.

Applications of surrogate models based on a combination of POD and RBF interpolation are found in many engineering fields. Among many other examples, nonlinear FE analyses are replaced by a POD+RBF-based surrogate model in a geotechnical application in the work of Khaledi et al. [20]. In their work a POD+RBF-based surrogate model is used to identify the material parameters of the soil in the simulation of tunnel excavations. Bocciarelli et al. [21] identify elasticity and plasticity material parameters of two aluminium alloys using an indentation test. In the work of Havinga et al. [22], a POD+RBF-based surrogate model of the force-displacement curve from a sheet bending process is used to perform real-time Bayesian estimation of process variations. Dang et al. [23] use a POD+RBF-based surrogate model to reduce the springback in U-shape draw bending. CFD analyses are replaced by a POD+RBF-based surrogate model by Kato et al. [24] to optimize a turbine blade.

The general off-line steps to construct a non-intrusive POD-based surrogate model are: obtaining a training set consisting of input and output data, preprocessing the output data, determining the predominant modes of the output data, interpolating the amplitudes of the different modes in the input space, and lastly post-processing the approximation. We will refer to the matrix, that collects the output data per input setting as its columns, as the snapshot matrix. The rows of the snapshot matrix correspond to the different (state) variables that need to be modelled [16,25]. The rows of the snapshot matrix are also referred to as *components* [26], *degrees of freedom* [5,23], *measurable quantities* [21] or *observation points* [27]. When different types of variables are present in the output data, one may use different separate snapshot matrices for each variable or gather them all in a single snapshot matrix. For instance, to model a displacement field, Hamdaoui et al. [5] use separate snapshot matrices for the displacements in *x*- and *y*-direction. While in the work of Buljak [28] the displacements in *x*- and *y*-direction are gathered in one single snapshot matrix. When different physical quantities need to be modelled one can also gather all data in one big snapshot matrix. However, in the surrogate modelling of FE analyses it is more common to use separate snapshot matrices for each physical quantity. For example separate snapshot matrices are used in Misiun et al. [29] to model stress components, plastic strain, a damage indicator and temperature; in Steffes-lai [30] to model the equivalent plastic strain, thickness and a damage criterion; and in Buljak [31] to model the loading & unloading curves, nodal displacements, equivalent stresses and principal stress components.

When such physical quantities are gathered in the same snapshot matrix their values may very much differ in magnitude, e.g. strains in the order of 10^{-1} and stresses in the order of 10^6 Pa (10^1–10^2 MPa). When a surrogate model of a CFD analysis is constructed the issue that some data is of different orders of magnitude compared to others is often solved by scaling the different physical parts. For example in the work of Zimmermann and Görtz [32] and Guéntot et al. [14] different scaling methods are described. In the latter a comparison is made between scaling the snapshot matrix per variable and per physical part with the mean, range and standard deviation. They found that there is not one superior scaling method, nevertheless the average errors are lower when scaling is applied to the snapshot matrix.

In this work, the effect of different approaches for preprocessing high-dimensional multivariate output data on the accuracy of POD surrogate models, that are constructed based on these data, are studied. In the next section different approaches in surrogate model

construction are presented. Two of these approaches are based on different preprocessing methods of the snapshot matrix and two approaches are based on decomposition of different parts of the snapshot matrix. The focus of this work is on how the snapshot matrix should be handled in the context of POD-RBF interpolation. The systematic comparison of the surrogate model results will contribute to a deeper understanding of the different approaches. One of the approaches is novel, while the others have been well documented in literature. The novel approach is relevant for the case that multiple physical quantities are simultaneously included in the data set, for example when the data set contains both stresses and strains. When a certain physical quantity is regarded as the main modelling objective, a reduced order model of the other quantities can be computed based on the covariance between the quantities, without requiring additional expensive computations for model reduction or fitting the interpolation models. These cheaper computations come at the cost of reduced accuracy, but may be appropriate for many applications, such as for visualization. In the section 'Demonstrator processes' two metal-forming processes are introduced for which surrogate models of the displacements, equivalent plastic strains and stress field are desired. The bases that are used for the different surrogate models are analyzed in the section 'Comparing different bases'. The different surrogate models are compared based on a validation set as presented in the section 'Validation set results of the demonstrator process'. The methods presented in this paper are generally applicable and are not limited to the specific case that is used as a demonstrator.

Constructing different surrogate models

The general steps to construct a POD-based surrogate model are: obtaining a training set, preprocessing the snapshot matrix, reducing the output, interpolating the amplitudes and lastly post-processing the approximation. All steps are depicted in Fig. 1, and will be discussed in more detail in this section.

Following the steps as presented in Fig. 1 and elaborating in the following sections, four different surrogate models will be constructed. In order to illustrate the theory with an example, the following output data field is presented. This example is used throughout the discussion, although the approaches are generic and can be applied to any high-dimensional multivariate output field. The different approaches discussed in this work can be used as guidelines to construct a surrogate model of other cases. It is expected that the suitability of the methods discussed in this work and of other methods that can perform the same task, may depend on the specific applications. The example output field of a single FE simulation consists of M variables that can be stored in the following $M \times 1$ result vector that can be partitioned according to 3 different physical fields:

$$\mathbf{y} = \begin{Bmatrix} \{\mathbf{y}_u\} \\ \{\mathbf{y}_\varepsilon\} \\ \{\mathbf{y}_\sigma\} \end{Bmatrix} \tag{1}$$

in which $\mathbf{y}_u$ contains the M_u displacements, $\mathbf{y}_\varepsilon$ contains the M_ε equivalent plastic strains, and $\mathbf{y}_\sigma$ contains the M_σ stress components.

The approaches described in the following sections can be applied to any other result field that is described by different physical fields. Different steps of the preprocessing and decomposition procedure may be either omitted or performed in several ways, lead-

ing to the following four approaches for construction of a POD-based surrogate model considered in this paper:

1. Only centering data by subtracting the mean value from the data (section 'Zero centering').
2. Centering, scaling and decomposing all data at once (section 'Scaling each physical part').
3. Centering data and decomposing each physical part independently (section 'Separate reduction of each physical part').
4. Centering data and assemble basis vectors from one physical quantity (section 'Assembly from one physical part').

The approaches 1 and 3 are commonly applied in surrogate modelling of FE analyses [5,28–31], while approach 2 is often applied in surrogate modelling of CFD analyses [14,32]. Scaling the data means that each variable, physical part or the total snapshot matrix is multiplied with a constant. Approach 4 is novel, in this approach one single physical part is used to assemble the basis vectors describing the total field. In the following sections, the steps of the surrogate model construction and the differences between the four approaches will be explained in detail.

Obtaining a training set

The input parameter space $\mathbf{x}$ is a N_{dim}-dimensional space that will be sampled with N_{exp} sample points. The sample points $\mathbf{x}_i$ in the initial sample set $\mathbf{X}$ can be generated using different sampling strategies. On each sample point a simulation is performed of which the data can be stored in the result vector $\mathbf{y}_i$. The snapshot matrix $\mathbf{Y}$ is obtained by collecting

Fig. 1 Construction of a POD-based surrogate model
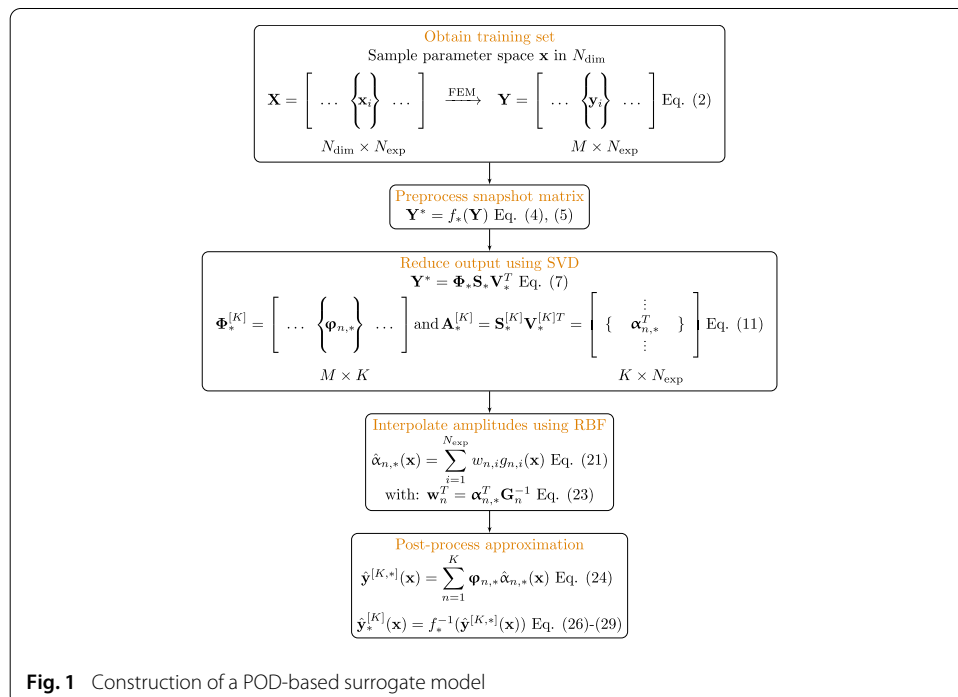
the simulation results as its columns:

$$\mathbf{X} = \left[\cdots \left\{ \mathbf{x}_i \right\} \cdots \right] \xrightarrow{\text{FEM}} \mathbf{Y} = \left[\cdots \left\{ \mathbf{y}_i \right\} \cdots \right] \tag{2}$$
$$\underbrace{}_{N_{\text{dim}} \times N_{\text{exp}}} \qquad\qquad \underbrace{}_{M \times N_{\text{exp}}}$$

Many different sampling strategies can be found in literature. In the construction of POD-based surrogate models it is common practice to combine different sampling strategies to obtain the initial sample set, e.g. in the work of Havinga [22] a full factorial design is combined with a space-filling Latin hypercube sampling (LHS) and in the work of Steffes-lai [30] a starpoint design is extended with more sample points using a custom infill strategy. A full factorial design consists of $2^{N_{\text{dim}}}$ sample points that are different combinations of the input parameters at their minimum or maximum value. The full factorial design is used to properly cover the boundaries of the parameter space [22]. A star point design consists of a center point, that is the nominal setting, and $2N_{\text{dim}}$ star points. The star points are constructed by setting each input parameter one-at-a-time to its minimum and maximum value, while other input parameters are kept at their nominal value. The star point design can be used for sensitivity analysis [30]. To generate a LHS the parameter space is divided into $N_{\text{exp}}^{N_{\text{dim}}}$ N_{dim}-dimensional hypercubes. Sample points are placed within the hypercubes under the condition that there are no sample points in the hypercubes perpendicular to a hypercube that is containing a sample point. If this is done such that the LHS is spacefilling, this generally leads to a uniform sampling over the whole parameter space.

Preprocessing the snapshot matrix
The snapshot matrix obtained based on the initial sample set as described in the previous section can be used for decomposition without any further modifications. For instance, no modifications to the snapshot matrix are made in the work of Bocciarelli et al. [21]. However, when a suitable preprocessing method is applied to the snapshot matrix, the information content of the first basis vectors can be increased. We define the preprocessing of the snapshot matrix as:

$$\mathbf{Y}^* = f_*(\mathbf{Y}) \tag{3}$$

In which $f_*(\cdot)$ is the preprocessing function that transforms the snapshot matrix into the preprocessed snapshot matrix $\mathbf{Y}^*$. The superscript * and subscript $_*$ denote the applied preprocessing method, e.g. '0' for zero centering and 'scaled' for scaling per physical part.

Zero centering
In the construction of non-intrusive POD-based surrogate models it is common practice to subtract the mean result from the snapshot matrix before decomposing [33]. Subtracting the mean centers the data around zero, the resulting matrix is therefore also referred to as the deviation matrix [14]. Subtracting the mean increases the information captured in the first basis vectors. When the mean is not subtracted, the first basis vector will dominate the decomposition and will point towards the mean [16,34]. As basis vectors will be orthogonal, the subsequent basis vectors will be constrained by the direction of the first basis vector, resulting in a suboptimal decomposition of the data set. An extensive

comparison between non-centred and centred PCA is given by Cadima and Jolliffe [35]. Although subtracting the mean has been shown beneficial in the construction of non-intrusive POD-based surrogate models [36], there are also applications known in which subtracting the mean is not compatible, for example in incremental SVD [37].

To obtain the zero centred snapshot matrix the mean over each row, and hence each variable, in the original snapshot matrix is subtracted.

$$\mathbf{Y}^0 = f_0(\mathbf{Y}) = \mathbf{Y} - \bar{\mathbf{y}}\mathbf{1}^T \quad \text{where:} \quad \bar{\mathbf{y}} = \frac{1}{N_{\text{exp}}}\mathbf{Y}\mathbf{1} \tag{4}$$

in which $\mathbf{1}$ is a $N_{\text{exp}} \times 1$ vector of ones. Note that the rank of the zero centered snapshot matrix is one lower than the rank of the original matrix.

Scaling the snapshot matrix

Scaling the entries of the zero centered snapshot matrix is a common preprocessing method in Principal Component Analysis. The scaling can be performed to the entire snapshot matrix, to each physical part separately or per variable in each row. The scaling can be calculated in many different ways. Scaling the data includes, but is not limited to, normalising the data.

The most simple scaling method is to divide the entire snapshot matrix by a constant. This is a common preprocessing method in covariance PCA. To perform covariance PCA the zero centered snapshot matrix is scaled with the square root of the number of experiments. When the snapshot matrix is divided by $\sqrt{N_{\text{exp}}}$ or $\sqrt{N_{\text{exp}} - 1}$ the matrix $\mathbf{Y}^T\mathbf{Y}$ will be a covariance matrix [17,38]. Dividing a matrix by a constant does not alter the singular vectors that will be found in the SVD. It will only change the magnitudes of the singular values by the same factor. Dividing by a constant therefore does not alter the basis vectors obtained in a decomposition and for that reason is left out in the remainder of this paper.

The scaling can also be performed per variable, that is each row in the snapshot matrix. For example to perform correlation PCA, all rows are divided by their Euclidean norm [17]. A major advantage of scaling by the norm is that the snapshot matrix is divided by something with the same physical quantity. Hence, the entries in the snapshot matrix become dimensionless. Instead of using the norm for scaling the data one can also use the mean, range or standard deviation. Note that the mean of a zero centered snapshot matrix is $\mathbf{0}$, and hence the mean $\bar{\mathbf{y}}$ as defined in equation (4) should be used for scaling. As pointed out by Skillicorn [33] dividing by the standard deviation will ensure that most of the values in each row will fall in the range -1 to $+1$. It has been found in previous work [36] that this is not beneficial for the quality of the surrogate model because the variation of the variables with relatively small variations will be amplified, and therefore these variables get a larger contribution to the basis vectors. To illustrate that this is not beneficial for the quality of the surrogate model, we refer to the industrial bending problem as described in "Demonstrator processes" section. The displacements in *y*-direction in the clamped part of the sheet metal will be very small or even zero. Therefore, the standard deviation will also be very small or zero. Scaling these displacements with their standard deviation will magnify their importance. Hence, when the snapshot matrix is decomposed the basis vectors will capture behavior (and noise) which in practice is barely present. Scaling each row is therefore also left out of consideration in this paper.

Scaling each physical part

The scaling can also be performed on each physical part in the snapshot matrix. As pointed out by Guéntot et al. [14] for a snapshot matrix with different physical quantities it is better to scale per data type. As different physical quantities are present in the result vector, they can be of different order. For example, in the demonstrator process the stress components in MPa are of order 10^2, while the strain components are of order 10^{-1}. This will cause the decomposition to be mainly determined by the stress terms. It is therefore proposed to scale each part of the snapshot matrix. In [14] this is referred to as physical scaling. The zero centered snapshot matrix in the example of this work with scaling applied to each part will take the form:

$$\mathbf{Y}^{\text{scaled}} = f_{\text{scaled}}(\mathbf{Y}) = \begin{bmatrix} \left[\mathbf{Y}_{\text{u}} - \overline{\mathbf{y}}_{\text{u}}\mathbf{1}^T\right] \cdot s_{\text{u}} \\ \left[\mathbf{Y}_{\varepsilon} - \overline{\mathbf{y}}_{\varepsilon}\mathbf{1}^T\right] \cdot s_{\varepsilon} \\ \left[\mathbf{Y}_{\sigma} - \overline{\mathbf{y}}_{\sigma}\mathbf{1}^T\right] \cdot s_{\sigma} \end{bmatrix} \tag{5}$$

Herein s_{u}, s_{ε} and s_{σ} are the scaling constants for the displacement, equivalent plastic strain and stress respectively. The scaling constants can be for example the mean, standard deviation or range of each physical part [14]. To the best of the authors knowledge, no applications can be found in literature of scaling the snapshot matrix per physical part to model FE simulation data. It is proposed to scale the different parts by their range. To calculate the range of each physical part the minimum over each row and column is subtracted from the maximum over each row and column. For the displacement field the physical scaling constant s_u can be calculated as:

$$s_{\text{u}} = (\max \mathbf{Y}_{\text{u}} - \min \mathbf{Y}_{\text{u}})^{-1} \tag{6}$$

The scaling constants for the other parts can be calculated in an equivalent way.

Decomposing the output

The predominant modes in the preprocessed snapshot matrices will now be determined. Hence, the zero centered snapshot ($\mathbf{Y}^0$, section 'Zero centering') and the snapshot matrix that is scaled per physical part ($\mathbf{Y}^{\text{scaled}}$, section 'Scaling each physical par'), will be decomposed. The snapshot matrix is decomposed into a new basis before reduction. In this work a singular value decomposition (SVD) is used to find the proper orthogonal basis vectors [39]. The SVD of the preprocessed snapshot matrix takes the following form:

$$\mathbf{Y}^* = \mathbf{\Phi}_* \mathbf{D}_* \mathbf{V}_*^T = \begin{bmatrix} \cdots & \left\{ \boldsymbol{\varphi}_{n,*} \right\} & \cdots \end{bmatrix} \begin{bmatrix} \ddots & & 0 \\ & d_{n,*} & \\ 0 & & \ddots \end{bmatrix} \begin{bmatrix} \vdots \\ \{ \mathbf{v}_{n,*}^T \} \\ \vdots \end{bmatrix} \tag{7}$$

$$M \times N_{\text{exp}} \qquad N_{\text{exp}} \times N_{\text{exp}} \quad N_{\text{exp}} \times N_{\text{exp}}$$

The preprocessed snapshot matrix is decomposed into three matrices: $\mathbf{\Phi}_*$ that contains the left singular vectors $\boldsymbol{\varphi}_{n,*}$ as its columns, $\mathbf{D}_*$ that contains the singular values $d_{n,*}$ on its diagonal and $\mathbf{V}_*^T$ that contains the right singular vectors $\mathbf{v}_{n,*}$ as its rows. The subscripts n and $*$ denote the n-th direction in the basis and the applied preprocessing method, respectively. As the singular values are sorted by size from largest to smallest, the most information will be captured by the first singular vectors.

Alternatively the so called *method of snapshots* can be used [16], though the resulting basis will be the same. In the method of snapshots the basis vectors are constructed from

the snapshot matrix by finding the eigenvalues and eigenvectors of the matrix $\mathbf{C}_* = (\mathbf{Y}^*)^T \mathbf{Y}^*$ [28]. The orthonormal basis vectors can be found using:

$$\boldsymbol{\varphi}_{n,*} = \mathbf{Y}^* \cdot \mathbf{v}_{n,*} \cdot \lambda_{n,*}^{-1/2} \tag{8}$$

In which $\mathbf{v}_{n,*}$ and $\lambda_{n,*}$ are the eigenvectors and eigenvalues of the matrix $\mathbf{C}_* = (\mathbf{Y}^*)^T \mathbf{Y}^*$ respectively. By comparing with equation (7), the link between POD and SVD can be deduced. The eigenvectors of the matrix $\mathbf{C}_*$ are the right singular vectors. Hence, the matrix $\mathbf{V}_*$ holds the collection of all N_{exp} eigenvectors $\mathbf{v}_{n,*}$. The left singular vectors are the eigenvectors of $(\mathbf{Y}^*)(\mathbf{Y}^*)^T$ [40]. The singular values are simply the square root of the eigenvalues, hence: $d_{n,*} = \sqrt{\lambda_{n,*}}$ [41].

Note that both the left and right singular vectors are orthonormal, in that case:

$$\mathbf{V}_*^T \mathbf{V}_* = \boldsymbol{\Phi}_*^T \boldsymbol{\Phi}_* = \mathbf{I} \tag{9}$$

Each left singular vector represents a vector in the result space (of length M). The left singular vector matrix can be seen as a new coordinate system in the result space, and can therefore be used as a basis for the data. This new basis can be truncated so that it will contain only the first K basis vectors $\boldsymbol{\varphi}_{n,*}$ with $n = 1...K$. We define the truncated basis with K basis vectors as:

$$\boldsymbol{\Phi}_*^{[K]} = \underbrace{\left[\left\{ \boldsymbol{\varphi}_{1,*} \right\} \cdots \left\{ \boldsymbol{\varphi}_{K,*} \right\} \right]}_{M \times K} \tag{10}$$

In which the superscript $[K]$ denotes the number of basis vectors in the basis. The amplitudes of the result vectors projected onto the basis vectors are found by multiplying the singular values with the right singular vectors.

$$\mathbf{A}_*^{[K]} = \mathbf{D}_*^{[K]} \mathbf{V}_*^{[K]T} = \underbrace{\left[\begin{matrix} \vdots \\ \{ \ \boldsymbol{\alpha}_{n,*}^T \ \} \\ \vdots \end{matrix} \right]}_{K \times N_{\text{exp}}} \tag{11}$$

The vector $\boldsymbol{\alpha}_{n,*}^T$ collects the amplitudes of the N_{exp} result vectors corresponding to basis vector $\boldsymbol{\varphi}_{n,*}$. The K-rank approximation of the preprocessed snapshot matrix $\mathbf{Y}^{[K,*]}$ can now be written as:

$$\mathbf{Y}^{[K,*]} = \boldsymbol{\Phi}_*^{[K]} \mathbf{A}_*^{[K]} \tag{12}$$

Reduction of the snapshot matrices

Both the zero centered snapshot ($\mathbf{Y}^0$, section 'zero centered') and the snapshot matrix that is scaled per physical part ($\mathbf{Y}^{\text{scaled}}$, section 'Scaling each physical part') are decomposed using the method described in the previous section, resulting in two different bases $\boldsymbol{\Phi}_0$ and $\boldsymbol{\Phi}_{\text{scaled}}$ respectively. These bases that are found using the two different preprocessing methods can both be truncated to K basis vectors to construct the corresponding K-rank surrogate models.

Separate reduction of each physical part

The third method for surrogate model construction uses decomposition of each physical part of the zero centered snapshot matrix separately:

$$
\begin{aligned}
\mathbf{Y}_u^0 &= \left[\mathbf{Y}_u - \overline{\mathbf{y}}_u \mathbf{1}^T \right] = \mathbf{\Phi}_u \mathbf{D}_u \mathbf{V}_u^T \\
\mathbf{Y}_\varepsilon^0 &= \left[\mathbf{Y}_\varepsilon - \overline{\mathbf{y}}_\varepsilon \mathbf{1}^T \right] = \mathbf{\Phi}_\varepsilon \mathbf{D}_\varepsilon \mathbf{V}_\varepsilon^T \\
\mathbf{Y}_\sigma^0 &= \left[\mathbf{Y}_\sigma - \overline{\mathbf{y}}_\sigma \mathbf{1}^T \right] = \mathbf{\Phi}_\sigma \mathbf{D}_\sigma \mathbf{V}_\sigma^T
\end{aligned}
\tag{13}
$$

Note that scaling the different physical parts does not have any influence on the separate bases as explained in "Scaling each physical part" section. As the physical parts are decomposed separately scaling each part will only change the magnitude of the corresponding singular values of the part.

Assembly from one physical part

In this section a new method for assembling the basis vectors from one physical part, hence a submatrix of the preprocessed snapshot matrix, will be described. The basis will be determined for one of the physical quantities, whereafter the full basis will be assembled from this basis, the full basis will describe the directions of variation in the data that covaries most with the variation of the main physical quantity. The reasoning is that certain physical quantities may be dependent on or covariant with others, and that the modes of dependent quantities may be indirectly captured in the basis of the main quantity. The resulting surrogate model will model the main quantity with maximum accuracy, without much loss of accuracy in the other physical parts.

Note that the reduction method can be applied independently of the chosen preprocessing method. By assembling from one physical part the covariances between the different physical parts are retained, while the dominance of other physical parts can be reduced. By means of example we are going to assemble the basis vectors from the zero centered equivalent plastic strain field. However, the method described below can be applied for assembly from any part of the snapshot matrix. One of the physical parts, in this case the zero centered equivalent plastic strain, is decomposed using SVD as described previously.

$$
\mathbf{Y}_\varepsilon^0 = \mathbf{\Phi}_\varepsilon \mathbf{D}_\varepsilon \mathbf{V}_\varepsilon^T
\tag{14}
$$

Note that the left singular vectors of the equivalent plastic strain have size $M_\varepsilon \times 1$. The basis vectors that span the entire result space can be assembled by multiplying the right singular vectors and the singular values from the zero centered equivalent strain with the total zero centered snapshot matrix ($\mathbf{Y}^0$):

$$
\mathbf{\Phi}_{(\varepsilon)} = \begin{bmatrix} \left[\mathbf{\Phi}_{(\varepsilon),u} \right] \\ \left[\mathbf{\Phi}_{(\varepsilon),\varepsilon} \right] \\ \left[\mathbf{\Phi}_{(\varepsilon),\sigma} \right] \end{bmatrix} = \mathbf{Y}^0 \mathbf{V}_\varepsilon \mathbf{D}_\varepsilon^{-1}
\tag{15}
$$

In which $\mathbf{\Phi}_{(\varepsilon)}$ denotes the basis that spans the entire results space assembled from the equivalent plastic strain. Due to the multiplication with the full preprocessed snapshot matrix, the basis vectors are not orthonormal anymore.

To calculate the amplitudes a least squares projection is used:

$$
\mathbf{A}_{(\varepsilon)} = (\mathbf{\Phi}_{(\varepsilon),\varepsilon}^T \mathbf{\Phi}_{(\varepsilon),\varepsilon})^{-1} \mathbf{\Phi}_{(\varepsilon),\varepsilon}^T \mathbf{Y}_\varepsilon^0
\tag{16}
$$

The obtained basis can be truncated to obtain the K-rank approximation of the preprocessed snapshot matrix assembled from the equivalent plastic strain field:

$$
\begin{array}{cc}
\mathbf{Y}^{[K,(\varepsilon)]} = & \mathbf{\Phi}^{[K]}_{(\varepsilon)} \quad \mathbf{A}^{[K]}_{(\varepsilon)} \\
M \times N_{\exp} & M \times K \quad K \times N_{\exp}
\end{array}
\tag{17}
$$

Interpolating the amplitudes

To predict the full result vector at any position of the input parameter space that is not part of the initial sample set $\mathbf{X}$, the amplitudes of the basis vectors will be interpolated. Each row of the amplitude matrix $\mathbf{A}_*$ in equation (11) consists of the amplitudes corresponding to basis vector n. The entries in vector $\boldsymbol{\alpha}_{n,*}$ correspond to each sample point $\mathbf{x}_i$, hence it collects the entries $\alpha_{in,*}$. To construct a continuous surrogate model, the amplitudes corresponding to the results of different input settings will be interpolated.

$$
\boldsymbol{\alpha}_{n,*} \rightarrow \hat{\alpha}_{n,*}(\mathbf{x})
\tag{18}
$$

In the remainder of this section the subscript that indicates the preprocessing method $*$ is left out for clarity. One can also interpolate the right singular vectors as they are simply the amplitudes divided by the singular values [30]. The interpolation can be done using different interpolation methods. In this work Radial Basis Functions (RBFs) are used to interpolate the amplitudes. An RBF can be any function that depends on the distance between an arbitrary point in the parameter space $\mathbf{x}$ and a sample point in the parameter space $\mathbf{x}_i$. The radial basis function corresponding to sample point $\mathbf{x}_i$ will be:

$$
g_i(\mathbf{x}) = g_i(||\mathbf{x} - \mathbf{x}_i||)
\tag{19}
$$

Several different basis functions have been proposed in literature. In multiple studies it has been shown that the performance of multiquadric RBF for scalar interpolation is generally good [2,42]. In a comparative study performed by Hamim [43] on the application of RBF in POD-based surrogate models, it was shown that multiquadric RBF perform best compared to other types of RBFs. Therefore a multiquadric RBF is chosen for interpolation of the amplitudes in all surrogate models:

$$
g_{n,i}(\mathbf{x}) = \sqrt{1 + (\mathbf{x} - \mathbf{x}_i)^T \theta_n (\mathbf{x} - \mathbf{x}_i)}
\tag{20}
$$

In which θ_n is the $N_{\dim} \times N_{\dim}$ diagonal global scaling matrix of the amplitude corresponding to basis vector n. The global scaling parameters are optimized per basis vector and per surrogate model based on the Leave-One-Out Cross Validation values as proposed in [44].

As the mean result vector is subtracted from the snapshot matrix, the mean amplitude will be (close to) 0. No polynomial detrending is applied before interpolating the amplitudes [45,46]. The interpolated amplitude is written as the sum of the RBFs $g_{n,i}(\mathbf{x})$ multiplied with a weight $w_{n,i}$:

$$
\hat{\alpha}_n(\mathbf{x}) = \sum_{i=1}^{N_{\exp}} w_{n,i} g_{n,i}(\mathbf{x})
\tag{21}
$$

One radial basis function is placed at the coordinate of each sample point in the input parameter space. The weights of the basis functions can be determined exactly, as the amplitudes are known at this same number of coordinates. To find the unknown weights all RBFs can be evaluated at the initial sample points $\mathbf{x}_i$ and collected in an $N_{\exp} \times N_{\exp}$

interpolation matrix $\mathbf{G}_n$ [31].

$$
\mathbf{G}_n = \begin{bmatrix} g_{n,1}(\mathbf{x}_1) & \cdots & g_{n,1}(\mathbf{x}_{N_{\mathrm{exp}}}) \\ \vdots & \ddots & \vdots \\ g_{n,N_{\mathrm{exp}}}(\mathbf{x}_1) & \cdots & g_{n,N_{\mathrm{exp}}}(\mathbf{x}_{N_{\mathrm{exp}}}) \end{bmatrix} \tag{22}
$$

The weights in equation (21) can be solved by setting $\hat{\alpha}_n(\mathbf{x}_i) = \alpha_{in}$ leading to:

$$
\mathbf{w}_n^T = \boldsymbol{\alpha}_n^T \mathbf{G}_n^{-1} \tag{23}
$$

Post-processing

When the amplitudes corresponding to all K basis vectors are interpolated the continuous approximation in a preprocessed basis reads:

$$
\hat{\mathbf{y}}^{[K,*]}(\mathbf{x}) = \sum_{n=1}^{K} \boldsymbol{\varphi}_{n,*} \hat{\alpha}_{n,*}(\mathbf{x}) \tag{24}
$$

To construct the final surrogate model, the approximation in the truncated basis must be mapped back to the same form as the initial snapshot matrix. The function that reverses the applied preprocessing method is called the post-processing function and is denoted with $f_*^{-1}(\cdot)$. The interpolated result vector can be written as:

$$
\hat{\mathbf{y}}_*^{[K]}(\mathbf{x}) = f_*^{-1}(\hat{\mathbf{y}}^{[K,*]}(\mathbf{x})) = f_*^{-1}\left(\sum_{n=1}^{K} \boldsymbol{\varphi}_{n,*} \hat{\alpha}_{n,*}(\mathbf{x}) \right) \tag{25}
$$

The surrogate model with K basis vectors from the zero centered snapshot matrix can be written as:

$$
\hat{\mathbf{y}}_0^{[K]}(\mathbf{x}) = \bar{\mathbf{y}} + \sum_{n=1}^{K} \boldsymbol{\varphi}_{n,0} \hat{\alpha}_{n,0}(\mathbf{x}) \tag{26}
$$

The surrogate model with K basis vectors from the zero centered and scaled snapshot matrix can be written as:

$$
\hat{\mathbf{y}}_{\mathrm{scaled}}^{[K]}(\mathbf{x}) = \bar{\mathbf{y}} + \left(\sum_{n=1}^{K} \boldsymbol{\varphi}_{n,\mathrm{scaled}} \hat{\alpha}_{n,\mathrm{scaled}}(\mathbf{x}) \right) \oslash \mathbf{s} \tag{27}
$$

In which $\oslash$ is the Hadamard division, respresenting element wise division and $\mathbf{s}$ is the $M \times 1$ scaling vector with M_{u} entries s_{u}, M_ε entries s_ε and M_σ entries s_σ which are calculated based on equation (6).

The surrogate model with K basis vectors from the zero centered separate snapshot matrices can be written as:

$$
\hat{\mathbf{y}}_{\mathrm{sep}}^{[K]}(\mathbf{x}) = \left\{ \begin{Bmatrix} \{\hat{\mathbf{y}}_{\mathrm{u}}^{[K]}(\mathbf{x})\} \\ \{\hat{\mathbf{y}}_\varepsilon^{[K]}(\mathbf{x})\} \\ \{\hat{\mathbf{y}}_\sigma^{[K]}(\mathbf{x})\} \end{Bmatrix} \right\} = \left\{ \begin{Bmatrix} \left\{ \bar{\mathbf{y}}_{\mathrm{u}} + \sum\limits_{n=1}^{K} \boldsymbol{\varphi}_{n,\mathrm{u}} \hat{\alpha}_{n,\mathrm{u}}(\mathbf{x}) \right\} \\ \left\{ \bar{\mathbf{y}}_\varepsilon + \sum\limits_{n=1}^{K} \boldsymbol{\varphi}_{n,\varepsilon} \hat{\alpha}_{n,\varepsilon}(\mathbf{x}) \right\} \\ \left\{ \bar{\mathbf{y}}_\sigma + \sum\limits_{n=1}^{K} \boldsymbol{\varphi}_{n,\sigma} \hat{\alpha}_{n,\sigma}(\mathbf{x}) \right\} \end{Bmatrix} \right\} \tag{28}
$$

When separate snapshot matrices are used a different number of basis vectors in the truncated basis K can be chosen for each physical part. Note that, when using separate bases for each physical quantity, one surrogate model must be fitted per basis vector per

physical quantity, whereas the other decomposition methods require fitting of only one surrogate model per basis vector for all physical quantities together.

The surrogate model with K basis vectors from the zero centered snapshot matrix that is assembled from the strain field can be written as:

$$\hat{\mathbf{y}}_{(\varepsilon)}^{[K]}(\mathbf{x}) = \bar{\mathbf{y}} + \sum_{n=1}^{K} \boldsymbol{\varphi}_{n,(\varepsilon)} \hat{\alpha}_{n,(\varepsilon)}(\mathbf{x}) \tag{29}$$

Demonstrator processes

As discussed earlier, the proposed methods for surrogate model construction described in this paper are generally applicable and not case specific. However, in order to illustrate the methods in concrete terms, two demonstrator process are introduced in this section.

Pure bending problem

The first demonstrator process is a rectangular beam under pure bending. The beam has height $h = 4$ mm and thickness $b = 2$ mm as depicted in Fig. 2. The beam is made of an imaginary material with Young's Modulus 2,10,000 MPa and a Poisson ratio of 0.3. Furthermore the relation between the stress σ and the plastic strain ε_p is modelled with the following hardening curve:

$$\sigma = \sigma_y + K(\varepsilon_0 + \varepsilon_p)^n \tag{30}$$

Herein σ_y is the yield stress offset, K is the strength coefficient, ε_0 is the prior plastic strain and n is the hardening exponent.

The Hill quadratic planar isotropic yield criterion is used to determine when the material is yielding:

$$\sigma_y = \sqrt{\frac{R(\sigma_{xx} - \sigma_{yy})^2 + \sigma_{xx}^2 + \sigma_{yy}^2}{1 + R}} \tag{31}$$

Herein R is the Lankford coefficient. The aforementioned five parameters that describe the hardening curve and the yield criterion vary according to Table 1. Hence, the dimensionality of the input parameters space is $N_{\text{dim}} = 5$ and a point in the parameter space can be denoted as:

$$\mathbf{x} = \{x_1, x_2, x_3, x_4, x_5\}^T = \{\sigma_y, K, \varepsilon_0, n, R\}^T \tag{32}$$

The solution to several outputs is calculated at 101 points across the beam height. Note that in Fig. 2 only 17 points are displayed for clarity of the illustration. The neutral axis (NA) that is depicted with the dashed line in Fig. 2 is defined as the point where the strain in x-direction $\varepsilon_x = 0$. The beam is bent to a maximum curvature of 0.2 mm^{-1} measured on the neutral axis, in a series of 101 subsequent steps, in which the curvature is gradually increased with a constant amount.
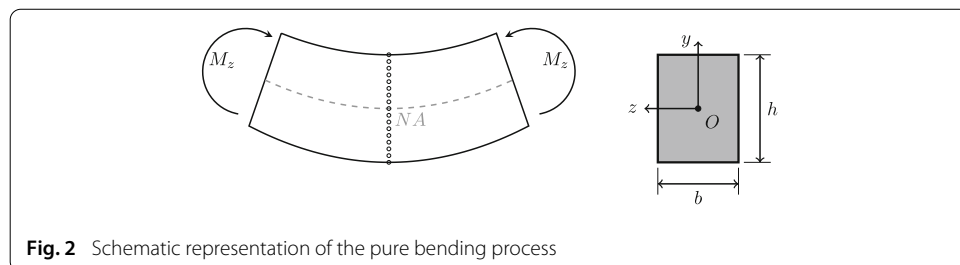
Fig. 2 Schematic representation of the pure bending process

Table 1 Ranges and dimensions of the input and output of the pure bending problem

	Variable	Minimum	Maximum	Unit
Input	Yield stress offset (σ_y)	100	200	MPa
	Strength coefficient (K)	200	250	MPa
	Prior plastic strain (ε_0)	0.01	0.1	[–]
	Hardening exponent (n)	0.1	0.4	[–]
	Lankford coefficient (R)	2/3	3/2	[–]
Output	Equivalent plastic strain (ε_p^{eq})	0.0250	0.444	[–]
	Engineering strain in y-direction (ε_y^{eng})	-0.1914	0.302	[–]
	Stress in x-direction (σ_{xx})	-460	423	MPa
	Stress in y-direction (σ_{yy})	-157	0.5	MPa

To maintain a pure bending situation without axial forces, the neutral line will shift in y-direction during the increments. The neutral line shift is solved iteratively based on the condition that the axial force should be equal to zero. Within each iteration, the strains in x-direction are fully defined by the curvature of the neutral line, and the neutral line shift. To assure the accuracy of the calculations, the local point density is increased in the elastic region. The stress and strain increments at all points can be solved using elastoplasticity theory, using associated flow, the Hill yield locus of Eq. (31), isotropic hardening, and the hardening rule of Eq. (30). The calculations have been implemented in a Matlab code.

The final solution of several outputs on the original locations of the 101 points will be modelled. The output of the model consists of the equivalent plastic strain ε_p^{eq}, the engineering strain in y-direction ε_y^{eng}, the stress in x-direction σ_{xx} and the stress in y-direction σ_{yy}. Hence the result of one evaluation can be collected in the $M \times 1$ result vector:

$$\mathbf{y} = \begin{Bmatrix} \{\mathbf{y}_\varepsilon\} \\ \{\mathbf{y}_\sigma\} \end{Bmatrix} \tag{33}$$

where $\mathbf{y}_\varepsilon$ is the $M_\varepsilon \times 1$ part of the result vector that corresponds to the two different strains and $\mathbf{y}_\sigma$ is the $M_\sigma \times 1$ part of the result vector that corresponds to the two stress components. With 101 points per output type the total length of the result vector is $M = 404$.

Construction of pure bending surrogate models

The pure bending model is used to construct different surrogate models thereof. As described in the section 'Obtaining a training set', the first step is to obtain a training set. To rule out the dependency on the training set when comparing the four different approaches in construction of a surrogate model, five distinct training sets will be used. Each training set consists of a sample set $\mathbf{X}$ and a snapshot matrix $\mathbf{Y}$.

Three different designs are combined to sample the $N_{\dim} = 5$ dimensional parameter space of the pure bending demonstrator: a star point design, a full factorial design and a space-filling LHS. The star point and the full factorial designs, consists of $N_{\text{star}} = 5 \cdot N_{\dim} + 1 = 11$ and $N_{\text{ff}} = 2^5 = 32$ sample points respectively. To combine the star point and full factorial design with the LHS, 250 different space-filling LHS of $N_{\text{lhs}} = 107$ sample points are generated in the input space and the 5 best are picked based on the maximum minimum Euclidean distance to the points in the star point and full factorial design. The total number of experiments in each training set will be $N_{\exp} = N_{\text{star}} + N_{\text{ff}} + N_{\text{lhs}} = 150$.

Table 2 Scaling constants for the different training sets of the pure bending problem

	Set 1	Set 2	Set 3	Set 4	Set 5
Strain (s_ε)	8.747	8.800	8.800	8.800	8.798
Stress ($s_\sigma/10^{-3}$)	1.571	1.426	1.667	1.735	1.627

Now that we have 5 different sample sets **X**, each sample point is evaluated using the pure bending model as described in the section 'Pure bending problem'. In total the model is evaluated $11 + 32 + 5 \cdot 107 = 578$ times. The result vectors from each sample set are placed in five different snapshot matrices. Each snapshot matrix can be partitioned into a submatrix corresponding to the strain field $\mathbf{Y}_\varepsilon$ and the stress field $\mathbf{Y}_\sigma$:

$$\mathbf{Y} = \begin{bmatrix} [\mathbf{Y}_\varepsilon] \\ [\mathbf{Y}_\sigma] \end{bmatrix} \tag{34}$$

To construct the four different surrogate models, first the snapshot matrices are zero centered as defined in equation (4). The first type of surrogate model will be based directly on the zero centered snapshot matrices.

In the construction of the second surrogate model the zero centered snapshot matrices will be scaled as defined in equation (5). The scaling constants s_ε and s_σ that are calculated using equation (6) for each training set can be found in Table 2.

To construct the third surrogate model separate zero centered snapshot matrices will be used for each physical part.

The fourth and fifth surrogate models are constructed by assembling the full basis vectors from the strain and stress field respectively. The ratio between the variables in the equivalent plastic strain field and the total number of variables in the result fields is:

$$\frac{M_\varepsilon}{M} = \frac{202}{404} = 0.5 \tag{35}$$

Hence, the total result field is assembled from approximately 50% of its total span. Equivalently, if the basis vectors are assembled from the stress, the total result field will also be approximated from 50% of its total span.

Summarizing, surrogate models of the pure bending problem are constructed using:

1. a zero centered snapshot matrix ($\hat{\mathbf{y}}_0^{[K]}(\mathbf{x})$).
2. a zero centered snapshot matrix that is scaled per physical part ($\hat{\mathbf{y}}_{\text{scaled}}^{[K]}(\mathbf{x})$).
3. separate zero centered snapshot matrices for each physical part ($\hat{\mathbf{y}}_{\text{sep}}^{[K]}(\mathbf{x})$).
4. a zero centered snapshot matrix with basis vectors assembled from strain field ($\hat{\mathbf{y}}_{(\varepsilon)}^{[K]}(\mathbf{x})$).
5. a zero centered snapshot matrix with basis vectors assembled from stress field ($\hat{\mathbf{y}}_{(\sigma)}^{[K]}(\mathbf{x})$).

based on the 5 training sets.

Industrial bending problem

For the second demonstrator we will investigate the bending of a metal flap. A schematic representation of the process can be found in Fig. 3. In the process a sheet metal work piece with initial sheet thickness (x_1) is bent downward to the final punch depth (x_2). Thereafter, the punch is released and the work piece finds a new equilibrium position after springback.

Fig. 3 Schematic representation of the demonstrator process and Finite Element mesh
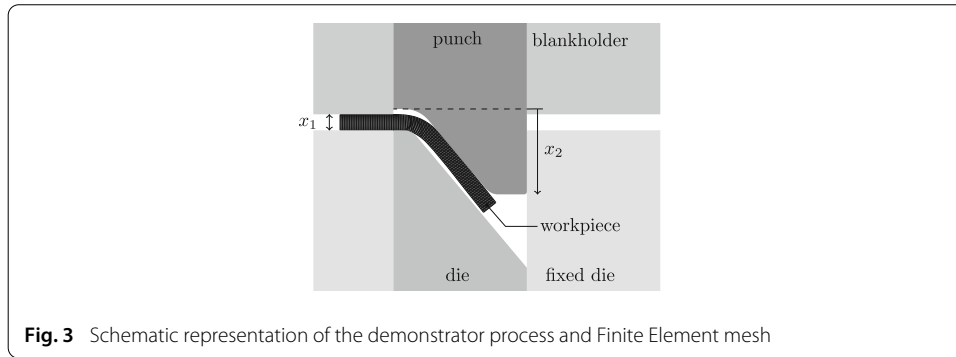
Table 3 Ranges and dimensions of the input and output of the industrial bending problem

	Variable	Minimum	Nominal	Maximum	Unit
Input	Thickness (x_1)	0.295	0.3	0.305	mm
	Punch depth (x_2)	1.5	1.55	1.6	mm
Output	Displacement (u)	− 1.69	n/a	0.07	mm
	Equivalent plastic strain (ε_p^{eq})	0	n/a	0.48	[–]
	Stress (σ)	− 661	n/a	566	MPa

The process is described by two parameters, the sheet thickness (x_1) and the final punch depth (x_2). Hence, we have a parameter space with $N_{\dim} = 2$. A point in the parameter space can be denoted as:

$$\mathbf{x} = \{x_1, x_2\}^T \tag{36}$$

The process is modelled with a 2D plane strain model using the FE analysis software MSC Marc/Mentat version 2016. The sheet metal is modelled with an elastic–plastic isotropic material model with Von Mises yield criterion and a tabular hardening curve. The work piece is meshed using 1200 quadrilateral elements and 1296 nodes. The elements are fully integrated using four integration points per element with constant dilatational strain.

The output of the FE-model consists of the nodal displacement field and the equivalent plastic strain field and stress field in the integration points. Hence the result after the final time step of the FE-model can be collected in the $M \times 1$ result vector:

$$\mathbf{y} = \begin{Bmatrix} \{\mathbf{y}_u\} \\ \{\mathbf{y}_\varepsilon\} \\ \{\mathbf{y}_\sigma\} \end{Bmatrix} \tag{37}$$

where $\mathbf{y}_u$ is the $M_u \times 1$ part of the result vector that corresponds to the displacement field of all nodes, $\mathbf{y}_\varepsilon$ is the $M_\varepsilon \times 1$ part of the result vector that corresponds to the equivalent plastic strain in all integration points and $\mathbf{y}_\sigma$ is the $M_\sigma \times 1$ part of the result vector that corresponds to the stress components in all nodes. With two degrees of freedom per node $M_u = 2592$. With four integration points per element $M_\varepsilon = 4800$. With four integration points per element and because of the plane strain analysis there are 4 nonzero stress components which makes $M_\sigma = 19200$. The overall length of a result vector is $M = M_u + M_\varepsilon + M_\sigma = 26592$. An overview of the different ranges and dimensions of both the input and output parameters is given in Table 3.

Table 4 Scaling constants for the different training sets

	Set 1	Set 2	Set 3	Set 4	Set 5
Displacement ($s_u/10^1$)	1.551	1.551	1.691	1.516	1.620
Equivalent plastic strain (s_ε)	4.784	4.776	4.771	4.742	4.770
Stress ($s_\sigma/10^{-4}$)	8.397	8.458	8.386	8.446	8.433

Construction of industrial bending surrogate models

Similarly as described in "Construction of pure bending surrogate models" section different surrogate models of the industrial bending problem will be constructed. Again the first step is to gather 5 training sets that consist of a sample set $\mathbf{X}$ and a snapshot matrix $\mathbf{Y}$. Each sample set will consist of a starpoint design with $N_{\text{star}} = 2 \cdot N_{\text{dim}} + 1 = 5$ sample points combined with a space-filling LHS of $N_{\text{lhs}} = 75$ sample points. The same procedure as used for the other demonstrator is used to combine the sample points in the input space. The total number of experiments in each training set is $N_{\text{exp}} = N_{\text{star}} + N_{\text{lhs}} = 80$.

Now that we have 5 different sample sets $\mathbf{X}$, each sample point is evaluated using FE analysis as described in the section 'Industrial bending problem'. In total $5 + 5 \cdot 75 = 380$ FE simulations are done on the different locations in the input space according to the proposed sampling. The result vectors from each sample set are placed in five different snapshot matrices. Each snapshot matrix can be partitioned into a submatrix corresponding to the displacement field $\mathbf{Y}_u$, the equivalent plastic strain field $\mathbf{Y}_\varepsilon$ and the stress field $\mathbf{Y}_\sigma$:

$$\mathbf{Y} = \begin{bmatrix} [\mathbf{Y}_u] \\ [\mathbf{Y}_\varepsilon] \\ [\mathbf{Y}_\sigma] \end{bmatrix} \tag{38}$$

To construct the four different surrogate models, first the snapshot matrices are zero centered as defined in Eq. (4). The first type of surrogate model will be based directly on the zero centered snapshot matrices.

In the construction of the second surrogate model the zero centered snapshot matrices will be scaled as defined in equation (5). The scaling constants s_u, s_ε and s_σ that are calculated using Eq. (6) for each training set can be found in Table 4.

To construct the third surrogate model separate zero centered snapshot matrices will be used for each physical part.

The fourth and last surrogate model is constructed by assembling the full basis vectors from the equivalent plastic strain field. The equivalent plastic strain is chosen as maximum accuracy of this physical part is required. The ratio between the variables in the equivalent plastic strain field and the total number of variables in the result fields is:

$$\frac{M_\varepsilon}{M} = \frac{4800}{26592} \approx 0.18 \tag{39}$$

Hence, the total result field is assembled from approximately 18% of its total span.

Summarizing, surrogate models are constructed using:

1. a zero centered snapshot matrix ($\hat{\mathbf{y}}_0^{[K]}(\mathbf{x})$).
2. a zero centered snapshot matrix that is scaled per physical part ($\hat{\mathbf{y}}_{\text{scaled}}^{[K]}(\mathbf{x})$).
3. separate zero centered snapshot matrices for each physical part ($\hat{\mathbf{y}}_{\text{sep}}^{[K]}(\mathbf{x})$).
4. a zero centered snapshot matrix with basis vectors assembled from equivalent plastic strain field ($\hat{\mathbf{y}}_{(\varepsilon)}^{[K]}(\mathbf{x})$).

based on the 5 training sets.

Comparing different bases

Due to the differences in training set, preprocessing method and decomposition methods surrogate models, different bases will be used for each surrogate model. The bases used in the construction of the different surrogate models as presented in "Construction of pure bending surrogate models" and "Construction of industrial bending surrogate models" sections are analyzed to investigate their impact on the surrogate model quality. We will first analyze the bases using the singular values as commonly done in literature ("Basis analysis using singular values" section), thereafter a new measure is introduced that takes into account the different physical parts in the snapshot matrices ("Basis analysis considering physical parts" section).

Basis analysis using singular values

It is useful to assess the contribution of each basis vector to the solution to determine how many basis vectors should be included in the truncated basis. Recall the relation between the singular values and the eigenvalues as described in "Decomposing the output" section. The eigenvalues are considered to give a measure for the contribution of the corresponding basis vector. To choose the number of basis vectors to be retained in the truncated basis, the ratio between the summation of K included eigenvalues and the summation of all of them is often used. The ratio is also referred to as *cummulative energy* [29,43], *contribution* [33] or in PCA the *variance explained* [38]. The ratio is defined as:
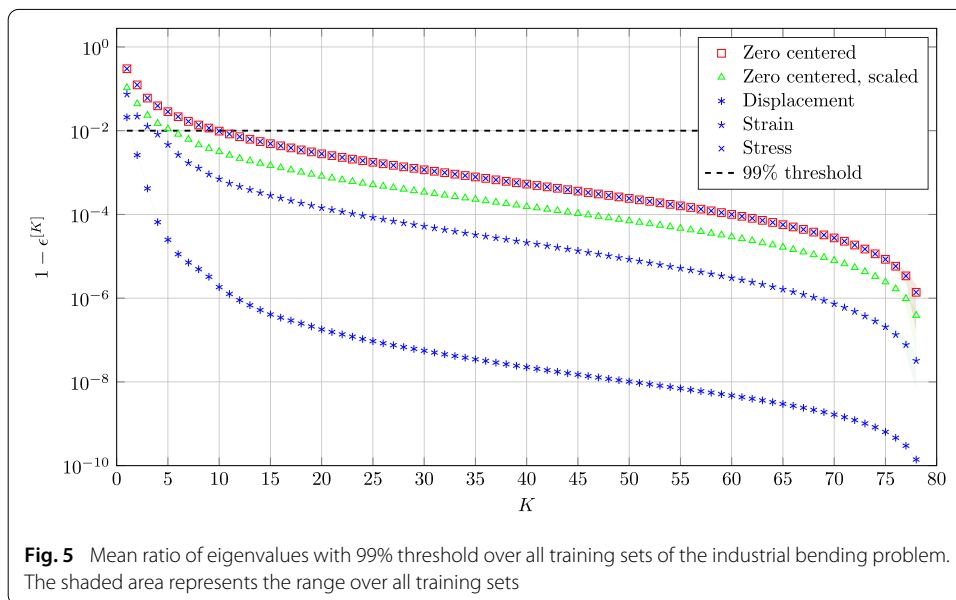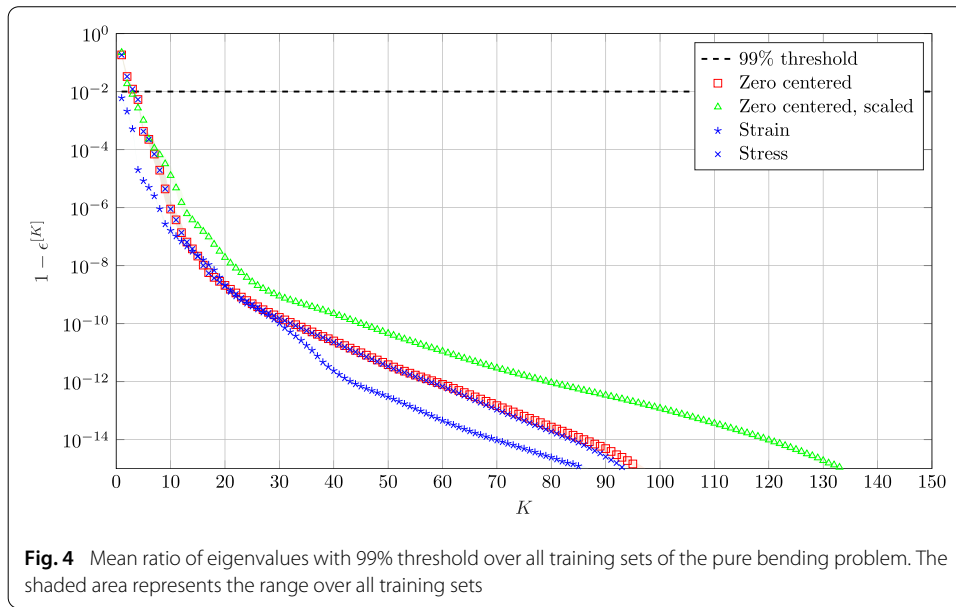
$$\epsilon^{[K]} = \frac{\sum_{n=1}^{K} \lambda_n}{\sum_{n=1}^{N_{\exp}-1} \lambda_n} \tag{40}$$

The ratio which is considered sufficient can be set and is called the cut-off ratio or threshold. For example, in the work of Hamim [43] the cut-off ratio is set to 99%, whereas in the work of Buljak [31] it is set to 99.999%. The ratios as defined in equation (40) are plotted for the different obtained bases in Fig. 4 for the pure bending problem and Fig. 5 for the industrial bending problem.

When Figs. 4 and 5 are analysed it is found that the eigenvalues of the zero centered snapshot matrix and the eigenvalues of the submatrix containing the zero centered stress field are almost identical. This can be explained by the fact that the numerical values of the stress data are several orders of magnitude larger than the numerical values of the strain and displacement data, and therefore the variation in stresses dominates the bases. In both problems the basis vectors of the stress field have lowest cumulative energy $\epsilon^{[K]}$, the corresponding basis vectors will therefore describe less variation compared to the other obtained basis vectors.

For the pure bending problem, the basis vectors of submatrix that contains the strain field have the highest contribution overall. The contribution of basis vectors from the zero centered and scaled snapshot matrix is the lowest when using 6 or more basis vectors.

Based on Fig. 4 it can be concluded that to retain 99% of the variation in the bases, from the zero centered snapshot matrix, the zero centered and scaled snapshot matrix and the snapshot matrix that only contains the stress components, 3 basis vectors are needed. The snapshot matrix that only contains the strain retains 99% of the variation with 1 basis vector only. Similarly the number of basis vectors that are needed to retain 99.999% of

Fig. 4 Mean ratio of eigenvalues with 99% threshold over all training sets of the pure bending problem. The shaded area represents the range over all training sets

Fig. 5 Mean ratio of eigenvalues with 99% threshold over all training sets of the industrial bending problem. The shaded area represents the range over all training sets

the variation can be determined using Fig. 4. The zero centered snapshot matrix, the zero centered and scaled snapshot matrix, the snapshot matrix that only contains the strain components and the snapshot matrix that only contains the stress components, need 9, 11, 5 and 9 basis vectors respectively to retain 99.999% of the variation.

For the industrial bending problem, both the zero centered snapshot matrix and the submatrix containing the stress field have the lowest cumulative energy. The basis vectors of the displacement field have the highest contribution. Scaling the snapshot matrix within each physical part increases the contribution of the first basis vectors. Hence, the decomposition is less dominated by the part that corresponds to the stress.

Based on Fig. 5 it can be concluded that to retain 99% of the variation in the bases, from the zero centered snapshot matrix and the snapshot matrix that only contains the stress

components, 10 basis vectors are needed. Using the zero centered and scaled snapshot matrix 99% of the variation will be retained using only 5 basis vectors. The snapshot matrices that only contain the strain and displacement components retain 99% of the variation with 4 or 2 vectors respectively.

Basis analysis considering physical parts

When the snapshot matrix is decomposed as a whole, the analysis based on the singular values does not take into account the contributions to the different physical parts. To analyse the contribution of each basis vector with respect to each physical part we will introduce the truncation error. The truncation error κE arises due to the truncation of the basis. If all basis vectors would be used, hence $K = N_{\exp} - 1$ the result vectors in the training set $\mathbf{y}_i$ can be reproduced without any error. The truncated approximation of the snapshot matrix is presented in Eq. (12), and hence we can write:

$$\mathbf{Y} = \mathbf{Y}^{[K]} + \kappa E^{[K]} \tag{41}$$

The truncation error for result vector $\mathbf{y}_i$ in the training set approximated using K basis vectors is:

$$\boxed{\kappa E_i^{[K]} = \mathbf{y}_i - \mathbf{y}_i^{[K]}} \tag{42}$$

This truncation error describes the error in all points, nodes or integration points of the different physical parts. We can split up the errors in a part corresponding to the displacement field $\kappa E_{u,i}^{[K]}$, the equivalent plastic strain field $\kappa E_{\varepsilon,i}^{[K]}$ and the stress field $\kappa E_{\sigma,i}^{[K]}$. The relative error is calculated by dividing the Euclidean norm of the error, by the Euclidean norm of the deviation in the training vector. This is done for each physical part to make the relative error dimensionless. For example, the error in the displacement field can be normalised with:

$$\kappa RE_i = \frac{\|\kappa E_{u,i}^{[K]}\|_2}{\|\mathbf{y}_{u,i} - \bar{\mathbf{y}}_u\|_2} \tag{43}$$

The normalised error for the other physical parts is calculated equivalently. The truncation Root Mean Square Relative Error ($\kappa RMSRE$) is calculated over all $N_{\exp} = 150$ and $N_{\exp} = 80$ results in the training set of the pure and industrial bending problems respectively. The relative errors of each physical part are squared and summed leading to:

$$\kappa RMSRE^{[K]} = \sqrt{\frac{1}{N_{\exp}} \sum_{i=1}^{N_{\exp}} \kappa RE_i^2} \tag{44}$$

Figures 6 and 7 presents the *RMSRE* of the different preprocessing and decomposition methods applied to the five different training sets of the pure and industrial bending problems respectively. In each subfigure the different preprocessing and decomposition methods are compared to separate decomposition of each physical part, which is the theoretical best result that can be achieved for each physical quantity using that training set. The solid line with markers in Figs. 6 and 7 describes the mean *RMSRE* over the 5 training sets. The shaded area describes the minimum and maximum *RMSRE* over all training sets. Due to the zero centering of all snapshot matrices, a full basis consists of $K = N_{\exp} - 1$ basis vectors. When a full basis is used, the truncation error drops to 0.
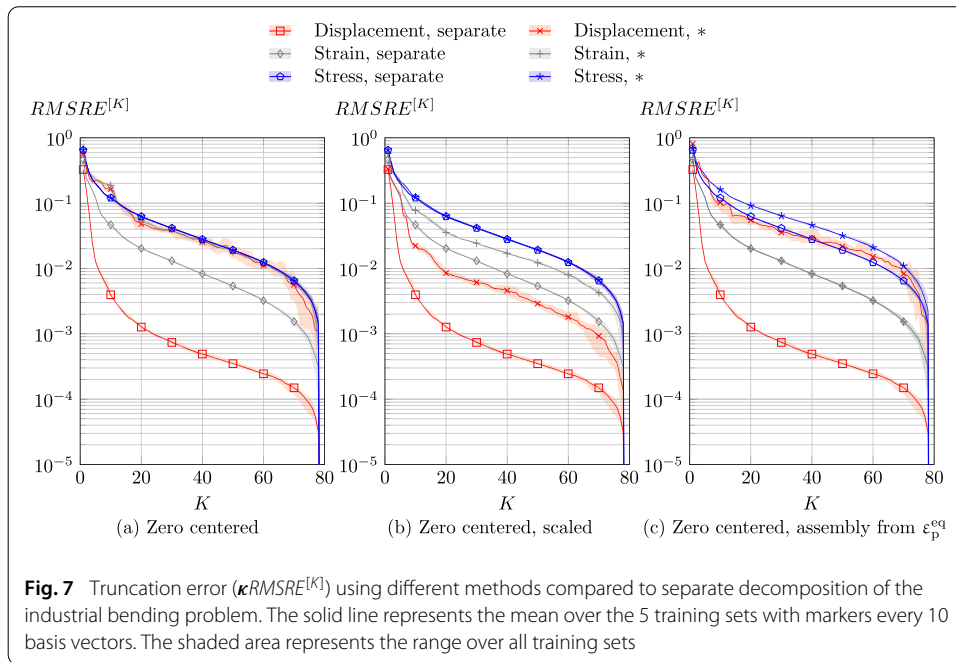
Note that having an *RMSRE* of 10^{-2} is comparable to having a cumulative energy (equation (40)) of 99%. For the pure bending problem, Fig. 6, to meet this 99% threshold
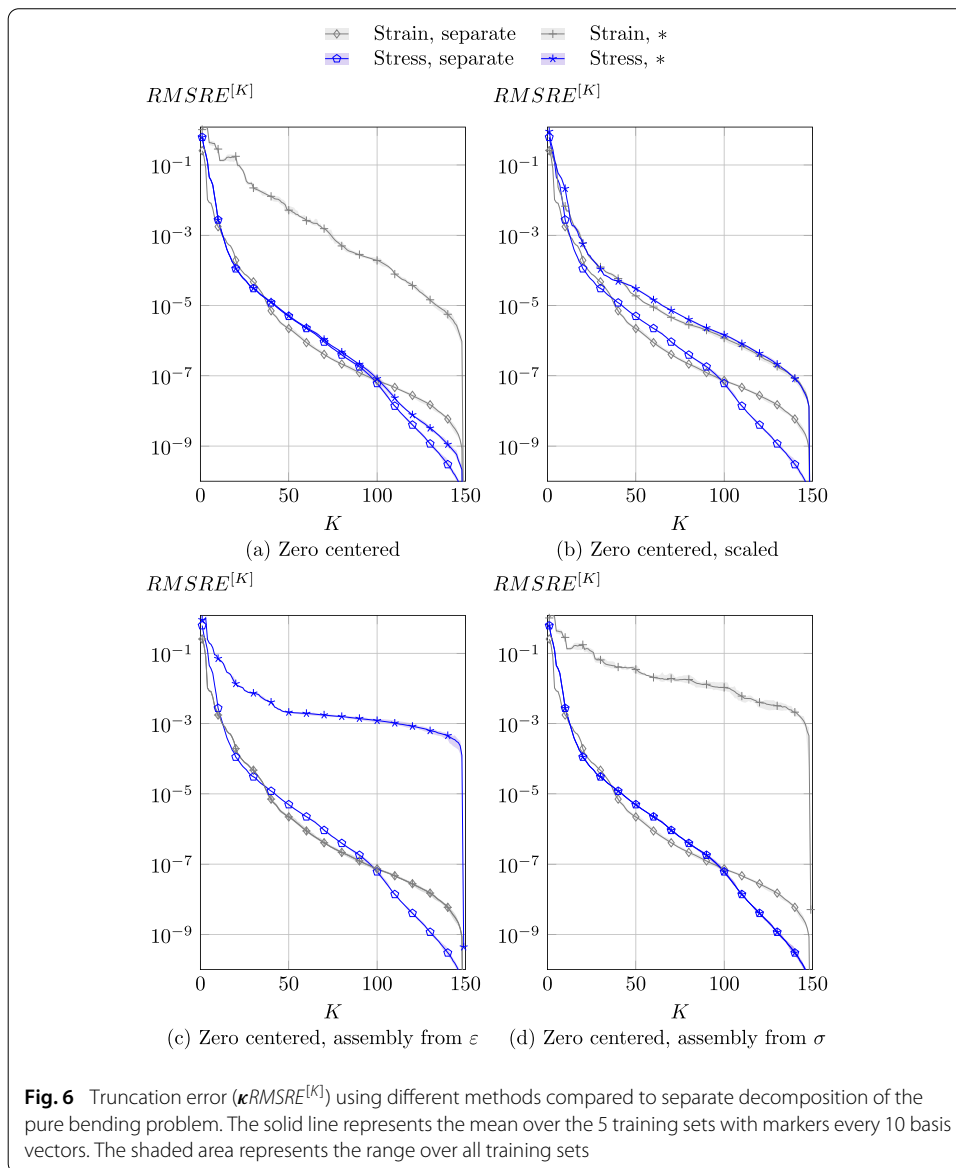
Fig. 6 Truncation error ($\kappa RMSRE^{[K]}$) using different methods compared to separate decomposition of the pure bending problem. The solid line represents the mean over the 5 training sets with markers every 10 basis vectors. The shaded area represents the range over all training sets

using separate snapshot matrices for each physical part, we need a basis with $K = 5$ and 8 for the strain and stress respectively. When separate snapshot matrices for each physical part are used for modelling the industrial bending problem, bases with $K = 5$, 36 and 65 for the displacement, strain and stress respectively are needed to meet this threshold (Fig. 7).

In both demonstrator problems, the decomposition of the zero centered snapshot matrix is dominated by the stress components (Figs. 6a and 7a). The basis vectors resulting from the zero centered snapshot matrix, describe less variation in the displacement and strain fields leading to higher *RMSRE* in those parts. When the basis vectors from the zero centered snapshot matrix are used to construct a surrogate model, the number of basis vectors needed to achieve error values similar to using separate snapshot matrices for the displacement and strain goes up to $K = 40$ for the pure bending problem and $K = 65$ for the industrial bending problem.

Fig. 7 Truncation error ($\kappa RMSRE^{[K]}$) using different methods compared to separate decomposition of the industrial bending problem. The solid line represents the mean over the 5 training sets with markers every 10 basis vectors. The shaded area represents the range over all training sets

When a zero centered and scaled snapshot matrix is used to obtain the basis vectors the approximation of the displacement and strain field is improved. The approximation of the stress components is slightly worsened for the pure bending problem (Fig. 6b), while not significantly worsened for the industrial bending problem (Fig. 7b).

Lastly, when the basis vectors are assembled from one physical part, generally this physical part can be approximated with the same accuracy as if separate snapshot matrices would be used (Figs. 6c, d, 7c). Assembling the basis vectors from one physical part worsens the approximation of the other part(s). For the industrial bending problem the approximation of the stress part of the snapshot matrix is worse, while the performance of the part corresponding to the displacement field is comparable to decomposing the zero centered snapshot matrix as a whole.

Validation set results of the demonstrator process

To validate the proposed methods for construction and decomposition of the snapshot matrix, a validation set is used. For the pure bending problem all remaining LHS samples are used. For example when validating the surrogate models that are based on training set 2, training sets 1, 3, 4 and 5 are used for validation. The total number of sample points in the validation set is thus $N_{\text{val}} = 4 \cdot 139 = 428$. The validation set of the industrial bending problem is obtained by sampling a 9×9 grid leading to an additional $N_{\text{val}} = 81$ simulations. The sample points in the validation set are denoted with $\mathbf{x}_p$. The result vectors from the simulation evaluated at each point $\mathbf{x}_p$ are considered as the reference solution and are denoted with:

$$\mathbf{y}_p \quad \text{with: } p = 1...N_{\text{val}} \tag{45}$$

As proposed by Abdi [17] a result vector from a validation set can be mapped into the truncated basis using a least squares projection. The amplitudes in the reduced basis

corresponding to the reference solution can be found using:

$$\alpha_{np,*} = \boldsymbol{\varphi}_{n,*}^T f_*(\mathbf{y}_p) \tag{46}$$

Note that for the basis vectors projected in the strain field the amplitudes of supplementary observations must be found using Eq. (16). The reference solution can be approximated in the truncated basis using any preprocessing method with K basis vectors as:

$$\mathbf{y}_{*,p}^{[K]} = f_*^{-1}(\boldsymbol{\Phi}_*^{[K]} \boldsymbol{\Phi}_*^{[K]T} f_*(\mathbf{y}_p)) \tag{47}$$

By means of example, a reference solution can be estimated with K basis vectors from the zero centered snapshot matrix as:

$$\mathbf{y}_{0,p}^{[K]} = \bar{\mathbf{y}} + \boldsymbol{\Phi}_0^{[K]} \boldsymbol{\Phi}_0^{[K]T} \left(\mathbf{y}_p - \bar{\mathbf{y}} \right) \tag{48}$$

We define the continuous approximation of a reference solution, that is the surrogate model evaluated at point $\mathbf{x}_p$, as:

$$\hat{\mathbf{y}}_*^{[K]}(\mathbf{x}_p) = f_*^{-1} \left(\sum_{n=1}^{K} \boldsymbol{\varphi}_{n,*} \hat{\alpha}_{n,*}(\mathbf{x}_p) \right) \tag{49}$$

Error measures

The different surrogate models approximate the reference solution $\hat{\mathbf{y}}_*^{[K]}(\mathbf{x}_p) \approx \mathbf{y}_p$. This approximation is not exact, several sources of errors can be distinguished. First there is an error due to the truncation of the basis, as described in "Basis analysis considering physical parts" section. If all basis vectors are used in the bases, the result vectors in the training set can be reproduced without any error because the obtained basis vectors can describe all variation in the training set. However, the basis can lack in richness to describe the variation present in the results from the validation set. In that case, even when a full basis is used in the surrogate model, the results in the validation set cannot be approximated without any error.

Secondly there is an interpolation error resulting from the RBF interpolation of the amplitudes as described in "Interpolating the amplitudes" section. The amplitudes can be approximated exactly on the sampling points in the training set, any point that is not in the training set will have an error. Another measure is simply the error between the reference solution and its approximation using a surrogate model. We will refer to this measure as the total reconstruction error. We will use the *RMSRE* as defined in Eq. (44) to quantify these errors. In the following sections the aforementioned error components will be formulated.

Truncation error

As the basis becomes richer when more basis vectors are added, the truncation error becomes smaller. However, it is not expected that a training set will contain all possible variations that may occur inside the parameter space. Therefore, the training data will not be able to capture all features that are present in the validation data. We can define the truncation error as the difference between the reference solution and its approximation using K basis vectors (Eq. (47)). The truncation error when K basis vectors are included in the approximation of reference solution $\mathbf{y}_p$ will be:

$$\boxed{\kappa E_p^{[K]} = \mathbf{y}_p - \mathbf{y}_{*,p}^{[K]}} \tag{50}$$

where $\mathbf{y}_p$ is the reference solution of sample point $\mathbf{x}_p$, $\mathbf{y}_{*,p}^{[K]}$ is the reference solution approximated in the truncated basis using any preprocessing method with K basis vectors. The error between the full basis ($K = N_{\exp} - 1$) and the reference solution is also called the *lack-of-richness error* [14].

By means of example, the error due to truncation of the surrogate model based on the zero centered snapshot matrix will become:

$$\kappa E_p^{[K]} = \mathbf{y}_p - \mathbf{y}_{0,p}^{[K]} = \mathbf{y}_p - \left(\bar{\mathbf{y}} + \boldsymbol{\Phi}_0^{[K]} \boldsymbol{\Phi}_0^{[K]^T} \left(\mathbf{y}_p - \bar{\mathbf{y}} \right) \right) \tag{51}$$

Total reconstruction error

The total reconstruction error is the difference between the reference solution and its interpolated approximation (Eq. (49)). The total reconstruction error of reference solution $\mathbf{y}_p$ approximated using a surrogate model with K basis vectors will be:

$$\boxed{\tau E_p^{[K]} = \mathbf{y}_p - \hat{\mathbf{y}}_*^{[K]}(\mathbf{x}_p)} \tag{52}$$

Again, by means of example, the total reconstruction error of validation sample $\mathbf{x}_p$ reconstructed using K basis vectors of the centered snapshot matrix will be:

$$\tau E_p^{[K]} = \mathbf{y}_p - \hat{\mathbf{y}}_0^{[K]}(\mathbf{x}_p) = \mathbf{y}_p - \left(\bar{\mathbf{y}} + \sum_{n=1}^{K} \boldsymbol{\varphi}_{n,0} \hat{\alpha}_{n,0}(\mathbf{x}_p) \right) \tag{53}$$

Interpolation error

The interpolation error is defined as the difference between the reference solution approximated using K basis vectors (Eq. 47), and the interpolated result vector (Eq. 49). As there is an interpolation error for every interpolated amplitude, with the addition of more basis vectors the interpolation error tends to increase. The interpolation error of validation sample $\mathbf{x}_p$ reconstructed using K basis vectors will be:

$$\boxed{\iota E_p^{[K]} = \mathbf{y}_{*,p}^{[K]} - \hat{\mathbf{y}}_*^{[K]}(\mathbf{x}_p) = \tau E_p^{[K]} - \kappa E_p^{[K]}} \tag{54}$$
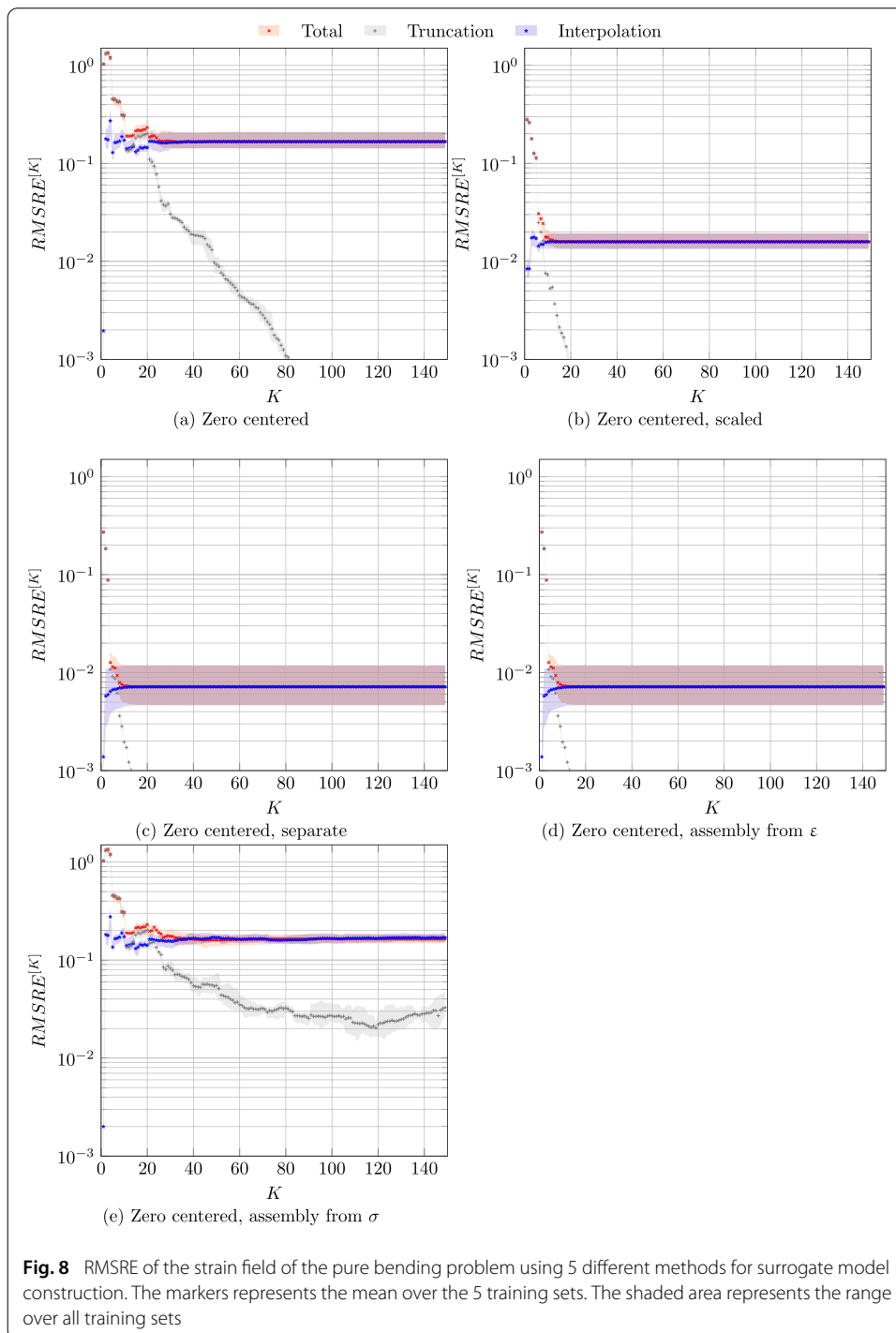
For example, the interpolation error using the zero centered snapshot matrix:

$$\begin{aligned} \iota E_p^{[K]} &= \mathbf{y}_{0,p}^{[K]} - \hat{\mathbf{y}}_0^{[K]}(\mathbf{x}_p) \\ &= \boldsymbol{\Phi}_0^{[K]} \boldsymbol{\Phi}_0^{[K]^T} (\mathbf{y}_p - \bar{\mathbf{y}}) - \sum_{n=1}^{K} \boldsymbol{\varphi}_{n,0} \hat{\alpha}_{n,0}(\mathbf{x}_p) \\ &= \boldsymbol{\Phi}_0^{[K]} \left(\boldsymbol{\Phi}_0^{[K]^T} (\mathbf{y}_p - \bar{\mathbf{y}}) - \hat{\alpha}_{n,0}(\mathbf{x}_p) \right) \end{aligned} \tag{55}$$
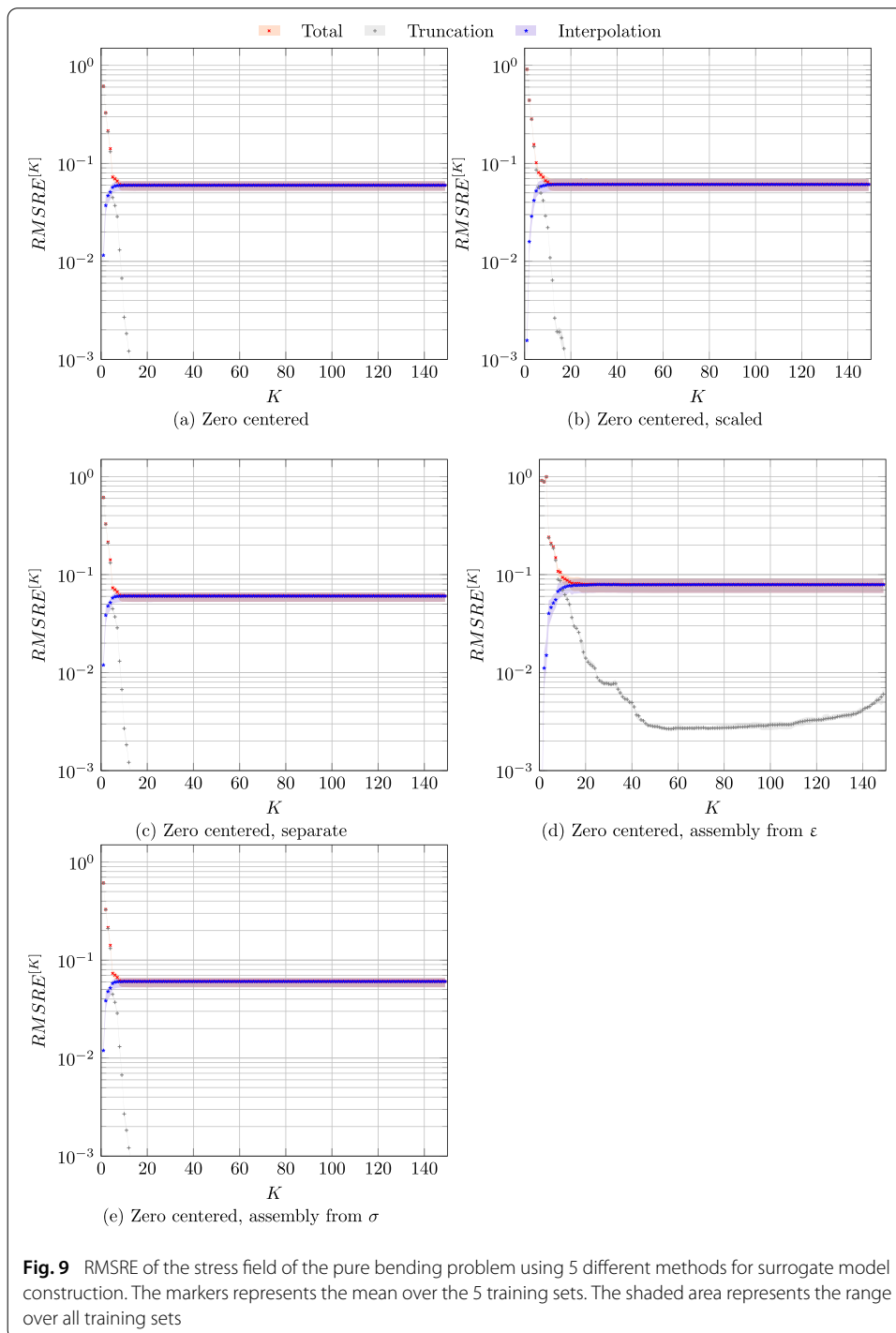
Error analysis considering physical parts

The Root Mean Square Relative Error (*RMSRE*) as proposed in Eq. (44) is used to quantify the different sources of error. Note that because in this case a validation set is used, the *RMSRE* is calculated over all $N_{\mathrm{val}} = 428$ and $N_{\mathrm{val}} = 81$ result vectors $\mathbf{y}_p$ for the pure and industrial bending problems respectively. Although Eq. (54) suggests that the total reconstruction *RMSRE* is simply the sum of the truncation and the interpolation *RMSRE*, this is not the case. As the calculation of the Euclidean norm of a vector is a non-linear operation, the total *RMSRE* will not be the sum of the interpolation error and the truncation error.

Fig. 8 RMSRE of the strain field of the pure bending problem using 5 different methods for surrogate model construction. The markers represents the mean over the 5 training sets. The shaded area represents the range over all training sets
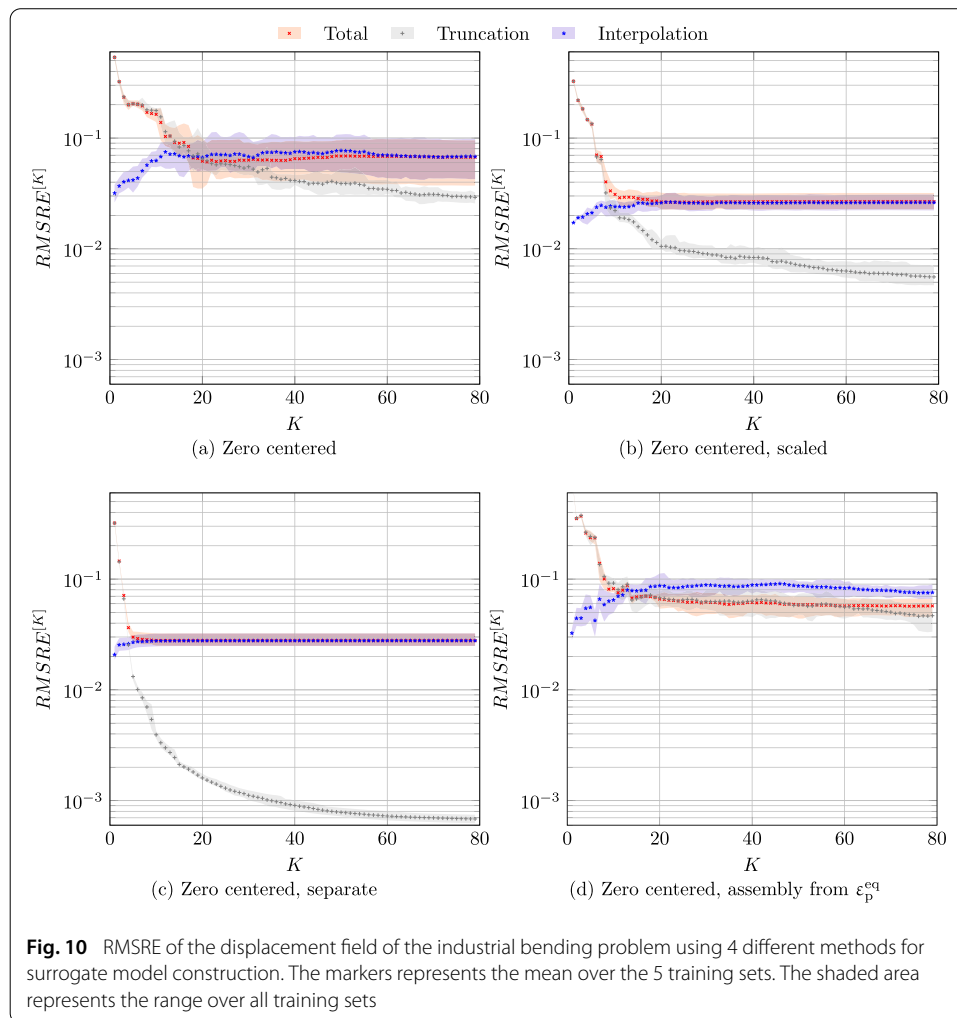
Figures. 8 and 9 display the *RMSRE* of the constructed surrogate models for the pure bending problem, based on the validation set, for strain and stress components respectively. Figs. 10, 11, and 12 display the *RMSRE* of the constructed surrogate models of the industrial bending problem, based on the validation set, for the displacement, equivalent plastic strain and stress components respectively. The markers represent the mean *RMSRE* over the 5 different training sets. The shaded area represents the minimum and maximum *RMSRE* found in the 5 different training sets. In Tables 5 and 6 the *RMSRE* of

Fig. 9 RMSRE of the stress field of the pure bending problem using 5 different methods for surrogate model construction. The markers represents the mean over the 5 training sets. The shaded area represents the range over all training sets

the different surrogate models using a basis that consists of all $K = N_{\exp} - 1$ basis vectors can be found.

The interpolation error as proposed in Eq. (54) can never be decreasing if an additional orthogonal basis vector is included in the surrogate model. With the addition of an orthogonal basis vector, the length of the vector describing the difference between the amplitude of the reference solution (Eq. 46) and the interpolated amplitude will stay the same or increase. This behavior can indeed be seen in Figs. 8 and 9c and Figs. 10, 11, 12c when

Fig. 10 RMSRE of the displacement field of the industrial bending problem using 4 different methods for surrogate model construction. The markers represents the mean over the 5 training sets. The shaded area represents the range over all training sets

separate snapshot matrices are used and thus only orthogonal basis vectors are added to the model. In all other surrogate models the interpolation error can decrease as the part of two basis vectors which correspond to one physical part are not orthogonal if the basis vector is determined based on the whole snapshot matrix.

When the snapshot matrices consisting of different physical parts are decomposed without taking into account the magnitudes of the physical parts, the components with larger magnitude will dominate the decomposition. Without scaling the data, the results for the stress are comparable for all preprocessing methods while the results for the displacement and strain are poorer. Overall, applying scaling to the snapshot matrix does not significantly deteriorates the performance of the surrogate models compared to separate decomposition.

When more basis vectors are included, the total and interpolation error stabilize around one value, because the interpolation error from previous modes dominates the error, making small improvements from additional vectors insignificant. Although in the work of Havinga [22] and Hamdaoui [5] different error measures are used, this flattening behavior can also be observed in their error plots. When the error has stabilized, adding more basis vectors will reduce the truncation error, but it will not result in a better performing surrogate model.
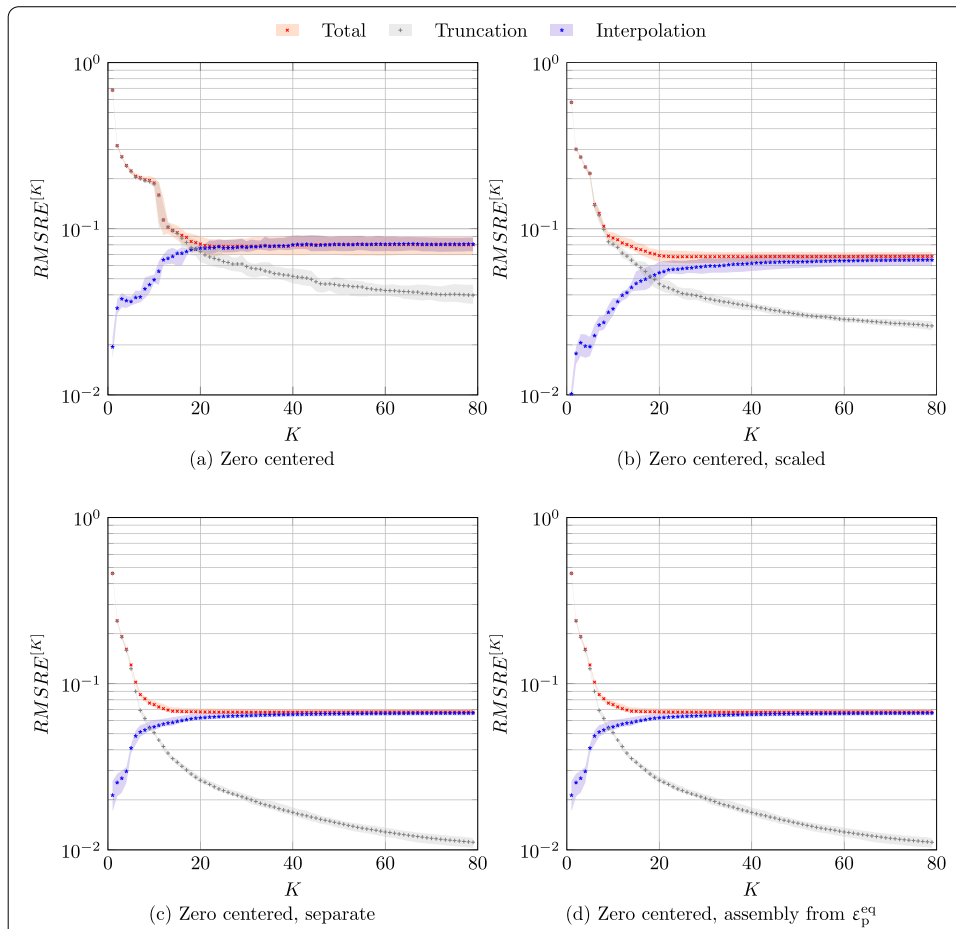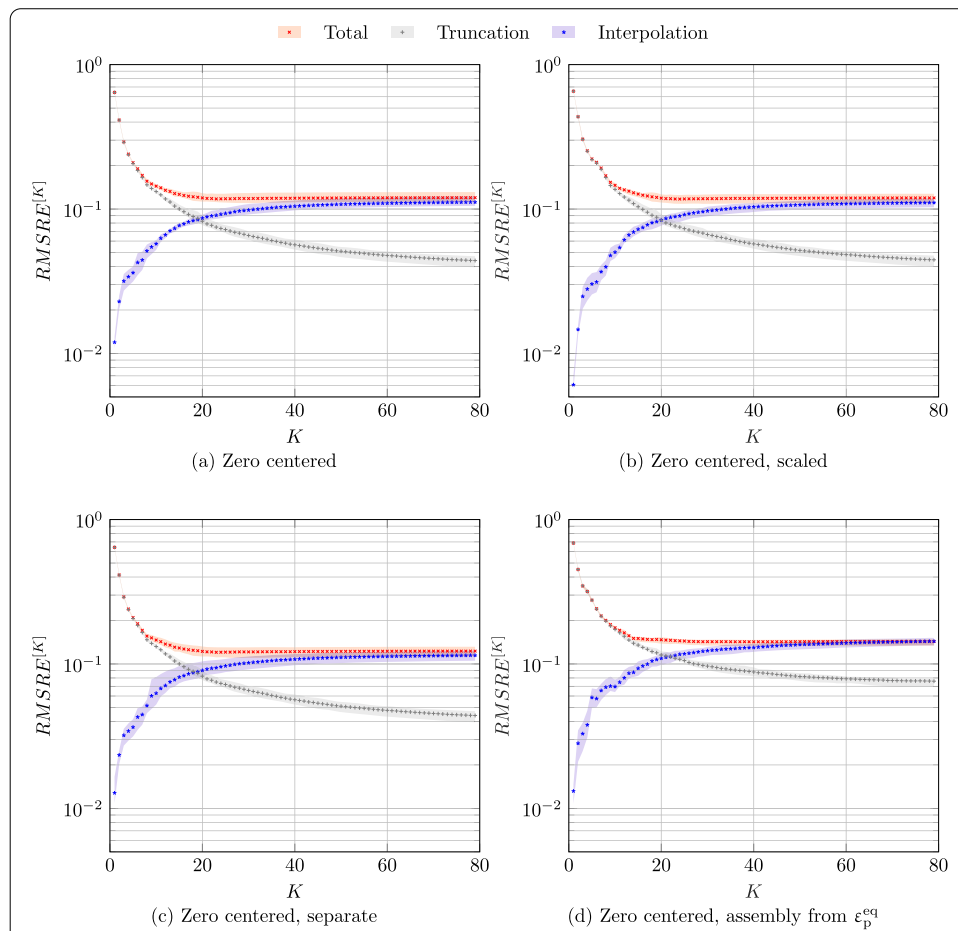
Fig. 11 RMSRE of the strain field of the industrial bending problem using 4 different methods for surrogate model construction. The markers represents the mean over the 5 training sets. The shaded area represents the range over all training sets

For the pure bending problem, the surrogate models that use separate decomposition or assembly from the strain perform best for the strain, as can be seen in Table 5. For the stress the surrogate model that is only zero centered performs best. For the industrial bending problem the surrogate model based on a separate snapshot matrices performs best for the equivalent plastic strain, while the surrogate model based on the scaled snapshot matrices performs best for the displacement and the stress, as can be seen in Table 6. A major advantage in scaling the snapshot matrix and decomposing all physical parts at once is in the fact that far less amplitude functions need to be interpolated. Hence one wins computational time without losing much accuracy.

Note however that the error from the scaled surrogate models of the industrial bending problem converge slower towards the stabilized error values. Hence more basis vectors will be needed to achieve the same performance compared to using separate snapshot matrices. Depending on the number of physical parts that need to be modelled, this slower convergence outweighs the computational costs of interpolating many more amplitudes.

We define the bandwidth as the difference between the minimum and maximum error found in the validation set results. A narrow bandwidth indicates that the results are robust and independent of the initial sampling in the training set, whereas a wider bandwidth

Fig. 12 RMSRE of the stress field of the industrial bending problem using 4 different methods for surrogate model construction. The markers represents the mean over the 5 training sets. The shaded area represents the range over all training sets

Table 5 Mean of the total reconstruction *RMSRE* ($\cdot 10^{-2}$) using a full basis ($K = 149$) over 5 training sets of the pure bending model

	Strain (ε)	Stress (σ)
Zero centered	16.676 (14.236− 20.854)	*5.980* (5.240− 6.584)
Zero centered, scaled	1.586 (1.340− 1.928)	6.104 (5.209− 6.951)
Zero centered, separate	*0.717* (0.465− 1.187)	6.052 (5.275− 6.583)
Zero centered, assembly from ε	*0.717* (0.465− 1.187)	7.904 (6.496− 9.138)
Zero centered, assembly from σ	16.410 (14.837− 18.275)	6.052 (5.275− 6.583)

Best performances are in italics, between brackets the minimum and maximum values over the five training sets are given

Table 6 Mean of the total reconstruction *RMSRE* ($\cdot 10^{-2}$) using a full basis ($K = 79$) over 5 training sets

	Displacement (u)	Equivalent plastic strain (ε_p^{eq})	Stress (σ)
Zero centered	6.71 (3.73− 9.55)	8.01 (6.93− 8.95)	12.00 (11.24− 13.10)
Zero centered, scaled	*2.67* (2.27− 3.19)	6.82 (6.38− 7.25)	*11.95* (11.05− 12.76)
Zero centered, separate	2.78 (2.49− 3.23)	*6.75* (6.53− 7.05)	12.28 (11.24− 13.16)
Zero centered, assembly from ε_p^{eq}	5.72 (4.58− 6.66)	*6.75* (6.53− 7.05)	14.32 (13.39− 14.81)

Best performances are in italics, between brackets the minimum and maximum values over the five training sets are given

indicates a stronger dependency to the randomly generated initial training sets. For both the pure bending problem and the industrial bending problem, the stress part has the most narrow bandwidth independent of the preprocessing method, which implies a more robust performance that is less dependent of the approach in surrogate model construction. For the pure bending problem, the surrogate model that uses a separate snapshot matrix for the strain or assembly from the strain has the widest bandwidth for the strain (Fig. 8c, d). For the industrial bending problem using separate snapshot matrices results in the smallest bandwidth for all physical parts. The bandwidth of the displacement field that is modelled using a zero centered snapshot matrix is the widest (Fig. 10a).

When the basis vectors are assembled from another physical part, the part that was dominating the decomposition deteriorates. For both the pure bending problem and the industrial bending problem, the RMSRE in the stress is larger when the basis vectors are assembled from the equivalent plastic strain compared to when a zero centered snapshot matrix is used. However, for the industrial bending problem when the basis vectors are assembled from the equivalent plastic strain, the RMSRE of the displacement field decreases with respect to the RMSRE from the zero centered snapshot matrix. Similarly, for the pure bending problem, when the basis vectors are assembled from the stress, the RMSRE of the strain field decreases with respect to the RMSRE from the zero centered snapshot matrix.

In the pure bending problem, the truncation error of the basis with basis vectors assembled from another physical part increases in the higher order basis vectors (Figs. 8e and 9d). This behaviour is a numerical artefact of the novel method. The basis vectors are sorted based on the variance in the data of the main physical part. The lower basis vectors describe the most variation in the data of the main physical part, while the higher basis vectors describe less variation in the data or even noise. The assembled basis vectors including the data from the secondary physical part are not orthonormal, allowing for larger fractions of the variation in the data of the secondary physical part to appear in the higher order basis vectors. These higher order basis vectors become unbalanced and contain relatively high values for the secondary physical parts, leading to an increase in the truncation RMSRE, despite the low amplitudes for these higher modes. This behaviour does however only seem to become significant for the highest modes, and may therefore not be relevant for practical applications of the method.

Comparing Figs. 8, 9 with Fig. 4 and Figs. 10, 11, 12 with Fig. 5 it can be concluded that the ratio based on the eigenvalues as in Eq. (40) does not give a good indication of how many basis vectors should be included to reach a certain accuracy, nor does it indicate when the error due to interpolation and the total error stabilize. At that specific K value adding additional basis vectors does not improve the surrogate model. Based on the ratio of the eigenvalues the error analysis is far more optimistic than the final performance.

Conclusion

In this work four different approaches for interpolation of field data from FE simulations are presented. Two of these approaches are based on different preprocessing methods of the snapshot matrix and two approaches are based on decomposition of different parts of the snapshot matrix. Based on 5 training sets, 5 surrogate models are constructed using

each approach for two different bending problems. The different approaches are analyzed based on a validation set.

When different physical fields are modelled, scaling the data improves the quality of the surrogate model in most physical parts of the snapshot matrix compared to only zero centering the data. In this study the surrogate model based on the scaled snapshot matrix of the industrial bending problem performs best for the displacement field and stress field. For both demonstrator problems, separate decomposition or assembly from the strain performs best for the strain.

Taking separate snapshot matrices has the advantage that one can choose per physical quantity how many basis vectors should be included. The disadvantage is that the basis vectors of each physical quantity must be interpolated independently, whereas decomposing data with different physical quantities together, only requires the construction of a single interpolation model per joint basis vector. Consequently, for the examples presented in this work, two or three times as many interpolation models must be fitted when separate decompositions with the same number of basis vectors are used. More amplitudes need to be interpolated without decreasing the error drastically. All surrogate models are based on the same training sets, hence the computational times for expensive model evaluations are the same. However, as there are two or three different physical parts, the computational costs of fitting the interpolation functions for the separate snapshot matrices are two or three times as high. When advanced interpolation algorithms are used, these computational costs can become significant compared to doing the expensive model evaluations. When different physical parts need to be modelled it is therefore recommended to use one snapshot matrix and scale the data per physical part, as it has the advantage of less computational cost in fitting the interpolation model and in evaluating the reduced order model, without a significant reduction in model accuracy. The novel approach, assembly from one physical field, can be used to assemble physical quantities based on their covariance with the physical quantity that is considered as the main modelling objective. Although there is a loss in accuracy in the physical fields that are assembled from the main physical quantity, the novel approach provides a cheap method to model different physical fields.

Declarations

Competing interests
The authors declare that they have no competing interests.

References
1.   Simpson TW, Poplinski JD, Koch PN, Allen JK. Metamodels for computer-based engineering design: survey and recommendations. Eng Computers. 2001;17(2):129–50. https://doi.org/10.1007/PL00007198.

2. Jin R, Chen W, Simpson TW. Comparative studies of metamodelling techniques under multiple modelling criteria. Struct Multidiscip Optim. 2001;23(1):1–13. https://doi.org/10.1007/s00158-001-0160-4.

3. Forrester AIJ, Sobester A, Keane AJ. Engineering Design via Surrogate Modelling: A practical guide. In: Engineering Design Via Surrogate Modelling, pp. 33–76, 2008. Constructing a Surrogate. https://doi.org/10.1002/9780470770801.ch2

4. Wiebenga JH, van den Boogaard AH, Klaseboer G. Sequential robust optimization of a V-bending process using numerical simulations. Struct Multidiscip Optim. 2012;46(1):137–53. https://doi.org/10.1007/s00158-012-0761-0.

5. Hamdaoui M, Le Quilliec G, Breitkopf P, Villon P. POD surrogates for real-time multi-parametric sheet metal forming problems. Int J Mater Form. 2014;7(3):337–58. https://doi.org/10.1007/s12289-013-1132-0.

6. Iuliano E, Quagliarella D. Proper orthogonal decomposition, surrogate modelling and evolutionary optimization in aerodynamic design. Comput Fluids. 2013;84:327–50. https://doi.org/10.1016/j.compfluid.2013.06.007.

7. Liang YC, Lee HP, Lim SP, Lin WZ, Lee KH, Wu CG. Proper orthogonal decomposition and its applications - part I: theory. J Sound Vibrat. 2002;252(3):527–44. https://doi.org/10.1006/jsvi.2001.4041.

8. Bialecki RA, Kassab AJ, Fic A. Proper orthogonal decomposition and modal analysis for acceleration of transient fem thermal analysis. Int J Numer Methods Eng. 2005;62(6):774–97. https://doi.org/10.1002/nme.1205.

9. Chinesta F, Keunings R, Leygue A. The proper generalized decomposition for advanced numerical simulations. Springer Briefs Appl Sci Technol. 2014. https://doi.org/10.1007/978-3-319-02865-1.

10. Radermacher A, Reese S. Model reduction in elastoplasticity: Proper orthogonal decomposition combined with adaptive sub-structuring. Comput Mech. 2014;54(3):677–87. https://doi.org/10.1007/s00466-014-1020-6.

11. Xiao M, Breitkopf P, Coelho RF, Villon P, Zhang W. Proper orthogonal decomposition with high number of linear constraints for aerodynamical shape optimization. Appl Math Comput. 2014;247:1096–112. https://doi.org/10.1016/j.amc.2014.09.068.

12. Amsallem D, Zahr M, Choi Y, Farhat C. Design optimization using hyper-reduced-order models. Struct Multidiscip Optim. 2015;51(4):919–40. https://doi.org/10.1007/s00158-014-1183-y.

13. Phalippou P, Bouabdallah S, Breitkopf P, Villon P, Zarroug M. Sparse pod modal subsets for reduced-order nonlinear explicit dynamics. Int J Numer Methods Eng. 2020;121(4):763–77. https://doi.org/10.1002/nme.6243.

14. Guénot M, Lepot I, Sainvitu C, Goblet J, Filomeno Coelho R. Adaptive sampling strategies for non-intrusive POD-based surrogates. Eng Comput. 2013;30(4):521–47. https://doi.org/10.1108/02644401311329352.

15. Benamara T, Breitkopf P, Lepot I, Sainvitu C. Multi-fidelity extension to non-intrusive proper orthogonal decomposition based surrogates. In: Proceedings of VII European Congress on Computational Methods in Applied Sciences and Engineering; 2016. p. 5–10.

16. Chatterjee A. An introduction to the proper orthogonal decomposition. Curr Sci. 2000;78(7):808–17. https://doi.org/10.1109/LPT.2009.2020494.

17. Abdi H, Williams LJ. Principal component analysis. Wiley Iinterdiscip Rev Comput Stat. 2010;2(4):433–70. https://doi.org/10.1002/wics.101.

18. Zhao D, Xue D. A comparative study of metamodeling methods considering sample quality merits. Struct Multidiscip Optim. 2010;42(6):923–38. https://doi.org/10.1007/s00158-010-0529-3.

19. Kianifar MR, Campean F. Performance evaluation of metamodelling methods for engineering problems: towards a practitioner guide. Struct Multidiscip Optim. 2020;61(1):159–86. https://doi.org/10.1007/s00158-019-02352-1.

20. Khaledi K, Miro S, König M, Schanz T. Robust and reliable metamodels for mechanized tunnel simulations. Computers Geotech. 2014;61:1–12. https://doi.org/10.1016/j.compgeo.2014.04.005.

21. Bocciarelli M, Buljak V, Moy CKS, Ringer SP, Ranzi G. An inverse analysis approach based on a POD direct model for the mechanical characterization of metallic materials. Comput Mater Sci. 2014;95:302–8. https://doi.org/10.1016/j.commatsci.2014.07.025.

22. Havinga J, Mandal PK, van den Boogaard AH. Exploiting data in smart factories: real-time state estimation and model improvement in metal forming mass production. Int J Mater Forming. 2020;13(5):663–73. https://doi.org/10.1007/s12289-019-01495-2.

23. Dang VT, Labergere C, Lafon P. POD surrogate models using adaptive sampling space parameters for springback optimization in sheet metal forming. Procedia Engineering. 2017;207:1588–93. https://doi.org/10.1016/j.proeng.2017.10.1053. International Conference on the Technology of Plasticity, ICTP 2017, 17-22 September 2017, Cambridge, United Kingdom.

24. Kato H, Funazaki K-i. Pod-driven adaptive sampling for efficient surrogate modeling and its application to supersonic turbine optimization. In: Proceedings of ASME Turbo Expo 2014: Turbine Technical Conference and Exposition, 2014; p. 1–10. https://doi.org/10.1115/GT2014-27229.

25. Braconnier T, Ferrier M, Jouhaud JC, Montagnac M, Sagaut P. Towards an adaptive POD/SVD surrogate model for aeronautic design. Comput Fluids. 2011;40(1):195–209. https://doi.org/10.1016/j.compfluid.2010.09.002.

26. Bolzon G, Buljak V, Maier G. Proper orthogonal decomposition in direct and inverse elastic-plastic analysis. In: X International Conference on Computational Plasticity, 2009; p. 1–4. http://congress.cimne.com/complas09/proceedings/pdfs/p542.pdf.

27. Khaledi K, Schanz T, Miro S. Application of metamodelling techniques for mechanized tunnel simulation. J Theor Appl Mech. 2014;44(1):45–54. https://doi.org/10.2478/jtam-2014-0003.

28. Buljak V. Proper Orthogonal Decomposition and Radial Basis Functions Algorithm for Diagnostic Procedure Based on Inverse Analysis. FME Trans. 2010;38(3):129–36.

29. Misiun G, Wang C, Geijselaers HJM, van den Boogaard AH. Interpolation of final geometry and result fields in process parameter space. In: MATEC Web of Conferences, vol. 80, 2016;pp. 503–510. https://doi.org/10.1051/matecconf/20168013006

30. Steffes-lai D. Approximation Methods for High Dimensional Simulation Results - Parameter Sensitivity Analysis and Propagation of Variations for Process Chains. PhD thesis, University of Cologne, 2014. https://books.google.nl/books/about/Approximation_Methods_for_High_Dimension.html?id=xD4LBAAAQBAJ&source=kp_cover&redir_esc=y.

31. Buljak V, Maier Giulio G. Proper Orthogonal Decomposition and Radial Basis Functions in material characterization based on instrumented indentation. Eng Struct. 2011;33(2):492–501. https://doi.org/10.1016/j.engstruct.2010.11.006.

32. Zimmermann R, Görtz S. Non-linear reduced order models for steady aerodynamics. Procedia Computer Sci. 2010;1(1):165–74. https://doi.org/10.1016/j.procs.2010.04.019.
33. Skillicorn D. Understanding Complex Datasets: Data Mining with Matrix Decompositions (Chapman & Hall/Crc Data Mining and Knowledge Discovery Series). Chapman & Hall/CRC, New York; 2007. https://doi.org/10.1201/9781584888338. http://www.amazon.ca/exec/obidos/redirect?tag=citeulike09-20&path=ASIN/1584888326.
34. Jolliffe IT. Principal Component Analysis: Second Edition. 2002;30:487.
35. Cadima J, Jolliffe IT. On relationships between uncentred and column-centred principal component analysis. Pak J Stat. 2009;25(4):473–503. https://doi.org/10.1016/B978-1-85617-912-6.00003-7.
36. de Gooijer BM, Hazrati J, Geijselaers HJM, van den Boogaard AH. On the choice of basis in proper orthogonal decomposition-based surrogate models. In: AIP Conference Proceedings, 2019;vol. 2113. https://doi.org/10.1063/1.5112635.
37. Phalippou P, Bouabdallah S, Breitkopf P, Villon P, Zarroug M. 'On-the-fly' snapshots selection for Proper Orthogonal Decomposition with application to nonlinear dynamics. Comput Methods Appl Mech Eng. 2020;367:113120. https://doi.org/10.1016/j.cma.2020.113120.
38. Jolliffe IT, Morgan BJT. Principal component analysis and exploratory factor analysis. Stat Methods Med Res. 1992;1(1):69–95. https://doi.org/10.1177/096228029200100105.
39. Abdi H. The Eigen-decomposition: Eigenvalues and Eigenvectors. In: Encyclopedia of measurements and statistics; 2007. p. 907–12. https://doi.org/10.4135/9781412952644.n413.
40. Golub GH, Van Loan CF. Matrix Computations. Baltimore: The Johns Hopkins University Press; 2013.
41. Abdi H. The Eigen-decomposition: Eigenvalues and Eigenvectors. In: Encyclopedia of measurements and statistics; 2007. p. 309–10. https://doi.org/10.4135/9781412952644.n150.
42. Franke R. Scattered Data Interpolation: Tests of Some Method. Math Comput. 1982;38(157):181. https://doi.org/10.2307/2007474.
43. Hamim SU, Singh RP. Taguchi-based design of experiments in training POD-RBF surrogate model for inverse material modelling using nanoindentation. Inverse Probl Sci Eng. 2017;25(3):363–81. https://doi.org/10.1080/17415977.2016.1161036.
44. Rippa S. An algorithm for selecting a good value for the parameter c in radial basis function interpolation. Adv Comput Math. 1999;11:193–210. https://doi.org/10.1023/A:1018975909870.
45. Buhmann MD. Radial Basis Functions: Theory and Implementations. Cambridge: Cambridge Monographs on Applied and Computational Mathematics. Cambridge University Press; 2003. https://doi.org/10.1017/CBO9780511543241.
46. Amouzgar K, Strömberg N. Radial basis functions as surrogate models with a priori bias in comparison with a posteriori bias. Struct Multidiscip Optim. 2017;55(4):1453–69. https://doi.org/10.1007/s00158-016-1569-0.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Submit your manuscript to a SpringerOpen journal and benefit from:

► Convenient online submission
► Rigorous peer review
► Open access: articles freely available online
► High visibility within the field
► Retaining the copyright to your article

Submit your next manuscript at ► springeropen.com