Cao *et al. Adv. Model. and Simul. in Eng. Sci.(2016)3:5*
DOI 10.1186/s40323-016-0057-9

Advanced Modeling and Simulation in Engineering Sciences
a SpringerOpen Journal

RESEARCH ARTICLE Open Access

A hybrid RNN-GPOD surrogate model for real-time settlement predictions in mechanised tunnelling

Ba-Trung Cao[1*], Steffen Freitag[1,2] and Günther Meschke[1]

*Correspondence:
ba.cao@rub.de
[1]Institute for Structural Mechanics, Ruhr University Bochum, Universitätstr. 150, 44801 Bochum, Germany
Full list of author information is available at the end of the article

Abstract

Realistic 3D simulations of the tunnelling process are increasingly required to investigate the interactions between machine-driven tunnel construction and the surrounding soil in order to provide reliable estimates of the expected settlements and associated risks of damage for existing structures, in particular in urban tunnelling projects. To accomplish the step from large-scale computational analysis to real-time predictions of expected settlements during tunnel construction, the focus of this paper is laid on the generation of a numerically efficient hybrid surrogate modelling strategy, combining Gappy proper orthogonal decomposition (GPOD) and recurrent neural networks (RNN). In this hybrid RNN-GPOD surrogate model, the RNN is employed to extrapolate the time variant settlements at several monitoring points within an investigated surface area and GPOD is utilised to predict the whole field of surface settlements based on the RNN predictions and a POD radial basis functions approximation. Both parts of the surrogate model are created based on results of finite element simulations from geotechnical and process parameters varied within the range of intervals given in the design stage of a tunnel project. In the construction stage, the hybrid surrogate model is applied for real-time reliability analyses of the mechanised tunnelling process to support the machine operator in steering the tunnel boring machine.

Keywords: Proper orthogonal decomposition, Recurrent neural network, Mechanised tunnelling, Surrogate model, Real-time application, Computational steering

Background

Mechanised tunnelling is a widely used construction method for underground infrastructure in particular in urban areas due to its effectiveness in controlling the advancement process and to limit the construction induced ground deformations. This construction method is characterised by a staged procedure of soil excavation at the tunnel face and lining erection, providing at the same time a continuous support of the soil by means of supporting fluids at the tunnel face and pressurised grouting of the tail gap. The interactions between the tunnel boring machine (TBM), the support measures and the soil, including the groundwater, are the determining factors for the efficiency, the safety of the tunnel advancement and the risk of damage on the existing infrastructure. Evidently,

these interactions are complex and require computational simulation methods based upon sufficiently 3D realistic numerical models for the individual components and their interactions. However, for a sufficient degree of spatial resolution of this highly nonlinear, time dependent problem, adequate simulation models are characterized by a large number of unknowns in the range of 10^6–10^8 unknowns.

Up to date, numerical simulations in tunnelling are restricted to the design stage of a project. The support of the steering of TBMs during construction mainly relies upon the interpretation of monitoring data during the construction by experienced experts. If results from large scale numerical models are to be used during construction to provide additional information on the potential consequences of decisions taken for the steering of the TBM (e.g., the surface settlement field ahead of the tunnel face), a real-time system response is required. However, performing large scale, computationally expensive models on site is unrealistic for real-time applications, which demand obtaining the system response in the range of seconds to minutes. To accomplish the step from large-scale computational analysis to real-time predictions of expected settlements during tunnel construction, model reduction strategies are required to substitute the original numerical model by surrogate models. Recently, the authors have proposed a hybrid surrogate model [1], which employs a combination of two different techniques: recurrent neural network (RNN) and proper orthogonal decomposition (POD), aiming to exploit the advantages of both methods. While RNNs are well suited for predictions using extrapolation of data (see [2]), the POD is able to deal with high dimensional outputs, see, e.g., [3,4].

For the specific problem at hand, two different variants of the POD method have been employed in [1]. The first one is the POD with interpolation to approximate surface settlements from previous stages of the tunnel excavation (excavation steps 1 to n). Frequently, radial basis functions (RBF) are considered as interpolation functions, resulting in the POD-RBF approach, see e.g., [3–5]. The second approach is the Gappy POD (GPOD) approach, first proposed in [6] and employed to tackle the missing data problem in image processing, or to reconstruct human faces from partial data. In [1], this approach has been utilised to approximate the field of settlements from data available from a limited number of monitoring points in step $n + 1$.

To allow for a monitoring and model based support of TBM steering, supplying information in real-time to the TBM driver on the expected (i.e., future) settlements, a model reduction technique which allows extrapolation in time with multiple outputs is proposed in this paper. To this end, the hybrid surrogate modelling strategy originally proposed in [1] is extended. More precisely, strategies to improve the predictions of hybrid surrogate models by means of problem specific enhancements for their components, POD-RBF and GPOD, are proposed.

For the POD using interpolation functions, the concept of RBF is replaced by an extension called Extended RBF (ERBF) [7]. This concept is capable of improving the prediction capability of the surrogate model by combining the effectiveness of the RBF and the flexibility of non-RBF approaches. This leads to a POD-ERBF network which has been shown to produce better prediction results as compared to the POD-RBF, see [8]. In association with the POD-ERBF algorithm, the reconstruction accuracy of the GPOD approach can be enhanced by adopting iterative schemes (IGPOD) to derive the POD basis, e.g., [9–12]. In the GPOD approach, an incomplete data vector is reconstructed based on the assumption that the vector inherits the characteristics of the known data set. The reconstruction

procedure is performed in one step based on the POD modes of the data set, e.g., [6,9]. A more precise approach is to include this partial missing vector into the known data set and to iterate until convergence. Everson and Sirovich [6] have performed an iterative approach based on an initial guess, which is denoted as the E–S method. The E–S method depends on the initial guess and therefore may lead to less accurate solutions because it only employs a certain number of POD modes in the reconstruction procedure. Based on the E–S method, Venturi and Karniadakis [11] proposed another extension, referred to as the V–K method, which is not dependent on the initial guess and improves the accuracy significantly by gradually increasing the number of POD modes. However, using the V–K method leads to a significant increase in computation time. Therefore, in order to replace classical GPOD, the IGPOD, based upon both the E–S and the V–K scheme is investigated in this paper in regards to prediction performance and computation time.

The remainder of this paper is organised as follows. "Numerical model for mechanised tunnelling" gives a brief description of the finite element model for process-oriented numerical simulations of TBM driven tunnelling. "Hybrid surrogate model" first summarises the algorithmic scheme suggested for the hybrid surrogate model and subsequently provides a synopsis of methods employed in the hybrid surrogate model: RNN, POD, POD-RBF, POD-ERBF, GPOD and IGPOD (E–S and V–K methods). "Application to mechanised tunnelling" is devoted to applications of the proposed surrogate model in mechanised tunnelling. Based upon the developed surrogate model, a software for real-time prognosis in TBM tunnelling is finally proposed in "Real-time simulation software for TBM steering support" to enable supporting decisions during mechanised tunnel construction.

Numerical model for mechanised tunnelling

A 3D numerical model *ekate* (enhanced KRATOS for advanced tunnelling engineering) developed specifically for process-oriented computational simulations of shield tunnelling processes at the Institute for Structural Mechanics of Ruhr University Bochum, see e.g., [13], is used in this paper. The model, which is conceptually based upon a previous model proposed by [14], has been supplemented with an automatic modeller [15], using a pre-processing procedure with a high degree of automation to generate complex 3D finite element (FE) models with small effort. The model is capable of simulating the excavation and construction procedure involved in mechanised tunnelling in soft, water saturated soils (i.e., the TBM advancement, continuous soil support, ring-wise installation of the linings and the related interactions of the TBM advance with the soil and the groundwater).

The soil is modelled as a three (or two) phase material for partially (or fully) saturated soils. Two elastoplastic models are available for the consideration of the inelastic response of soft soils: the clay and sand model and the Drucker-Prager model (Fig. 1(1)). The TBM is represented as a deformable body (Fig. 1(2)) moving through the soil with frictional surface-to-surface contact along the shield skin, allowing that the deformation of the soil naturally follows the real, tapered geometry of the TBM and that the effect of overcutting is captured. In order to realistically model the movement of the TBM and its interaction with the soil, an algorithm to control the individual jack thrusts is used to keep the TBM on the designed alignment path. After each TBM advance, the excavation at the cutting face,

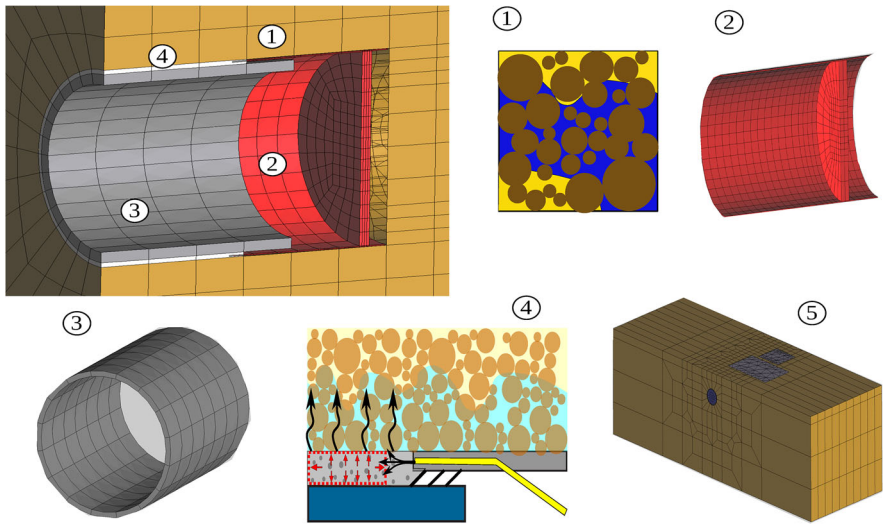

Fig. 1 FE model for shield tunnelling *ekate* with respective components. (*1*) surrounding soil, (*2*) TBM, (*3*) tunnel lining, (*4*) pressurisation of time-dependent grouting material and (*5*) representation of existing buildings

the tail void grouting and the erection of a new lining ring during standstill are taken into account by re-zoning the finite element mesh and adjusting all boundary conditions to the new situation. The tail gap grouting (see Fig. 1(4)) is modelled as a fully saturated two-phase material with a hydrating matrix phase, considering the temporal evolution of stiffness and permeability of the cementitious grout. In the simulation model *ekate*, buildings are represented by substitute models with appropriate stiffness and thickness to consider the soil-structure interaction (see Fig. 1(5)). In the presented FE formulation, isotropic shell, or, alternatively, volume elements with respective properties are adopted interacting with the soil through a mesh independent surface-to-surface contact algorithm, which prevents the penetration of the foundation of the building into the soil. For more details on the numerical model *ekate* and its ability to predict tunnelling induced settlements in the context of different applications we refer to [16–18].

For real-time applications during tunnel construction, the FE model must be substituted by a fast surrogate model which is able to approximate the physical behaviour of all relevant components involved in the simulation model for mechanised tunnelling. The surrogate model is developed to predict time variant settlement fields, which is realised by a mapping of time constant and time variant low dimensional inputs onto time variant high dimensional outputs. For this mapping, a hybrid surrogate modelling strategy is proposed in the next section.

Hybrid surrogate model

In this section, a hybrid surrogate model is proposed, which combines the concepts of artificial neural networks and POD for the prediction of surface settlements in TBM driven tunnel constructions. A RNN and the GPOD approach are combined to obtain a hybrid RNN-GPOD surrogate model to substitute the FE model for real-time applications. The hybrid surrogate model is generated offline, during the design stage of a tunnel project, and applied online, during the construction stage, to predict the time variant settlement field.

Figure 2 contains a schematic illustration of the real-time prognosis scheme including the offline and online stages. The process-oriented FE model *ekate* described in the previous section is used to simulate the mechanised tunnelling processes with varying model and process parameters. The objective of the surrogate modelling is to predict the complete spatio-temporal field of surface settlements for further excavation steps for an arbitrary set of input parameters with similar accuracy as the original FE model. The input parameters (a mix of time variant process parameters, e.g., excavation rate, support and grouting pressures, and time constant geotechnical parameters, e.g., material parameters of the soil layers) are selected based on sensitivity analyses. In the context of real-time applications this implies, that firstly, the POD-RBF network is employed to approximate the spatial field of surface displacements from time step 1 to the current time step n with the actual operational parameters obtained directly from the construction site, i.e., from the control unit of the TBM. Meanwhile, for the generation of user defined steering scenarios, the surface displacements of selected monitoring points for the subsequent time step n+1 are predicted by utilising a RNN. Finally, the GPOD approach is adopted to predict the complete field of tunnelling induced surface settlements for the set of user defined steering scenarios based on a combination of the results from the two previous methods.

The predicted results are used to support the machine driver in selecting the steering parameters for the next excavation step $n + 1$. Then these data, i.e., the finally selected input parameters and the corresponding settlement predictions for time step $n + 1$ are included into the available data set and the procedure is repeated for the subsequent excavation steps. The algorithm of the proposed RNN-GPOD approach is summarised in Table 1.

It should be noted, that also monitoring data can be used to update the surrogate model, see [19]. However, the focus of this paper is on improving the prediction quality of the POD-RBF and the GPOD components of the hybrid surrogate model. The individual

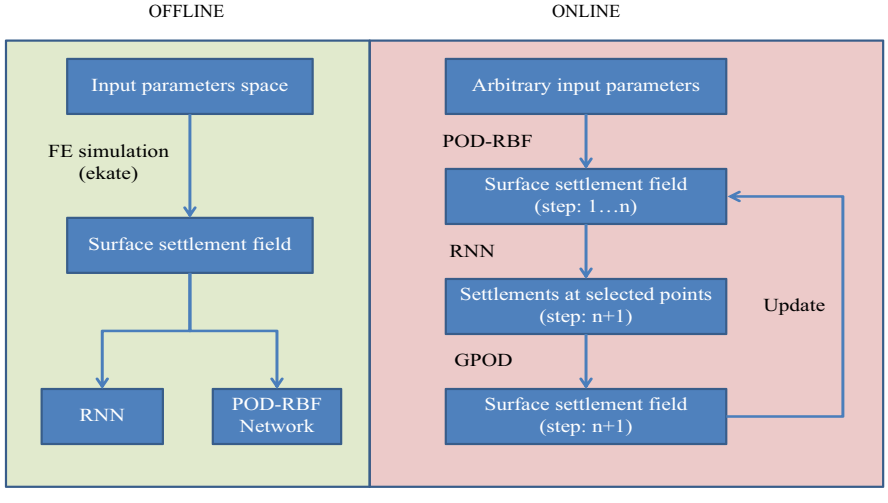

Fig. 2 Hybrid surrogate model scheme. Offline stage: performing FE simulations for a predefined range of varying input parameters, training of the surrogate model; online stage: generation of settlement field prediction

Table 1 RNN-GPOD approach to predict the complete field of surface displacements

Offline stage

1. Generate a numerical model representing the tunnelling process within a selected section of the project from time step 1 to *n*
2. Define investigated input parameters and the corresponding range of values
3. Run numerical simulations with different input parameters
4. Store the numerical results of displacements of the complete surface area of the analysis section
5. Provide data of selected monitoring points for the training of a RNN
6. Provide data of the complete displacements field for training of the POD-RBF

Online stage

1. Input: an arbitrary set of input parameters
2. Approximate the complete displacement field from time step 1 to *n* (by POD-RBF)
3. Predict the displacements of selected monitoring points for the next time step *n*+1 (by RNN)
4. Predict the displacements field for the next time step *n*+1 (by GPOD)
5. Update the complete displacements field from time step 1 to *n*+1
6. Repeat steps 3–5

components of the hybrid RNN-GPOD surrogate model are explained below, highlighting the proposed extensions and improvements.

Recurrent neural network

Surrogate models based on artificial neural networks are widely used in civil engineering and structural mechanics, see e.g., [20] and [21] for an overview. RNN have been proposed, see e.g., [2], for the approximation of dependencies between structural processes, i.e., to predict dependencies between time variant structural actions such as mechanical and thermal loadings and time variant structural responses, e.g., displacements and structural stresses. RNNs have two advantages which are exploited in the proposed hybrid surrogate model: (1) they are able to learn dependencies between data series without considering time as an additional input parameter and (2) they allow for extrapolations to predict the future structural response. The basic architecture of RNNs is similar to feed forward neural networks. But context neurons are added to the hidden and output neurons in order to consider history dependencies in structural processes, see Fig. 3.

In each time step n, the $j = 1, \ldots, J$ inputs (time constant geotechnical parameters and time variant tunnelling process parameters) ${}^{[n]}x_j$ of the RNN are processed layer by layer through the network to obtain the outputs (surface settlements at selected monitoring points) ${}^{[n]}s_k$. The signals of neuron i in layer $m = 2, \ldots, M$ are computed by

$$ {}^{[n]}x_i^{(m)} = \varphi_i^{(m)} \left(\sum_{h=1}^{H} \left[{}^{[n]}x_h^{(m-1)} \cdot w_{ih}^{(m)} \right] + \sum_{q=1}^{Q} \left[\sum_{d=1}^{D} \left[{}^{[n-d]}x_q^{(m)} \cdot {}^{d}c_{iq}^{(m)} \right] \right] + b_i^{(m)} \right). \quad (1) $$

Different types of activation functions $\varphi_i^{(m)}(.)$ (e.g., hyperbolic tangent function, logistic sigmoid function, area hyperbolic sine function, linear function) can be used to process the input signals to the corresponding output signal of neuron i. In Eq. (1), $x_h^{(m-1)}$ are the output signals of the $h = 1, \ldots, H$ neurons in the previous layer $(m-1)$, which are multiplied by the synaptic weights $w_{ih}^{(m)}$, ${}^{[n-d]}x_q^{(m)}$ are the $[n-1], \ldots, [n-D]$ delayed prior output signals of the $q = 1, \ldots, Q$ neurons in the same layer (m), which are multiplied by the context weights ${}^{d}c_{iq}^{(m)}$, and $b_i^{(m)}$ is a bias value of neuron i. It should be noted, that the output signals of the neurons in the first layer $(m = 1)$ are identical to the input signals

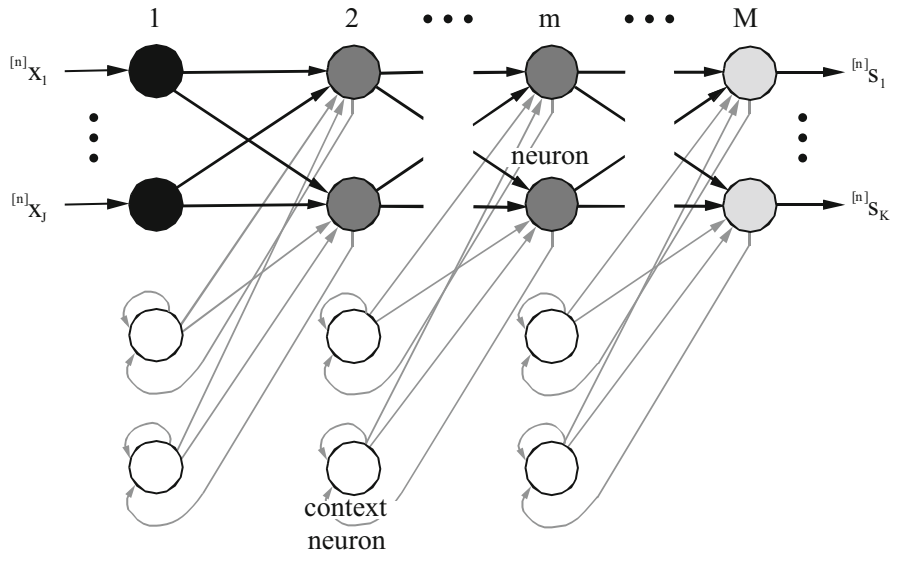

Fig. 3 Recurrent neural network. Prediction of settlements $^{[n]}s_1$ to $^{[n]}s_K$ (at K monitoring points) in time step [n] influenced by the input parameters $^{[n]}x_1$ to $^{[n]}x_J$ of time step [n] and prior values of the input parameters [of time steps (1,…,n)] considered by context neurons

of the RNN ($^{[n]}x_j^{(1)} = {}^{[n]}x_j$) and that the output signals of the neurons in the output layer ($m = M$) are identical to the output signals of the RNN ($^{[n]}x_k^{(M)} = {}^{[n]}s_k$), see Fig. 3. For the prediction of the current structural response considering the complete structural history, a delay d of one or more time steps can be used to compute the network signals, see [21] for more details. The free network parameters (i.e., the synaptic weights $w_{ih}^{(m)}$, the context weights $^{d}c_{iq}^{(m)}$ for each $d = 1,\ldots,D$ delayed time step and the bias values $b_i^{(m)}$) can be adjusted by a number of training strategies, e.g., by backpropagation algorithms or particle swarm optimisation approaches.

Proper orthogonal decomposition

POD is a mathematical procedure that allows to perform a decomposition of a large set of data to describe the original system with a much smaller number of unknowns by projecting them onto subspaces. The method was proposed under different names for applications in various fields including principal component analysis (PCA) [22] in statistics, singular value decomposition (SVD) [23] in linear algebra and Karhunen–Loeve decomposition (KLD) [24,25] in signal processing. Nowadays the method is applied extensively in various branches of computational mechanics, such as fluid dynamics [6,26], aerodynamics [9] and nonlinear solid mechanics [27,28], see also [29] for a comprehensive review of POD applications.

The main idea of the POD is to find a useful set of basis vectors and the dimension of the subspace which is necessary to obtain a good approximation quality of the system. A widely used method to generate POD vectors is the method of snapshots introduced by Sirovich [30]. It is characterised by computing a set of system states or snapshots $\mathbf{U}$ and extracting an optimal set of basis vectors $\mathbf{\Phi}$, such that the error between the original and the projection onto subspace is minimised by

$$min. \parallel \mathbf{U}^i - \sum_{k=1}^{K} (\mathbf{U}^i \cdot \Phi^k)\Phi^k \parallel_{L^2}^2 . \tag{2}$$

The POD basis vectors must satisfy the two following conditions. Firstly, the observed data of the snapshots can be represented exactly by the POD basis functions. In other words, the snapshots and the linear subspaces spanned by the POD basis vectors coincide exactly [26]. Secondly, the first p POD basis vectors always capture more energy on average than p vectors of any other orthonormal basis [26].

Given a collection of N snapshots of the system obtained by varying the input parameters, where each set of input parameters results in a snapshot with M output values, the data can be organised into a rectangular M by N matrix $\mathbf{U}$, denoted as the snapshot matrix. The POD basis can be established by means of a SVD of the matrix $\mathbf{U}$ or by solving the eigenvalue problem of the sample covariance matrix $\mathbf{C} = \mathbf{U}^{\mathrm{T}} \cdot \mathbf{U}$. Employing the latter technique in this paper, the POD basis is expressed by a linear transformation

$$\mathbf{\Phi} = \mathbf{U} \cdot \mathbf{V} \tag{3}$$

of the snapshots $\mathbf{U}$. The coefficient matrix $\mathbf{V}$ satisfies the following eigenvalue problem

$$\mathbf{C} \cdot \mathbf{V} = \mathbf{\Lambda} \cdot \mathbf{V}, \tag{4}$$

where $\mathbf{\Lambda}$ is a diagonal matrix storing the eigenvalues λ_i of $\mathbf{C}$. $\mathbf{\Lambda}$ can be interpreted as the variance of the data set in the direction of the corresponding POD modes. $\mathbf{\Lambda}$ also measures the energy captured by the respective POD mode. The sum of all eigenvalues defines the total energy. The relative energy captured by the ith basis vector is $\Lambda_i / \sum_{j=1}^{N} \Lambda_j$. Introducing a truncated POD vector $\overline{\mathbf{\Phi}}$ approximating the exact basis vector $\mathbf{\Phi}$, the exact and approximated predictions of the snapshot matrix $\mathbf{U}$ can be represented by

$$\mathbf{U}_{M\times N} = \mathbf{\Phi}_{M\times N} \cdot \mathbf{A}_{N\times N}, \tag{5}$$

$$\mathbf{U}_{M\times N} \approx \overline{\mathbf{\Phi}}_{M\times K} \cdot \overline{\mathbf{A}}_{K\times N}, \tag{6}$$

with $K \ll N$ chosen according to the desired level of energy to be captured. Matrices $\mathbf{A}$ and $\overline{\mathbf{A}}$ are regarded as the amplitude matrix and *truncated* amplitude matrix, respectively. In other words, a *truncated* POD matrix $\overline{\mathbf{\Phi}}$ taken from the K first columns of the full matrix $\mathbf{\Phi}$ is used to approximate the original data matrix $\mathbf{U}$.

POD with interpolation

To predict the system behaviour related to intermediate values of input parameters that are not included in the snapshot data, interpolation is performed to determine the *truncated* amplitude matrix $\overline{\mathbf{A}}$, assuming that $\overline{\mathbf{A}}$ is a smooth function of input parameters. More specifically, each amplitude vector is defined as a linear combination of a set of vectors $\mathbf{F}_i = [f_1(\mathbf{z}^i) \cdots f_j(\mathbf{z}^i) \cdots f_N(\mathbf{z}^i)]^T$, with $f_j(\mathbf{z})$ as predefined interpolation functions of the input parameters $\mathbf{z}^i$ and an unknown coefficient matrix $\mathbf{B}$.

$$\overline{\mathbf{A}}_i = \mathbf{B} \cdot \mathbf{F}_i \tag{7}$$

POD-RBF The choice of $f_j(\mathbf{z})$ can be arbitrary. A suitable strategy is to employ RBF as interpolation functions. The idea of using RBF as approximation functions was first introduced by Hardy [31] to fit irregular topographical data. Since then, RBF have been successfully

applied to surrogate modelling strategies in the context of multidisciplinary design optimisation (MDO), see [32–34]. RBF are expressed in terms of the Euclidean distance $r = ||z - z^i||$ of a point z from a given point z^i, denoted as the center. Typical RBFs are:

i. Gaussian $\psi(r) = e^{-\frac{r^2}{c^2}}$ (8)

ii. Multiquadric $\psi(r) = \sqrt{r^2 + c^2}$ (9)

iii. Inverse multiquadric $\psi(r) = \frac{1}{\sqrt{r^2 + c^2}}$ (10)

iv. Thin plate splines $\psi(r) = r^k \ln r, \quad \mathrm{k} \in [2, 4, \ldots]$ (11)

Interpolation is also effective in inverse problems, where the objective is to find a set of parameters that produces computed results matching measured results. RBFs together with POD technique have been used in [3] and [4] to build a once-for-all trained POD-RBF network for solving inverse problems, choosing the inverse multiquadric function as interpolation function. Each element of the vector $\mathbf{F}_i$ is computed according to

$$f_j(\mathbf{z}^i) = f_j(|\mathbf{z}^i - \mathbf{z}^j|) = \frac{1}{\sqrt{|\mathbf{z}^i - \mathbf{z}^j|^2 + c^2}}\,, \tag{12}$$

with c as a smoothing factor. The matrix $\mathbf{F}$ containing the interpolation functions can be constructed from all vectors $\mathbf{F}_i$ of input parameters that are used to generate the snapshots. Using $\mathbf{F}$, the *truncated* amplitude matrix $\overline{\mathbf{A}}$ is given by

$$\overline{\mathbf{A}} = \mathbf{B} \cdot \mathbf{F}\,, \tag{13}$$

with

$$\overline{\mathbf{A}} = \overline{\mathbf{\Phi}}^{\mathrm{T}} \cdot \mathbf{U}\,. \tag{14}$$

From Eqs. (13) and (14), the coefficient matrix $\mathbf{B}$ is determined. Finally, the approximation of the output system response corresponding to an arbitrary set of input parameters is obtained by

$$\mathbf{U}_i \approx \overline{\mathbf{\Phi}} \cdot \mathbf{B} \cdot \mathbf{F}_i\,. \tag{15}$$

POD-ERBF In the context of multidisciplinary design optimisation (MDO), Mullur and Messac [7] introduced the method of extended RBFs (ERBF) by combining RBFs with non-RBFs. This allows a greater flexibility in the model generation, because the typical RBF approach provides an interpolating surface that is unique for a given set of data, i.e., it yields only a unique interpolative solution to the surrogate modelling problem. In addition, the typical RBF approach does not allow to express desirable properties of the surrogate models. Consequently, the ERBF surrogate model possesses the effectiveness of RBF together with the flexibility of non-RBFs. Recently, an approach combining ERBF and POD as a surrogate model for application in mechanised tunnelling has been proposed in [8]. This modified version of POD-RBF method, which has shown improvements in prediction accuracy of the generated surrogate models for both linear and nonlinear functions in multi dimensional spaces will be utilised in this paper.

The main idea of ERBF is to supplement the RBF with a non-RBF term, which are not functions of the Euclidean distance r. Instead, they are functions of ξ^i, which is the vector pointing from the coordinate of a generic point z in the input space to a data point z^i,

defined as $\xi^i = z - z^i$. Figure 4 depicts the difference between the Euclidean distance r and the relative coordinates ξ used in ERBF. The non-RBF terms ϕ_{ij} for the ith data point and the jth dimension are assumed in the format:

$$\phi_{ij}(\xi_j^i) = \alpha_{ij}^L \phi^L(\xi_j^i) + \alpha_{ij}^R \phi^R(\xi_j^i) + \beta_{ij}\phi^\beta(\xi_j^i). \tag{16}$$

ϕ^L, ϕ^R and ϕ^β are components defined in Table 2. The parameter γ in Table 2 can be considered as a smoothness parameter. Figure 5 shows four distinct regions (I–IV) of non-RBF with superscripts L and R corresponding to left and right parts. More details on the design of non-RBF parameters are contained in [7,35].

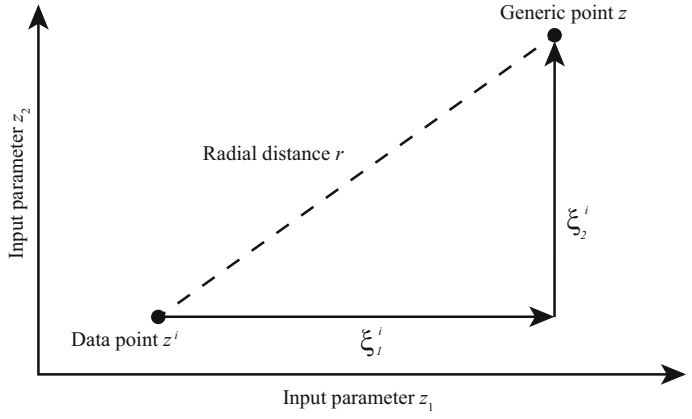

Fig. 4 Euclidean distance r and relative coordinates ξ, according to [7]

Table 2 Non-radial basis functions [7]

Region	Range of ξ_j^i	ϕ^L	ϕ^R	ϕ^β
I	$\xi_j^i \le -\gamma$	$(-n\gamma^{n-1})\xi_j^i + \gamma^n(1-n)$	0	ξ_j^i
II	$-\gamma \le \xi_j^i \le 0$	$(\xi_j^i)^n$	0	ξ_j^i
III	$0 \le \xi_j^i \le \gamma$	0	$(\xi_j^i)^n$	ξ_j^i
IV	$\xi_j^i \ge \gamma$	0	$(n\gamma^{n-1})\xi_j^i + \gamma^n(1-n)$	ξ_j^i

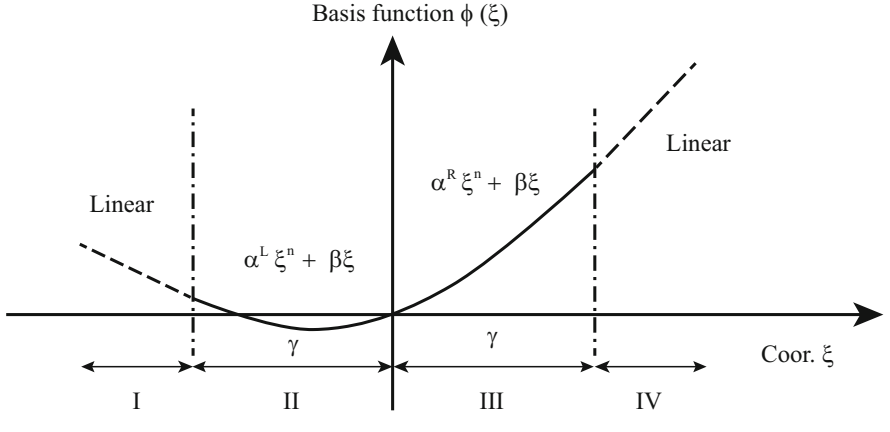

Fig. 5 Non-radial basis functions, according to [7]. The functions are linear in regions I and IV; and *nth* degree monomials plus linear terms in regions II and III

According to the POD-ERBF, the approximation is performed by a linear combination of radial and non-RBFs. The following equation provides the approximated amplitude value of an arbitrary point using the ERBF approach:

$$a(\mathbf{z}) \approx \sum_{i=1}^{N} c^i f_j(\mathbf{z}) + \sum_{i=1}^{N}\sum_{j=1}^{S} \left\{ \alpha_{ij}^L \phi^L(\xi_j^i) + \alpha_{ij}^R \phi^R(\xi_j^i) + \beta_{ij}\phi^\beta(\xi_j^i) \right\}, \tag{17}$$

with S as the dimension of input parameters. The key idea of the POD with interpolation is to modify the amplitude matrix $\overline{\mathbf{A}}$ as a smooth function. Using the ERBF approach, Eq. (7) can be re-written in full and truncated matrix form as follows:

$$\left[\overline{\mathbf{A}}\right]_{K\times N} = [\mathbf{C}]_{K\times N}[\mathbf{F}]_{N\times N} + [\mathbf{D}]_{K\times 3SN}[\mathbf{G}]_{3SN\times N} \tag{18}$$

$$\left[\overline{\mathbf{A}}\right]_{K\times N} = [\mathbf{CD}]_{K\times(S+3SN)}[\mathbf{FG}]^T_{(S+3SN)\times N} \tag{19}$$

$\mathbf{C}$ and $\mathbf{D}$ are unknown coefficients matrices. The matrix $\mathbf{F}$ is generated according to Eq. (12) and the matrix $\mathbf{G}$ is calculated as

$$[\mathbf{G}]_{3SN\times N} = \left\{ \mathbf{G}^1 \quad \mathbf{G}^2 \quad \dots \quad \mathbf{G}^N \right\} \tag{20}$$

with

$$\left[\mathbf{G}^k\right]_{3SN\times 1} = \left\{ \mathbf{G}^{Lk} \quad \mathbf{G}^{Rk} \quad \dots \quad \mathbf{G}^{\beta k} \right\}^T \quad k = 1, \dots, N \tag{21}$$

and

$$\left[\mathbf{G}^{Lk}\right]_{1\times SN} = \left[\phi^L\left(z_1^k - z_1^1\right) \phi^L\left(z_2^k - z_2^1\right) \dots \phi^L\left(z_S^k - z_S^1\right) \dots \phi^L\left(z_S^k - z_S^N\right) \right] \tag{22}$$

$$\left[\mathbf{G}^{Rk}\right]_{1\times SN} = \left[\phi^R\left(z_1^k - z_1^1\right) \phi^R\left(z_2^k - z_2^1\right) \dots \phi^R\left(z_S^k - z_S^1\right) \dots \phi^R\left(z_S^k - z_S^N\right) \right] \tag{23}$$

$$\left[\mathbf{G}^{\beta k}\right]_{1\times SN} = \left[\phi^\beta\left(z_1^k - z_1^1\right) \phi^\beta\left(z_2^k - z_2^1\right) \dots \phi^\beta\left(z_S^k - z_S^1\right) \dots \phi^\beta\left(z_S^k - z_S^N\right) \right] \tag{24}$$

Considering $[\mathbf{C}\ \mathbf{D}] = [\mathbf{Q}]$ and $[\mathbf{F}\ \mathbf{G}]^T = \mathbf{R}$ such that $\mathbf{QR} = \overline{\mathbf{A}}$, the unknown matrix $\mathbf{Q}$ can be obtained as proposed in [7]:

$$\mathbf{Q} = \overline{\mathbf{A}}\left[\mathbf{R}^+\right], \tag{25}$$

with $[\mathbf{R}^+] = (\mathbf{R}^T\mathbf{R})^{-1}\mathbf{R}^T$ as the pseudo-inverse of matrix $\mathbf{R}$ [7]. Consequently, $\mathbf{C}$ and $\mathbf{D}$ are determined from $\mathbf{Q}$. Finally, the system response of an arbitrary set of input parameters $\mathbf{z}_a$ in the parameter space is approximated by:

$$\mathbf{U}_a = \overline{\Phi}(\mathbf{CF}_a + \mathbf{DG}_a) \tag{26}$$

with $\mathbf{F}_a$ and $\mathbf{G}_a$ determined from Eq. (12) and Eqs. (21)–(24), respectively.

POD with missing data

Without any missing data, an arbitrary snapshot $\mathbf{U}_j$, which belongs to a set of snapshots, can be approximated as a linear combination of the first K POD basis vectors $\overline{\mathbf{\Phi}}$ and an amplitude vector $\overline{\mathbf{A}_j}$ as described in "Proper orthogonal decomposition". The amplitude vector is calculated by minimising the error norm

$$min. \parallel \mathbf{U}_j - \overline{\mathbf{\Phi}} \cdot \overline{\mathbf{A}_j} \parallel^2_{L^2} . \tag{27}$$

The same least square approach can be effectively used to restore missing data of an incomplete data snapshot $\mathbf{U}^*$ by

$$min. \parallel \mathbf{U}^* - \overline{\mathbf{\Phi}} \cdot \overline{\mathbf{A}}^* \parallel_{L^2}^2 . \tag{28}$$

However, due to missing elements, the L^2 norm in Eq. (27) cannot be evaluated correctly. As a remedy, the Gappy POD procedure, introduced by Everson and Sirovich [6], employs the concept of a gappy norm based on available data. This concept was successfully applied in aerodynamics [9] for the data recovery of the flow field. The procedure is summarised below.

It is assumed, that an output system response snapshots set $\mathbf{U}$, where all snapshots are completely known, together with the POD basis $\mathbf{\Phi}$ is given. $\mathbf{U}^*$ is another solution vector, which has some elements missing, specified by a corresponding mask vector $\mathbf{m}$. The vector $\mathbf{m}$ is defined to indicate the locations of missing data in the data set

$$\begin{aligned} \mathbf{m}_i &= 0, \quad \text{for locations of unknown or missing data} \\ \mathbf{m}_i &= 1, \quad \text{for locations of known data} \end{aligned} \tag{29}$$

The gappy norm is defined with a gappy inner product $(\cdot, \cdot)_n$, such that

$$\parallel \mathbf{U}_j \parallel_n^2 = (\mathbf{U}_j, \mathbf{U}_j)_n = (\mathbf{m} \circ \mathbf{U}^*, \mathbf{m} \circ \mathbf{U}^*)_{L^2} = \parallel \mathbf{m} \circ \mathbf{U}^* \parallel_{L^2}^2, \tag{30}$$

where $\circ$ denotes point-wise multiplication. The complete (repaired) vector from $\mathbf{U}^*$ can be reproduced using the assumption that $\mathbf{U}^*$ can be characterised with the existing snapshots set $\mathbf{U}$. The intermediate repaired vector $\widetilde{\mathbf{U}}^*$ can be expressed in terms of the *truncated* POD basis vectors $\overline{\mathbf{\Phi}}$ as

$$\widetilde{\mathbf{U}}^* \approx \overline{\Phi} \cdot \overline{\mathbf{A}}^*. \tag{31}$$

By minimising the error norm $E = ||\mathbf{U}^* - \widetilde{\mathbf{U}}^*||_n^2$, the coefficient vector $\overline{\mathbf{A}}^*$ can be computed. A solution to this *least squares* or *linear regression* problem is given by a linear system of equations

$$\mathbf{M} \cdot \overline{\mathbf{A}}^* = \mathbf{R}. \tag{32}$$

Herein, $\mathbf{M} = (\overline{\mathbf{\Phi}}^{\mathrm{T}}, \overline{\mathbf{\Phi}})_n$ and $\mathbf{R} = (\overline{\mathbf{\Phi}}^{\mathrm{T}}, \mathbf{U}^*)_n$ are gappy inner products. Evaluating Eq. (31) with $\overline{\mathbf{A}}^*$ computed from Eq. (32), the intermediate repaired vector $\widetilde{\mathbf{U}}^*$ is obtained. Finally, by replacing the missing elements in $\mathbf{U}^*$ by those in $\widetilde{\mathbf{U}}^*$, a complete vector containing the system response is reconstructed.

Iterative GPOD In the method described above, the reconstruction procedure is accomplished in one step with the assumption that the "repaired" vector can be characterised with the already known snapshots set. This method can be extended to the case, where the missing vector itself is included into the snapshots sets. This requires an iterative procedure to construct the POD basis. Two algorithms, the E–S and the V–K method proposed in the literature, are briefly summarised below.

In the E–S method [6], initial guesses are made at missing locations, which are iteratively updated:

(a) Use time-average values as initial guess at missing locations to obtain $\widetilde{\mathbf{U}}^*$.
(b) Include $\widetilde{\mathbf{U}}^*$ into snapshot $\mathbf{U}$ to form $\widetilde{\mathbf{U}}$.
(c) Perform POD on $\widetilde{\mathbf{U}}$ to obtain $\mathbf{\Phi}$.
(d) Choose the number of modes K to be employed in the reconstruction.

(e) Compute $\overline{\mathbf{A}}^*$ from Eq. (32) with $\mathbf{M} = (\overline{\mathbf{\Phi}}^{\mathrm{T}}, \overline{\mathbf{\Phi}})$ and $\mathbf{R} = (\overline{\mathbf{\Phi}}^{\mathrm{T}}, \mathbf{U}^*)$.
(f) Compute $\widetilde{\mathbf{U}}$ from Eq. (31) and overwrite the previous guess at missing locations.
(e) Proceed until convergence, the eigenvalues no longer change. If no convergence is obtained, go to (b).

An extension of the E–S method, which is not dependent on an initial guess and which improves the accuracy significantly, has been proposed by Venturi and Karniadakis in [11]. The main steps of this extended procedure, referred to as the V–K method, are described below:

(a) Perform the standard E–S procedure, but employ only $K = 2$ modes in the reconstruction.
(b) Take the converged result from the previous step as a new initial guess.
(c) Perform the E–S procedure, but employ now $K = 3$ modes in the reconstruction.
(d) Proceed until convergence (the eigenvalues no longer change).

Application to mechanised tunnelling

In this section, the performance of the proposed hybrid surrogate model is demonstrated by means of an application concerned with the numerical simulation of the advancement process of a TBM driven tunnel. The main goal is to demonstrate the capability of the enhanced surrogate model, supplemented with ERBF and IGPOD, to provide reliable predictions of the expected settlements induced by mechanised tunnelling. To this end, the model predictions will be evaluated by comparing the predictions with results from a surrogate model proposed in [1] and with reference results obtained from the original process-oriented finite element model *ekate*.

Figure 6 illustrates the geometry of a synthetic example of a tunnel section of 48 m length, constructed by a tunnel boring machine (TBM) and its discretisation by means

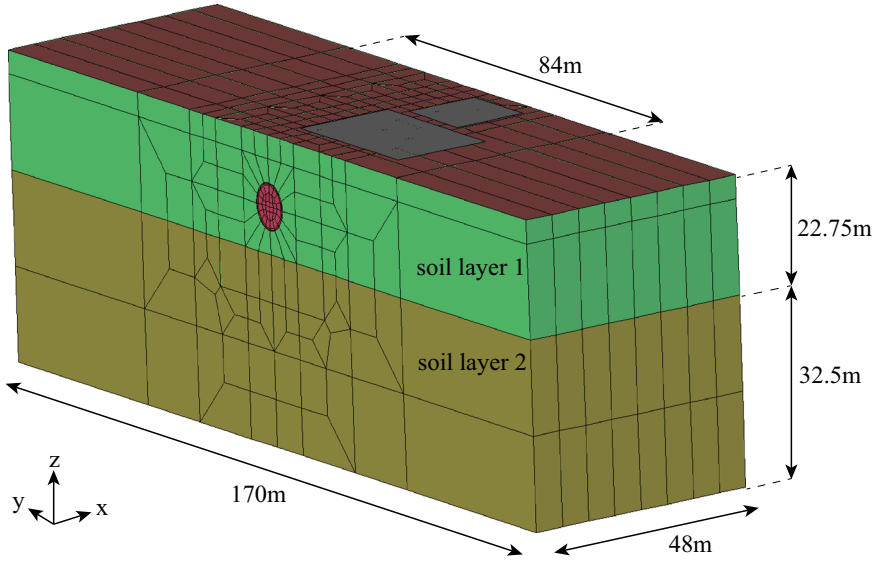

Fig. 6 Numerical simulation model of a tunnel section: layout and geometry. The tunnel is excavated through the top soil layer in 32 discrete excavation steps, with a ring length of 1.5 m. The groundwater table is at the top surface; the stiffness of existing buildings is considered by substitute models

of 11,072 quadrilateral two-field finite elements with quadratic approximations for the displacements and linear approximations of the water pressure. The tunnel with a diameter of 8.5 m is excavated with an overburden of 8.5 m. The tunnel lining has a thickness of 0.3 m; each lining ring has a length of 1.5 m. The effect of the stiffness of two existing buildings located at the surface is considered by placing rectangular plate-like substitute models with an equivalent thickness of 5 m and a stiffness of 50 GPa at the top of the discretised soil body with a length of 48 m and a width of 170 m (Figs. 6, 7). The numerical simulation of the mechanised tunnelling process is characterised by a step-by-step procedure consisting of different phases: soil excavation at the tunnel face, application of the face support pressure, advancement of the TBM, installation of a tunnel ring and application of the tail void grouting pressure. The soil model comprises of two layers of soft cohesive soils. The tunnel is excavated completely within the top layer. The groundwater table is assumed to be situated at the ground surface. The constitutive behaviour of the soil is modelled by means of a Drucker–Prager plasticity model with the friction angle, cohesion, and Poisson's ratio taken as 30°, 10 kPa and 0.3 for layer 1 and 28°, 50 kPa and 0.3 for layer 2, respectively. For the tunnel lining and the TBM shield, a linear elastic material law is assumed. The support pressure, which is necessary to prevent tunnel face collapse, is assumed to remain at a value of 180 kPa at the tunnel axis during the tunnel advance. In contrast, the grouting pressure filling the annular gap, which is important to avoid large settlements at the ground surface and large deformations of the surrounding soil, is considered as a time variant process parameter. Because of the fact that beyond a distance of 42 m in Y-direction from both sides of the tunnel axis the surface settlements are almost zero, an effective rectangular surface area with 105 grid points as illustrated in Fig. 7 is considered for the generation of the surrogate model.

In the present demonstration example, it is assumed that the current state of the TBM advance corresponds to the 22nd step of the excavation process and that the history of the settlements of the surface monitoring points are known. The proposed hybrid surrogate modelling approach is employed to predict the complete surface displacement field in the subsequent excavation step (step 23) from input parameter selected within a specific

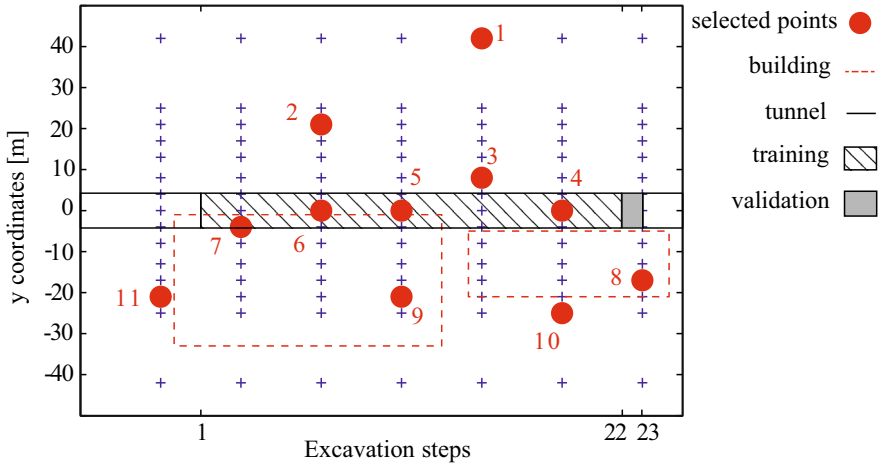

Fig. 7 Investigated surface area for settlement predictions. The surface is represented by 105 grid points (*plus symbol*); 11 points are selected as monitoring points (*red points*)

range. In this application example, 11 monitoring points are selected for predicting the future tunnelling induced settlements by the RNN approach. The number of monitoring points is selected to ensure that the RNN is capable to provide good predictions and to have an appropriate accuracy of the GPOD. In particular, the reconstruction quality of the GPOD would be better, if there are more available data points. However, the training and prediction of the RNN might be more complicated. In addition, the position of 11 points are chosen based on the usual position of measurement sensors on the surface in a real tunnel project. Subsequently, the complete surface displacement field will be approximated with the GPOD method.

The following steps describe how to apply the proposed strategy to mechanised tunnelling following the algorithm in Table 1.

Offline stage

- In order to set up the surrogate model, 60 numerical simulations are performed to obtain data in the offline stage. Each simulation corresponds to a combination of two varying parameters: the elastic modulus E_1 of soil layer 1 and the grouting pressure ${}^{[n]}GP$.
- The range of variation of these parameters are 20–110 MPa for E_1 and 130–230 kPa for ${}^{[n]}GP$, respectively. E_1 can take 1 out of 10 particular values in the range from 20 to 110 MPa, whereas for ${}^{[n]}GP$ one of the six scenarios of time varying pressures is taken, see Fig. 8.
- The numerical results of 60 simulations are split randomly into two separate data sets. The first set containing results from 54 simulations is used to establish the surrogate model. In contrast, the results from six cases are stored for validation.
- In addition, only the displacements of the complete surface field from time step 1–22 are kept in the first set for training and generating the surrogate model. Meanwhile, the validation set contains surface settlements from time step 1–23. The results from time step 23 are used to validate the prediction capability of the proposed surrogate

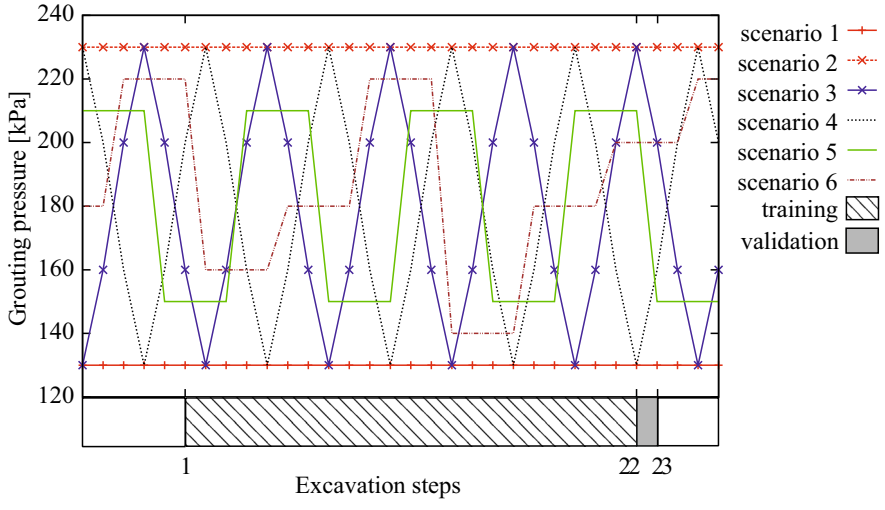

Fig. 8 Scenarios adopted for the grouting pressure during tunnel advancement

modelling strategy. The performances of the POD-RBF and the POD-ERBF networks are compared with numerical results from time step 1–22.

Online stage

- An arbitrary set of input parameters within the investigated ranges is selected: a value of E_1 in the range from 20 to 110 MPa and values of ${}^{[n]}GP$ from time step 1–22 in the range of 130–230 kPa.
- Approximate the complete displacement field from time step 1–22 by POD-RBF based on the data of 54 numerical simulations.
- With an arbitrary value of ${}^{[23]}GP$ in the range of 130–230 kPa, the RNN predicts the settlements at 11 selected monitoring points of time step 23.
- The complete displacement field of time step 23 is predicted using GPOD.

It should be noted that the same section of the mechanised tunnelling project has been investigated previously in [1] with a similar surrogate modelling strategy using POD-RBF and GPOD. This allows to compare the POD-RBF and GPOD approach with the enhanced by POD-ERBF and IGPOD (E–S and V–K) methods, respectively. The error between prediction and FE result is calculated as

$$error = \sqrt{\frac{\sum_{i=1}^{N}(\mathbf{U}_i^{FE} - \mathbf{U}_i^{*})^2}{\sum_{i=1}^{N}(\mathbf{U}_i^{FE})^2}} \times 100\,\%. \tag{33}$$

The POD-RBF network is used to approximate the surface settlements within all 105 grid points from all previous excavation steps (steps 1–22), containing a total number of 2310 settlement values. Figure 9 shows a comparison between the results from the POD-ERBF and the POD-RBF with the FE results for all 105 surface points and all time steps 1–22 for validation case 1, i.e., the value of E_1 is 90 MPa and ${}^{[n]}GP$ is chosen according to pressure scenario 5, see Fig. 8. Table 3 shows the prediction performances of POD-RBF and POD-ERBF of the six validation cases and the corresponding average values. Only a slight improvement of the predictions of the surface settlements is observed when employing the POD-ERBF instead of POD-RBF. Considering all 22 steps in the evaluation of the error according to Eq. (33), the average error decreases from 1.91 to 1.46 % when using

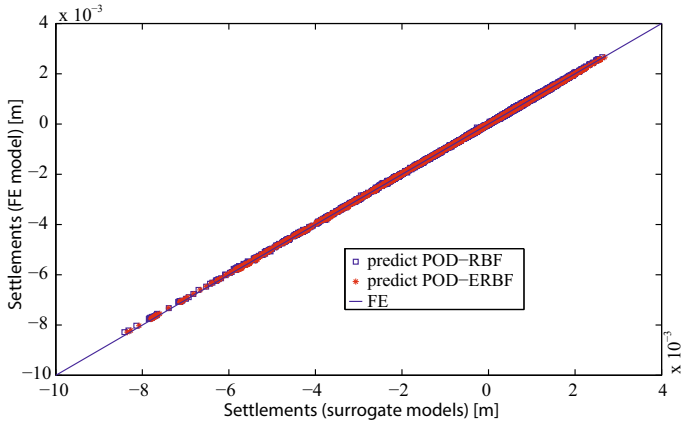

Fig. 9 Comparison of RBF and ERBF with FE results. Predictions are performed for 22 steps with errors of 1.50 % by using POD-RBF and of 0.87 % by using POD-ERBF (for validation case 1)

Table 3 Prediction performances of POD-RBF and POD-ERBF for all validation cases

Error (%)	POD-RBF	POD-ERBF
Validation case 1	1.50	0.87
Validation case 2	2.51	2.46
Validation case 3	1.80	1.38
Validation case 4	2.04	1.52
Validation case 5	2.06	1.57
Validation case 6	1.57	0.97
Average	1.91	1.46

the POD-ERBF. A comparison between the predicted results for the surface settlements in all surface points using the GPOD and the IGPOD approaches with the reference results from the original FE simulation is presented in Table 4. As expected, the IGPOD provide better predictions with the method V–K giving the best approximation results for all validation cases. Figure 10 depicts the predicted surface settlements in comparison with the FE results for validation case 1. It can be concluded that the POD-ERBF and the IGPOD have the possibility to enhance the prediction performances of the corresponding components (POD-RBF and GPOD) of the proposed surrogate model individually. The next paragraph investigates the consequence when integrating POD-ERBF and IGPOD into the hybrid surrogate model.

Table 4 Prediction performances of GPOD and IGPOD for all validation cases

Error (%)	POD-RBF			POD-ERBF		
	GPOD	E–S	V–K	GPOD	E–S	V–K
Validation case 1	5.18	5.03	1.36	5.20	5.06	1.18
Validation case 2	7.82	7.80	1.18	7.65	7.63	1.01
Validation case 3	4.81	4.70	1.38	4.55	4.42	0.95
Validation case 4	4.33	4.22	2.13	4.49	4.33	1.15
Validation case 5	7.45	7.24	2.15	6.90	6.75	1.09
Validation case 6	5.40	5.36	1.81	5.60	5.55	2.31
Average	5.83	5.72	1.67	5.73	5.62	1.28

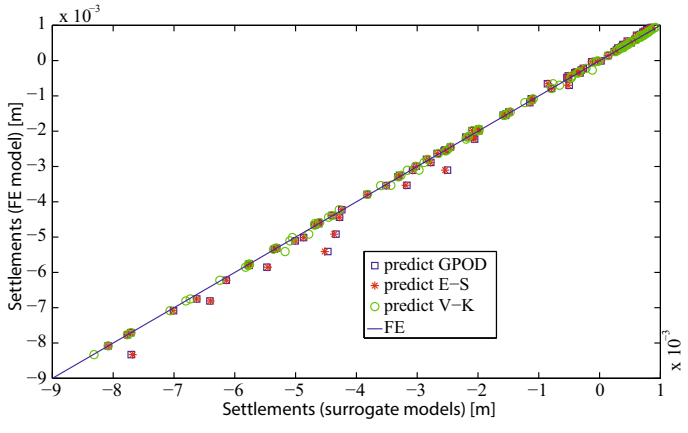

Fig. 10 Comparison of GPOD and iterative GPOD methods (IGPOD) with FE results. Prediction of surface settlements in step 23. The GPOD, the IGPOD based upon the Everson–Sirovich(E–S) method and the Venturi–Karniadakis (V–K) method lead to an error of 5.18; 5.03 and 1.36 %, respectively (for validation case 1)

The investigation is carried out by combining the POD-RBF and the POD-ERBF separately with the GPOD and the E–S and V–K enhancements, respectively. This leads to six surrogate models in total. According to Table 4 the models with the E–S method produce slightly better results compared with models employing the GPOD method. This can be explained by the fact that reconstruction results from both GPOD and E–S depends on the initial guess, see [11]. However, with errors of only 1,18 and 1,36 % as compared to error levels larger than 5 %, the best accuracy is obtained from POD-RBF and POD-ERBF models enhanced by the V–K method. This agrees well with observations in [11], where the V–K method was also shown to produce the relative best results as it does not depend on an initial guess. The reconstruction problem is considered for a rectangular matrix (105×23) with 23 time steps and 105 settlement points in which all elements from column 1 to column 22 and 11 elements of column 23 are known. The 94 missing elements of time step 23 over a total of $105 \times 23 = 2415$ elements constitute a small gappiness of 4 %. This may explain, why the prediction accuracy and computation time of surrogate models in this example differ only slightly. The average computation time for one run of the hybrid surrogate model is summarised in Table 5. It is noted, that the previous approach using RBF and GPOD needs less computing time as compared to the iterative and more complex POD-ERBF approach in conjunction with IGPOD. However, with a maximum computing time of less than one second, the criterion for real-time application is still fulfilled. Therefore, the enhanced hybrid surrogate model will be chosen as the main ingredient for a new software tool proposed in the next section, which is targeted to on site applications to support the steering of TBMs.

Real-time simulation software for TBM steering support

Based on the enhanced IGPOD-ERBF algorithm described in 'Hybrid surrogate model', a real-time simulation software is developed with the aim to support the steering of the tunnel boring machine during mechanised tunnelling. The goal of this software is to predict the system response, i.e., the surface settlements (and/or tunnel lining forces etc.), resulting from the TBM-soil interaction in real-time as the response to changes of operational parameters, such as the face pressure or the grouting pressure.

The example described in this section is similar to the application example in the previous section. However, two varying operational parameters in mechanised tunnelling are considered here. They are the support pressure ${}^{[n]}SP$ and the grouting pressure ${}^{[n]}GP$ which constitute the main operational parameters to control the settlements caused by the tunnel drive. The software consists of three main modules: "Overview" module, "Monitoring" module and "Prediction" module. For the prediction of time variant surface settlement fields for the next excavation step, the user has to follow the sequence of all three modules.

The "Overview" modules gives a brief description about basic parameters of the selected tunnel section of a tunnelling project, such as cover depth, tunnel radius, ring length,

Table 5 Computation time of different surrogate models (in average)

	POD-RBF			POD-ERBF		
	GPOD	E–S	V–K	GPOD	E–S	V–K
Computation time [s]	0.01	0.03	0.06	0.61	0.73	0.92

current position, etc. In addition, the module also allows to record the complete history of the operational parameters of the tunnel construction described above, the support pressure ${}^{[n]}SP$ and the grouting pressure ${}^{[n]}GP$. The offline data obtained from numerical simulations performed in the design stage are loaded to set up the surrogate model. Figure 11 shows a screenshot of the "Overview" module with two graphs representing the history of ${}^{[n]}SP$ and ${}^{[n]}GP$ upto the current state of the project.

The "Monitoring" module is used to store and visualise the locations of all monitoring points and the evolution of the settlements in the already constructed tunnel section, see Fig. 12. The settlement history of these monitoring points will be extrapolated with

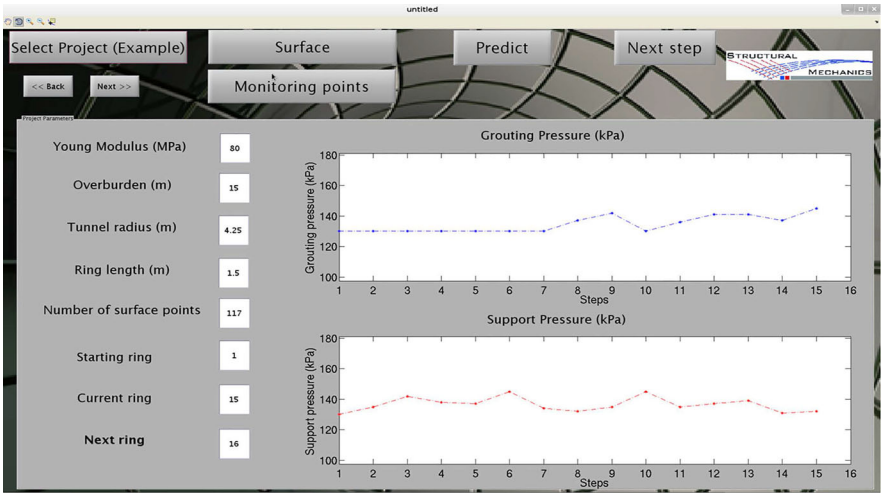

Fig. 11 TBM real-time steering software: overview module. This module provides information about project parameters, the TBM position (here: ring number 15) and the history of two operational parameters (grouting pressure and face support pressure)

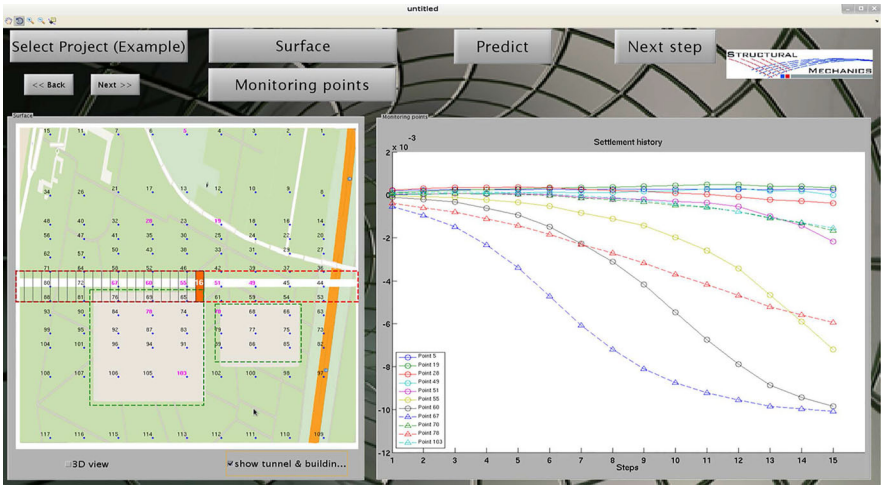

Fig. 12 TBM real-time steering software: monitoring module. This module provides on the *left panel* the location of monitoring points and the layout of the tunnel drive including existing buildings, and on the *right panel* the evolution of the settlement trough. This graph is updated in subsquent steps according to the choice of steering parameters

predicted values from RNN for the next time step corresponding to the chosen support pressure and grouting pressure. In practice, the predicted values from RNN are updated with real measurements at monitoring points. Nevertheless, in this example this step is skipped since there are only synthetic data from numerical simulations.

The most important part of the software is the "Prediction" module, where the user can see almost immediately the complete surface settlements of the next time step caused by changes of the grouting pressure or the face support pressure applied through movements of the corresponding sliders on the GUI. The module offers the possibility for the user to define a limit state, such as a maximum tolerated settlement in any point on the surface. Figure 13 demonstrates that by defining a limit state and varying the grouting or the face support pressure, red points appearing on the left panel may appear to indicate that in these points the tolerated displacements would be exceeded if the chosen combination of pressures would be applied. In summary, with the developed software users can obtain in real-time the surface behaviour resulting from changing the operational parameters in mechanised tunnelling. Therefore, it provides TBM drivers an assistant tool for decision-making during the TBM steering.

The tests presented in this paper are executed on a standard computer with IntelChip 2×1.70 GHz, 2×4 GB RAM. The software can run in almost every standard laptop or even tablet since it does not require any expensive hardware.

Conclusion

In this paper, a hybrid surrogate modelling strategy based upon the combination of a RNN and the GPOD has been proposed to enable real-time prognoses during mechanised tunnelling. The proposed approach is an extension of preliminary work by the authors, characterised by supplementing the previous surrogate model with Extended RBFs and an iterative Gappy POD. It was proven, that the suggested enhancements improve the prediction capability of the surrogate model. Although more complex formulations and an iterative concept are involved, the new hybrid surrogate model still provides predictions

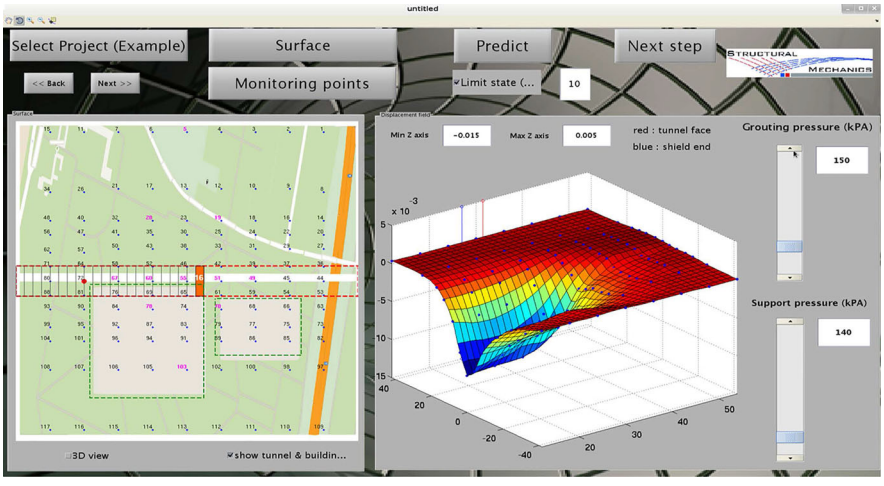

Fig. 13 TBM real-time steering software: prediction module. This module provides real-time predictions of the surface settlements corresponding to operational parameters selected by the user via sliders (*right panel*)

within a reasonable time in the context of applications in mechanised tunnelling, targeted to predict tunnelling induced settlements in real-time. In particular, the surrogate model combining the POD-ERBF and the V–K method yields the best prediction results in the presented application. The new approach has been integrated into a software which provides, in real-time, the expected tunnelling induced settlements for varying operational parameters chosen by the user and thus enables to support the steering of tunnel boring machines. The developed software provides a new option for tunnel engineers on the construction site to select the appropriate parameters such, that tolerated settlements will not be exceeded during the upcoming excavation steps.

The proposed hybrid surrogate model will be further extended to consider uncertain geotechnical parameters. First steps include interval and fuzzy data within the presented RNN-GPOD approach to allow predicting time variant interval settlement fields by means of interval arithmetic operations computed in real-time. In contrast to more time consuming optimisation approaches for interval analysis with the hybrid RNN-GPOD surrogate model, see [1], the problem of overestimation has to be considered within interval arithmetic. In addition, a strategy is currently developed to continuously reduce the predicted interval uncertainty by updating the settlement field with monitoring data.

Authors' contributions
All authors, BTC, SF and GM, participated equally in the manuscript writing. All authors read and approved the final manuscript.

Author details
[1]Institute for Structural Mechanics, Ruhr University Bochum, Universitätstr. 150, 44801 Bochum, Germany, [2]Collaborative Research Center for Interaction Modelling in Mechanised Tunnelling, Ruhr University Bochum, Universitätstr. 150, 44801 Bochum, Germany.

Acknowlegements
Financial support was provided by the German Research Foundation (DFG) in the framework of project C1 of the Collaborative Research Center SFB 837 "Interaction Modelling in Mechanised Tunnelling". This support is gratefully acknowledged.

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

Received: 14 October 2015 Accepted: 28 January 2016
Published online: 05 March 2016

References

1. Freitag S, Cao BT, Ninić J, Meschke G. Hybrid surrogate modelling for mechanised tunnelling simulations with uncertain data. Int J Reliab Saf. 2015;9(2/3):154–73.
2. Freitag S, Graf W, Kaliske M. Recurrent neural networks for fuzzy data. Integr Comput Aided Eng. 2011;18(3):265–80.
3. Ostrowski Z, Bialecki RA, Kassab AJ. Solving inverse heat conduction problems using trained pod-rbf network inverse method. Inverse Probl Sci Eng. 2008;16(1):39–54.
4. Buljak V, Maier G. Proper orthogonal decomposition and radial basis functions in material characterization based on instrumented indentation. Eng Struct. 2011;33:492–501.
5. Ostrowski Z, Bialecki RA, Kassab AJ. Estimation of constant thermal conductivity by use of proper orthogonal decomposition. Comput Mech. 2005;37:52–9.
6. Everson R, Sivorich L. Karhunen–Loeve procedure for gappy data. J Opt Soc Am A Opt Image Sci Vis. 1995;12(8):1657–64.
7. Mullur AA, Messac A. Extended radial basis functions: more flexible and effective metamodeling. AIAA J. 2005;43(6):1306–15.
8. Khaledi K, Miro S, König M, Schanz T. Robust and reliable metamodels for mechanized tunnel simulations. Comput Geotech. 2014;61:1–12.
9. Bui-Thanh T, Damodaran M, Willcox K. Aerodynamic data reconstruction and inverse design using proper orthogonal decomposition. AIAA J. 2004;42:1505–16.
10. Willcox K. Unsteady flow sensing and estimation via the gappy proper orthogonal decomposition. Comput Fluids. 2006;35(2):208–26.
11. Venturi D, Karniadakis GE. Gappy data and reconstruction procedures for flow past a cylinder. J Fluids Mech. 2004;519:315–36.
12. Gunes H, Sirisup S, Karniadakis GE. Gappy data: to krig or not to krig? J Comput Phys. 2006;212:358–82.

13. Nagel F, Stascheit J, Meschke G. Process-oriented numerical simulation of shield tunneling in soft soils. Geomech Tunn. 2010;3(3):268–82.
14. Kasper T, Meschke G. A 3D finite element model for TBM tunneling in soft ground. Int J Numer Analyt Methods Geomech. 2004;28:1441–60.
15. Stascheit J, Nagel F, Meschke G, Stavropoulou M, Exadaktylos G. An automatic modeller for finite element simulations of shield tunnelling. In: Eberhardsteiner J, Beer G, Hellmich C, Mang HA, Meschke G, Schubert W, editors. Computational modelling in tunnelling (EURO:TUN 2007). Vienna; 2007.
16. Meschke G, Ninic J, Stascheit J, Alsahly A. Parallelized computational modeling of pile-soil interactions in mechanized tunneling. Eng Struct. 2013;47:35–44.
17. Cao BT, Chmelina K, Stascheit J, Meschke G. Enhanced monitoring and simulation assisted tunnelling (emsat). In: Meschke G, Eberhardsteiner J, Soga K, Schanz T, Thewes M, editors. Computational methods in tunnelling and subsurface engineering (EURO:TUN 2013). 2013. p. 41–50.
18. Ninić J, Koch C, Freitag S, Meschke G, König M. Simulation- and monitoring-based steering for mechanized tunneling using project data of Wehrhahn-Linie. In: Proceedings of the 11th World Congress on Computational Mechanics (WCCM 2014), Barcelona; 2014 **(online available)**.
19. Freitag S, Cao BT, Ninić J, Meschke G. Hybrid rnn-gpod surrogate model for simulation and monitoring supported tbm steering. In: Tsompanakis Y, Kruis J, Topping BHV, editors. Proceedings of the 4th International conference on soft computing technology in civil, structural and environmental engineering (CSC2015). Prague: Civil-Comp Press; 2015. paper 2. doi:10.4203/ccp.109.2.
20. Adeli H. Neural networks in civil engineering: 1989–2000. Comput Aided Civ Infrastruct Eng. 2001;16:126–42.
21. Freitag S. Artificial neural networks in structural mechanics. In: Tsompanakis Y, Kruis J, Topping BHV, editors. Computational technology reviews, vol. 12. Stirlingshire: Saxe-Coburg Publications; 2015. p. 1–26. doi:10.4203/ctr.12.1.
22. Hotelling H. Analysis of a complex system of statistical variables into principal components. J Educ Psychol. 1933;24:417–441498520.
23. Golub GH, Reinsch C. Singular value decomposition and least squares solutions. Numer Math. 1970;14(5):403–20.
24. Karhunen K. Zur spektraltheorie stochastischer prozesse. Ann Acad Sci Fennicae 37. 1946.
25. Loeve M. Probability theory II. 4th ed. Graduate texts in mathematics, vol. 46. New York: Springer; 1978.
26. Holmes P, Lumley JL, Berkooz G. Turbulence, coherent structures. Dynamical systems and symmetry. New York: Cambridge University Press; 1996.
27. Radermacher A, Reese S. Model reduction in elastoplasticity: proper orthogonal decomposition combined with adaptive sub-structuring. Comput Mech. 2014;54(3):677–87.
28. Galland F, Gravouil A, Malvesin E, Rochette M. A global model reduction approach for 3d fatigue crack growth with confined plasticity. Comput Method Appl Mech Eng. 2011;200(5–8):699–716.
29. Mifsud M. Reduced-order modelling for high-spped aerial weapon aerodynamics. Ph.D. Thesis. Cranfield University-College of Aeronautics; 2008.
30. Sirovich L. Turbulence and the dynamics of coherent structures: i, ii and iii. Q Appl Math. 1987;45:561–71.
31. Hardy RL. Multiquadric equations of topography and other irregular surfaces. J Geophys Res. 1971;76:1905–15.
32. Jin R, Chen W, Simpson T. Comparative studies of metamodelling techniques under multiple modelling criteria. Struct Multidiscip Optim. 2001;23(1):1–13.
33. Simpson TW, Peplinski JD, Koch PN, Allen JK. Metamodels for computer-based engineering design: survey and recommendations. Eng Comput. 2001;17:129–50.
34. Hussain MF, Barton RR, Joshi SB. Metamodeling: radial basis functions, versus polynomials. Eur J Oper Res. 2002;138:142–54.
35. Mullur AA, Messac A. Metamodeling using extended radial basis functions: a comparative approach. Eng Comput. 2006;21:203–17.