Materials & Design 215 (2022) 110428

Contents lists available at ScienceDirect

Materials & Design

journal homepage: www.elsevier.com/locate/matdes

Optimisation of the hydrogen bake-out treatment in steels via Gaussian processes

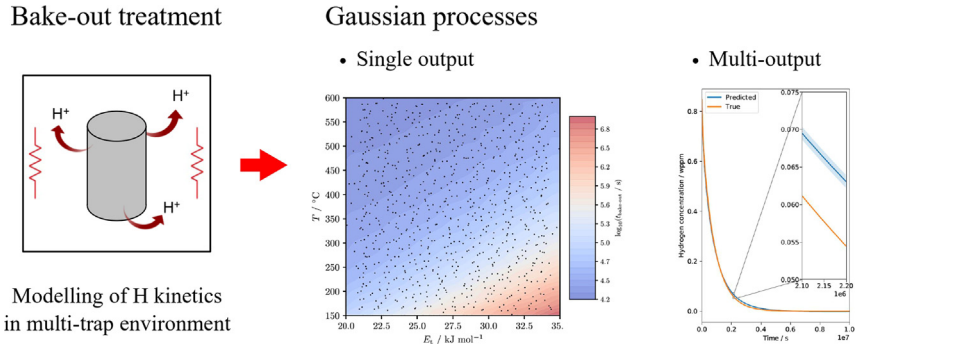
Edmund Cheung [a], Fernando D. León-Cázares [a,*], Enrique I. Galindo-Nava [a,b,*]

[a] Department of Materials Science and Metallurgy, University of Cambridge, 27 Charles Babbage Rd, Cambridge CB3 0FS, UK
[b] Department of Mechanical Engineering, University College London, Torrington Place, London WC1E 7JE, UK

HIGHLIGHTS

- Gaussian process surrogate models optimise bake-out treatments of steel with hydrogen traps.
- Incorporating one dimension at a time results in better fits during training of the models.
- Single output models with experimental design accurately predict optimal bake-out times.
- Permutations of the training sets with respect to trapping features shorten the training times.
- Multi-output models via principal component analysis predict the hydrogen content evolution.

GRAPHICAL ABSTRACT



ARTICLE INFO

Article history:
Received 21 September 2021
Revised 11 January 2022
Accepted 26 January 2022
Available online 01 February 2022

Keywords:
Hydrogen diffusion
Steel
Bake-out heat treatment
Gaussian processes
Principal component analysis

ABSTRACT

The presence of hydrogen in structural alloys reduces their ductility, a phenomenon called hydrogen embrittlement. Bake-out heat treatments are employed during processing to allow hydrogen trapped in microstructural features to effuse from the samples, but the optimal times and temperatures depend on the kinetics of hydrogen diffusion in the material. In this work, Gaussian process surrogate models are employed to emulate the outputs of microstructure-sensitive diffusion differential equations in steel. Training the models by sequentially increasing the number of dimensions results in better performances and shorter training times. Two main approaches are developed: single output models with experimental design for the prediction of optimal bake-out times, and multi-output principal component analysis models for the prediction of hydrogen concentration evolution. A novel approach is implemented to shorten the training times of multi-trap models by exploiting the symmetry of the equations with respect to different kinds of traps. The resulting models pave the way for the implementation of Gaussian processes on more computationally expensive diffusion simulations for the optimisation of heat treatments and other applications.

1. Introduction

Hydrogen causes degradation of mechanical properties in steels, which is known as hydrogen embrittlement (HE). This leads to components failing in a brittle manner and without evident warning signs. To avoid this, it is important to consider the kinetics

* Corresponding authors at: Department of Materials Science and Metallurgy, University of Cambridge, 27 Charles Babbage Rd, Cambridge CB3 0FS, UK (E.I. Galindo-Nava).
E-mail addresses: fdl22@cam.ac.uk (F.D. León-Cázares), e.galindo-nava@ucl.ac.uk (E.I. Galindo-Nava).

https://doi.org/10.1016/j.matdes.2022.110428
0264-1275/

of hydrogen diffusion, which is affected by the microstructure due to preferential segregation to defects such as dislocations, grain boundaries, precipitates, etc. [1], also known as trapping. To tackle this, safety critical components, such as welded steels used in nuclear reactor pressure vessels, undergo 'bake-out' treatments, which allow hydrogen to effuse out of the component and suppresses trapping at defects. Therefore, research on the modelling and optimisation of the bake-out process must incorporate microstructural trapping of hydrogen [1–3].

Gaussian processes (GPs) have been used extensively in machine learning applications, but have only recently had widespread adoption in materials science. Computationally expensive simulations have been modelled using surrogate GPs to predict materials properties and behaviour [4–6]. Advantages of GP models are the quantification of the uncertainty and the fast emulation time compared to that of larger and more complex physical simulations. In the context of hydrogen diffusion and trapping, these often require numerical integration to solve nonlinear diffusion equations [7]. To the knowledge of the authors, GPs have never been applied to the modelling of hydrogen diffusion in metals.

The current work presents a 'proof-of-concept' application of GP modelling to hydrogen diffusion and trapping in metallic systems. The goal is to develop numerically efficient data-based methods that can be implemented in the design of hydrogen bake-out treatments for structural components.Using data from computationally inexpensive simulations, multiple GP surrogate models are trained to predict the time taken for successful heat treatments and the evolution of hydrogen content in samples during degassing. We introduce novel approaches to model these phenomena, and compare their performance and speed in a systematic way. A case study on two steels is included to showcase how the methods developed can be applied to optimise bake-out treatments.

2. Methods

2.1. Hydrogen diffusion modelling

Modelling of hydrogen diffusion requires the modification of Fick's laws to explicitly implement the effects of trapping. This is to make the distinction between the solute dissolved in normal lattice sites, also known as diffusible hydrogen, and that trapped at microstructural features or crystal defects. HE is widely recognised to besensitive to the former[8], but the overall kinetics are slowed down by the trapping and detrapping processes[9,10]. Note that hydrogen accumulation at grain boundaries can also affect the embrittlement behaviour, a phenomenon known as hydrogen-enhanced decohesion [11].

The model derived by Turk et al. [3] is adopted to perform the simulations of hydrogen kinetics. This is based on two main assumptions. Firstly, traps are considered to be point defects. Secondly, Oriani's equilibrium approximation [1] ensures that a local thermodynamic equilibrium is maintained between lattice and trap sites. Such equilibrium reads

$$\frac{\theta_l}{1-\theta_l} = \frac{\theta_t}{1-\theta_t} K, \tag{1}$$

where the subscripts l and t denote the lattice and trapping sites, respectively, θ is the fractional occupancy of the corresponding sites, K the equilibrium constant

$$K = \exp\left(-\frac{\Delta E_t}{RT}\right), \tag{2}$$

R the molar gas constant, T the absolute temperature and ΔE_t the trapping energy. The occupancy $\theta_i = c_i/N_i$ can also be expressed in terms of the concentration c_i and site density N_i of each site type i. Eq. (2) was derived more rigorously by Svoboda and Fischer [12], who formulated the Gibbs free energy for arbitrarily many trapping sites and minimised it using the method of Lagrange multipliers to arrive at Eq. (2) [13]. The local equilibrium assumption is a valid approximation for a variety of defects, because the detrapping kinetics are often faster than the lattice hydrogen diffusion, except for very strong traps[14], defined by a large value of ΔE_t. Therefore, models of hydrogen diffusion in steel will often use this assumption [15,16].

For a material with m traps, Fischer et al. [13] showed that Eq. (1) can be rewritten as

$$\theta_{tk} = \frac{\theta_l}{K_k + \theta_l(1-K_k)}, \tag{3}$$

where the index k (between 1 and m) denotes the individual trap. The overall hydrogen concentration c is the sum of the lattice and trapped hydrogen concentrations

$$c = c_l + \sum_{k=1}^{m} c_{tk}. \tag{4}$$

With the above considerations, Turk et al. [3] derived the diffusion equation used in this work. The one-dimensional isothermal form of the differential equation is

$$\frac{\partial c_l}{\partial t} = A^{-1}\frac{\partial}{\partial x}\left(D\frac{\partial c_l}{\partial x}\right), \tag{5}$$

where the lattice concentration $c_l(x,t)$ is a function of distance x and time t, $D = D_0\exp(-Q/(RT))$ is the coefficient of diffusion of hydrogen in the lattice sites, D_0 its prefactor and Q the activation energy for diffusion. The parameter

$$A = 1 + \sum_{k=1}^{m}\left(\frac{N_{tk}}{N_l}\frac{K_k}{\left(K_k + \frac{c_l}{N_l}(1-K_k)\right)^2}\right) \tag{6}$$

is a tortuosity factor that characterises the convoluted pathways followed by the solute due to the presence of point traps, effectively slowing down the hydrogen diffusion rate across this heterogeneous medium. The reader is referred to the original work [3] for a complete derivation of equation (5). This was previously used to model hydrogen diffusion in ferritic VC-containing steels [3], Ni-Fe-Cr alloys [17], martensite [2] and ferritic-austenitic steels [7].

Eq. (5) is implemented in this work using an explicit finite difference method (FDM) [18] for a thin plate geometry of thickness $2L$ where hydrogen effuses from both surfaces. The simulation takes an initial uniform lattice hydrogen concentration, $c_l(x,0) = c_{l0}$, and simulates the evolution of the hydrogen concentration profile over the interval $[0,L]$ along x, as the sample is symmetric about L. The boundary conditions are zero solute at the free surface and zero flux at the symmetry plane, *i.e.*

$$c_l(0,t) = 0 \tag{7a}$$

$$\left.\frac{\partial c_l}{\partial x}\right|_{x=L} = 0. \tag{7b}$$

Fig. 1 shows a schematic diagram with the geometry of the problem and the expected evolution of the lattice concentration profiles. The spatial domain is discretised using a second-order central difference, with a mesh size of $\Delta x = L/20$. The time step size Δt is chosen by considering the von Neumann criterion for numerical stability of this FDM. For differential Eq. (5), this becomes

$$A^{-1}D\frac{\Delta t}{\Delta x^2} \leqslant \frac{1}{2}, \tag{8}$$

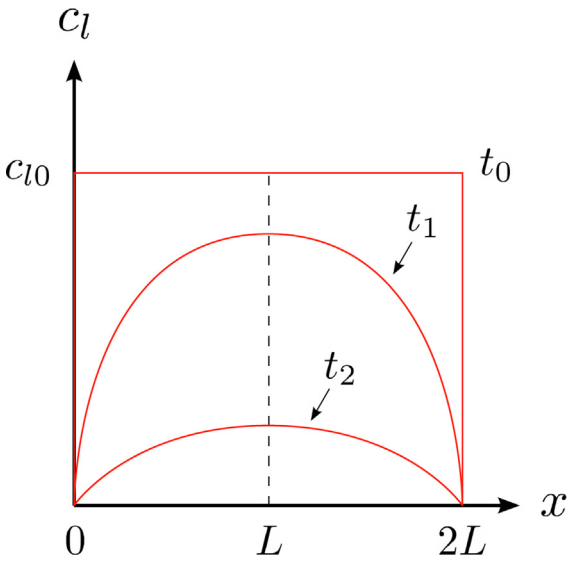


Fig. 1. Geometry of the diffusion problem showing the evolution of the lattice concentration profiles ($t_0 < t_1 < t_2$). The dashed line represents the symmetry plane of the sample.

where the value of A in Eq. (6) was calculated with $c_l = c_{l0}$ to ensure convergence at all times.

A criterion is required for the optimisation of the bake-out treatment. HE is sensitive to the lattice hydrogen, and studies have shown that there is a threshold concentration above which embrittlement becomes significant [19–22]. Nonetheless, the bake-out treatment also targets the effusion of trapped hydrogen, as otherwise this can diffuse back into the matrix and cause HE during in-service life. Thus, the total hydrogen concentration c_{total}, defined as the integral of c from Eq. (4) over the sample thickness, is numerically computed after every time step of the simulation. The bake-out time $t_{\text{bake-out}}$ is then the time taken for c_{total} to fall below a threshold value c_{thresh}, at which point the simulation is stopped. A value of $c_{\text{thresh}} = 1$ wppm is chosen for validation purposes. The specific threshold is in reality sensitive to chemistry, microstructure and service conditions. Values reported for the onset of HE vary from less than 1 wppm [19,21] to 5–10 wppm [20,22]. In this work, this variable is considered a 'design' parameter, as the analysis focusses on the overall behaviour of $t_{\text{bake-out}}$ for given input parameters.

2.2. Gaussian process modelling

The GP methods covered in this work are regression and surrogate modelling, with experimental design, a methodology to systematically choose the training points more efficiently (see Section 2.2.1 for a more detailed description). These are employed to evaluate the output of underlying physical simulations in a computationally cheaper way. A prior distribution over continuous functions is assumed to be a Gaussian Process, i.e.

$$\mathbf{f}|\mathbf{X} \sim \mathcal{N}(\boldsymbol{\mu}, \mathbf{k}), \tag{9}$$

where $\mathbf{f} = (f(\mathbf{x}_1), f(\mathbf{x}_2), \ldots, f(\mathbf{x}_n))$ is a vector of outputs from n training points, $\mathbf{X} = (\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n)^T$ a matrix of inputs, $\boldsymbol{\mu} = (\mu(\mathbf{x}_1), \mu(\mathbf{x}_2), \ldots, \mu(\mathbf{x}_n))$ a mean vector for some function $\mu(\mathbf{x})$, and $\mathbf{k}$ the covariance matrix, defined by the covariance function $k(\mathbf{x}_i, \mathbf{x}_j)$. We can make the standard assumption of a mean function of zero without loss of generality. For a prediction of $f(\mathbf{x}_*)$ at a new unseen data point $\mathbf{x}_*$, the posterior mean and variance conditioned on the observations $\mathbf{y} = (y_1, y_2, \ldots, y_n)$ at training inputs $\mathbf{X}$ can be found as

$$y_* = \mathbf{k}(\mathbf{x}_*)^T \mathbf{K}^{-1} \mathbf{y}, \tag{10a}$$

$$\sigma_*^2 = k(\mathbf{x}_*, \mathbf{x}_*) - \mathbf{k}(\mathbf{x}_*)^T \mathbf{K}^{-1} \mathbf{k}(\mathbf{x}_*), \tag{10b}$$

respectively, where $\mathbf{k}(\mathbf{x}_*) = (k(\mathbf{x}_*, \mathbf{x}_1), \ldots, k(\mathbf{x}_*, \mathbf{x}_n))^T$, $\mathbf{K} = \mathbf{k} + \sigma_{\text{noise}}^2 \mathbf{I}$, σ_{noise} is the noise, and $\mathbf{I}$ the $n \times n$ identity matrix. In this work, a squared exponential covariance function is used

$$k(\mathbf{x}_p, \mathbf{x}_q) = \sigma_f^2 \exp\left(-\frac{1}{2} \sum_{i=1}^{d} \left(\frac{x_p^{(i)} - x_q^{(i)}}{\ell^{(i)}} \right)^2 \right), \tag{11}$$

where σ_f is a scaling term for the variance and $\boldsymbol{\ell} = \left(\ell^{(1)}, \ldots, \ell^{(d)}\right)$ a vector with the 'length scale' parameters for the d dimensions of the $\mathbf{x}$ input vectors. To choose the hyperparameters of our model we consider the log likelihood

$$\log p(\mathbf{y}|\mathbf{X}, \boldsymbol{\ell}, \sigma_f) = -\frac{1}{2} \mathbf{y}^T \mathbf{K}^{-1} \mathbf{y} - \frac{1}{2} \log |\mathbf{K}| - \frac{n}{2} \log 2\pi \tag{12}$$

and optimise it over the hyperparameters [23].

Different models were implemented and compared. These are presented in Table 1 and detailed in the subsections below. All GP models were performed using GPy [24], a Python package that provides a framework to implement GPs. The values of σ_f and the length scales of ℓ were all initialised to unity, and the noise was fixed at a small value $\sigma_{\text{noise}} = 10^{-5}$ to prevent fitting issues when performing GP regression[25]. Fitting of all models was quantified from the true (directly from simulations) and predicted outputs, denoted y_i^{true} and y_i^{pred}, respectively, using the coefficients of determination

$$R^2 = 1 - \frac{\sum_{i=1}^{n} \left(y_i^{\text{true}} - y_i^{\text{pred}} \right)^2}{\sum_{i=1}^{n} \left(y_i^{\text{true}} - \bar{y} \right)^2}, \tag{13}$$

where

$$\bar{y} = \frac{1}{n} \sum_{i=1}^{n} y_i^{\text{true}}, \tag{14}$$

and the root-mean-square errors (RMSE)

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^{n} \left(y_i^{\text{true}} - y_i^{\text{pred}} \right)^2} \tag{15}$$

2.2.1. Single output models

GP regression surrogate models were developed with the bake-out time $t_{\text{bake-out}}$ as a univariate output. Experimental design was used to efficiently sample the input space. This works by building GP surrogate models for the output of a simulation, denoted Y. Therefore, the objective is to find a function f such that

Table 1
Overview of the surrogate models developed.

Model	Inputs	Output
Single output		
• Single trap	$T, Q, E_t, N_t, L, c_{l0}$	$t_{\text{bake-out}}$
- All-in-one		
- Dimension-by-dimension		
• Multi-trap	$T, Q, \mathbf{E}_t, \mathbf{N}_t, L, c_{l0}$	$t_{\text{bake-out}}$
Multi-output	$T, Q, E_t, N_t, L, c_{l0}$	$c_{\text{tot}}(t)$

$f : \mathbf{x} \rightarrow Y(\mathbf{x})$, where $Y(\mathbf{x}) \in \mathbb{R}$ is a univariate output of the simulation for the given input $\mathbf{x} \in \mathscr{X} \subseteq \mathbb{R}^D$, $\mathscr{X}$ is a D-dimensional domain of interest and $\mathbf{x}$ are locations in this D-dimensional space [26]. GP surrogate models can then be trained on the input–output pair from the simulations outlined in Section 2.1. Sequential design, as outlined by Gramacy [27], selects the inputs $\mathbf{x}$ for the simulations in order to sample $\mathscr{X}$ efficiently, *i.e.* sampling as few points as possible while attaining the greatest reduction in the uncertainty of the surrogate model.

The experimental design by Seo et al. [28] is employed in the current work. This translates the 'active learning MacKay' (ALM) method, originally devised for neural networks [29], into GPs by means of the acquisition function

$$J(\mathbf{x}) = \sigma_n^2(\mathbf{x}), \tag{16}$$

where the variance $\sigma_n^2(\mathbf{x})$ is calculated from Eq. (10b) using the surrogate model training data $\{X_n, Y_n\}$. The subscript n in this context denotes the variance calculated from n data points. The point with the highest predictive variance is chosen as the subsequent training point,

$$\mathbf{x}_{n+1} = \text{argmax}_{\mathbf{x}\in\mathscr{X}} \sigma_n^2(\mathbf{x}). \tag{17}$$

This is then evaluated using the simulation as $y_{n+1} = Y(\mathbf{x}_{n+1})$ to form the new dataset $\{X_{n+1}, Y_{n+1}\}$. The process is then repeated with Eqs. (16) and (17) for a total of n_{train} steps. A flow chart of the training algorithm for the single output model is shown in Fig. 2(a).

Simulations were initially performed considering a single trap. This applies when a microstructural feature is responsible for most of the trapping in an alloy, or when multiple traps have overlapping (or similar) trapping energies. An extension is introduced later to deal with the multi-trap case.

Single output single trap models have input vectors $\mathbf{x_i} = (T_i, Q_i, E_{t,i}, N_{t,i}, L_i, c_{l0,i})$. Two approaches were followed to train such 6-parameter models: 'all-in-one' (a-in-o) and 'dimension-by-dimension' (d-by-d). In the former scenario, all dimensions were varied simultaneously from the beginning as is typically done in other applications [5,6,30]. The latter approach initially varies only one of the parameters to estimate its hyperparameter, whilst keeping all others constant with values equal to the mean of their corresponding ranges. Then, it incorporates an additional variable but uses the hyperparameter learned from the lower dimension model. This sequence is repeated until all the variables are incorporated into the analysis and their hyperparameters are estimated. Two different orders for the implementation of the parameters were tested. Both approaches were performed with the same number of training points. The d-by-d resulted in a considerably better fitting (as shown in Section 3), so this method was chosen for all following models.

The experimental design loop was initialised using 1000 points (except for 1- and 2-parameter models, which used 100 points) with Latin hypercube sampling (LHS) to ensure uniformity [5,31]. Training was performed with the Python package Emukit [32] running on GPy. The training and test sets were obtained from separate LHS instances.

For materials with non-overlapping trap energies, GPs with additional dimensions are needed. For example, a model with three traps requires four extra dimensions compared to the single trap case (two new trap energies and densities), which would drastically increase the computational time following the same d-by-d method. A clever approach is sought to address this, stemming from the fact that all traps are incorporated into the diffusion Eq. (5) in the same way.

A single output multi-trap (with three different traps) model is built, as an extension from the single trap case, exploiting the underlying symmetry of the diffusion equation with regards to all trap energies and site number densities. The input vectors are $\mathbf{x_i} = (T_i, Q_i, \mathbf{E_{t,i}}, \mathbf{N_{t,i}}, L_i, c_{l0,i})$, with trap vectors $\mathbf{E_{t,i}} = (E_{t,1,i}, E_{t,2,i}, E_{t,3,i})$ and $\mathbf{N_{t,i}} = (N_{t,1,i}, N_{t,2,i}, N_{t,3,i})$. The model varies the 10 dimensions simultaneously setting the initial characteristic lengths of all trap

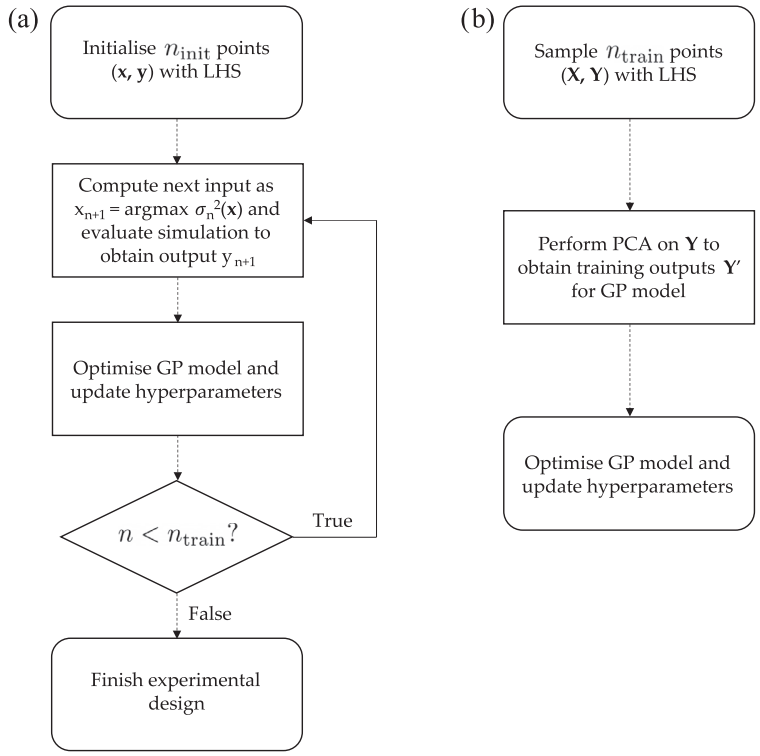


Fig. 2. Flow charts of theGP model training processes for (a) single output and (b) multi-output models.

energies and densities, and the remaining parameters, equal to those previously obtained for the single trap case, *i.e.* $l_{E_{t,1}} = l_{E_{t,2}} = l_{E_{t,3}} = l_{E_t}$ and $l_{N_{t,1}} = l_{N_{t,2}} = l_{N_{t,3}} = l_{N_t}$. For the initialisation stage, a set of n_{init} training points are obtained via LHS and these are simulated. The number of points at this stage can be multiplied (sixfold for three traps) without additional simulations by permuting the traps, given that any combination of a set of input trap parameters in Eq. (5) should output the same result. This makes the training stage considerably faster and ensures that the characteristic lengths of all traps remain equal to each other. The subsequent training stage with experimental design is then performed in the same way as for previous dimensions in the d-by-d approach, setting different ranges for the individual traps.

2.2.2. Multi-output model

This GP regression surrogate model attempts to capture the evolution of the total hydrogen concentration $c_{\text{total}}(t)$, simulated for $M = 10^5$ time steps Δt, via principal component analysis (PCA). This is achieved by applying a singular value decomposition (SVD) to a concentration matrix $\mathbf{c} \in \mathbb{R}^{(M\times N)}$ [5,33] used as training data, where N is the total number of training instances. For each row in the concentration matrix, the outputs are centred by subtracting the row's mean and dividing by its standard deviation. This gives a centred and scaled output

$$\bar{\mathbf{c}} = (\bar{c}^{(1)}, \ldots, \bar{c}^{(N)}) \in \mathbb{R}^{M\times N}. \tag{18}$$

PCA is performed by defining the matrix $\mathbf{W} \in \mathbb{R}^{M\times M}$ as the linear transformation

$$\mathbf{W}\bar{\mathbf{c}} = \mathbf{Y}, \tag{19}$$

where $\mathbf{Y}$ is the principal component (PC) score matrix. To compute the matrix $\mathbf{W}$,

$$\bar{\mathbf{c}} = \mathbf{U}\boldsymbol{\Sigma}\mathbf{V}^T, \tag{20}$$

which consists of orthogonal matrices $\mathbf{U} \in \mathbb{R}^{M\times M}$ and $\mathbf{V} \in \mathbb{R}^{N\times N}$, as well as the matrix $\boldsymbol{\Sigma} \in \mathbb{R}^{M\times N}$. Multiplying Eq. (20) by $\mathbf{U}^{-1}$ and using the fact that $\mathbf{U}$ is orthogonal, such that $\mathbf{U}^{-1} = \mathbf{U}^T$, one obtains

$$\mathbf{U}^T\bar{\mathbf{c}} = \boldsymbol{\Sigma}\mathbf{V}^T. \tag{21}$$

Thus, $\mathbf{W} = \mathbf{U}^T$ and $\mathbf{Y} = \boldsymbol{\Sigma}\mathbf{V}^T$.

The PCAs were truncated to the first $M\prime = 5$ PCs, capturing 99.99% of the cumulative explained variance and reducing the basis to $\mathbf{W}' \in \mathbb{R}^{M\prime\times N}$. Dimension reduction of the output matrix makes the problem tractable, although the PCs do not necessarily imply any physical meaning [34]. The GP regression can then be performed on the corresponding PC scores $\mathbf{Y}' \in \mathbb{R}^{M\prime\times N}$.

Experimental design was not employed in this model due to computational efficiency. This would require performing PCA after every training point is computed, with a complexity that scales with the size of the training data as $\mathcal{O}(MN^2)$ [35]. Instead, all the training and test points were sampled via LHS, followed by a single optimisation stage. The d-by-d approach was used with 500 training points for the first dimension and 2000 for the rest, as it is easier to train for lower dimensions. This model only used a single trap type. A flow chart of the multi-output model is shown in Fig. 2 (b).

2.2.3. Parameter definition

The parameters and their ranges for the GP models are shown in Table 2. The lattice diffusion activation energy selected corresponds to the range of values reported for ferrite [36]. The temperature range is characteristic of bake-out treatments, seeking to avoid phase transformations in steel, and the thicknesses chosen cover a wide range of components for different applications. The range of the initial hydrogen concentration includes various values used in the literature [2,3,7,17]. For the multi-trap case, the ranges of the trapping parameters correspond to grain boundaries, dislocations and regions of retained austenite, respectively [2]. For the single trap case, the trapping energy values chosen correspond to grain boundaries and dislocations.

Table 2
Parameters and ranges of the values used for the single output and multi-output models. The first column indicates the order in which parameters are added for training

#	Parameter	Range
1	T/°C	150–600
2	Q/kJ mol^{-1}	3.85–5
3	E_t/kJ mol^{-1}	20–35
4	N_t/mol m^{-3}	0–500
5	L/mm	1–200
6	c_{l0}/ mol m^{-3}	0–10
7[a]	$E_{t,1}$/kJ mol^{-1}	20–35
7[a]	$E_{t,2}$/kJ mol^{-1}	20–23
7[a]	$E_{t,3}$/kJ mol^{-1}	50–60
7[a]	$N_{t,1}$/mol m^{-3}	0–100
7[a]	$N_{t,2}$/mol m^{-3}	0–9000
7[a]	$N_{t,3}$/$\times 10^{-3}$ mol m^{-3}	0–5

[a] Only for the single output multi-trap case

3. Results

3.1. Single output models

Single output models focus merely on process optimisation, predicting only the recommended time for a succesful bake-out treatment. A detailed summary of the training steps used, performance and broken down computation times are shown in Table 3. Sample, optimise and acquisition times refer to those running the hydrogen diffusion simulations, optimising and updating the hyperparameters of the GP model, and optimising over the acquisition function, respectively, for the corresponding $n_{\text{init}} + n_{\text{train}}$ points. For the d-by-d models, the times shown refer to those spent upon adding each individual dimension.

The a-in-o approach fitted very poorly, with an R^2 close to zero. This is due to the large size of the multidimensional domain over which it operates. Note that the poor fitting is independent of the way in which the data was obtained, as the acquisition function employed avoids this. Thus, a better fit could only be obtained by drastically increasing the number of training points. The d-by-d method showed higher R^2, indicating a better fit, for the same number of training points as in the a-in-o approach. All subsequent GP models were therefore trained using the d-by-d approach.

The effect of the order in which the parameters are added in the d-by-d approach was also investigated. The results of two different runs are shown in Fig. 3, where combination 1 follows the order given in Table 3 and combination 2 has those reverted. The order of the parameters only mildly affects the R^2 values (Fig. 3)) and has a negligible effect on the final length scales learned (Fig. 3 (b)). Thus, finely tuning the parameter order is deemed unnecessary.

The correlations between the true and predicted bake-out times of the single output d-by-d models are shown in Fig. 4. The plots show the results at the end of the individual training stages, each varying an additional dimension. A clustering of points is observed for shorter bake-out times. Moreover, increasingly poorer fits are observed as more parameters are added into the training of the model.

Table 3
Numbers of training steps used, performance and computation times of the single output models (s.t. = single trap, m.t. = multi-trap).

Model	Param.	n_{init}	n_{train}	R^2	RMSE	Training time/ s		
						Sample	Optimise	Acquisition
s.t., a-in-o	6[a]	1000	2000	−0.017	3.63×10^6	81.0	4333.9	1060.6
s.t., d-by-d	1	100	100	1.0000	1.23×10^2	4.9	1.3	8.3
s.t., d-by-d	2	100	100	0.9998	2.06×10^3	5.0	0.9	6.3
s.t., d-by-d	3	1000	1000	0.9993	2.29×10^4	61.4	1081.8	403.4
s.t., d-by-d	4	1000	1000	0.9893	1.27×10^5	66.3	1026.7	373.8
s.t., d-by-d	5	1000	1000	0.9769	3.02×10^5	76.4	1224.0	528.6
s.t., d-by-d	6	1000	2000	0.9072	6.78×10^5	146.1	4328.3	1406.3
m.t., d-by-d	10	1200	4200	0.8845	1.54×10^6	441.5	25620.1	6343.3

[a] All parameters are added simultaneously in the a-in-o model

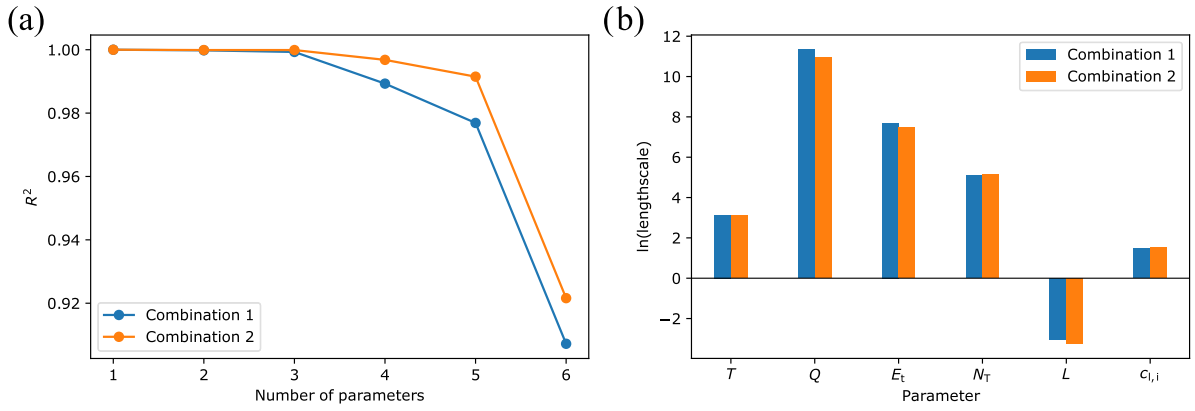


Fig. 3. Comparison of different combination of parameter orders. Figure (a) shows the variation of R^2 with the number of parameters used to train single trap type model and Figure (b) shows the learned length scales.

3.2. Multi-output model

The multi-output GPs offer additional information compared to the single output models. The PC scores are projected back to hydrogen evolution profiles, which could find potential applications in tracking the kinetics of the process rather than only the bake-out times. The numbers of training points, performance and training times of these models are shown in Table 4. Sample, PCA and optimise times refer to those running the hydrogen diffusion simulations, performing the PCA transformation on the training dataset, and optimising the GP model, respectively. Unlike in the single output case, the latter time here refers to the single optimisation step performed for all the training data.

While the aim of this work is to predict the hydrogen evolution profiles, it is useful to first investigate the performance of the GP models in the PC space. Both the training and validation data sets were projected onto the PC basis and residuals were computed for the predicted PC scores. The correlations between the true and predicted PC scores (in the 5 directions) of the multi-output model are shown in Fig. 5. The behaviours are similar to those observed for the single-output case, with increasingly larger errors for additional dimensions. The model with 6 parameters has a particularly poor R^2 score, likely due to the extra dimension compared to the previous case, so more training points would be required for a good fit.

The PC score prediction is not necessarily an indication of the model performance upon projecting back onto the concentration space via the inverse PCA transformation. Fig. 6 shows a representative set of predicted hydrogen evolution profiles reconstructed from the PC scores of the 5-parameter model, together with the results from the physical simulations. The plots correspond to the test points with the minimum, median and maximum RMSE, based on a test set of 1000 points. The regressions from the PC scores resemble the simulated hydrogen concentration curves, and larger uncertainties are localised around points with different locations in each simulation. The minimum error profile, in Fig. 6 (a), shows accurate fitting within the model uncertainty. The median case, Fig. 6(b), shows relatively accurate fitting, although it does deviate from 1σ. Nevertheless, the overall behaviour of the true profile is reasonably reproduced in the predicted profile. In Fig. 6(c), the case with the maximum error exhibits high model uncertainty at low times, with noticeable localisation of high uncertainty regions.

4. Discussion

4.1. Assessment of GP models

The a-in-o approach to GP modelling is by far the most widely used in the literature. The idea of running a single cycle that incorporates all dimensions at once is simpler to implement and requires less input from the user. However, this approach proves too computationally expensive for this application, as it works with a higher number of dimensions than many other GP models used in materials science (*e.g.* [5,30]). A considerably larger training dataset would be needed to improve the predictive capabilities of this model to reasonable levels of confidence.

The d-by-d approach devised and implemented in this study shows a better performance. This is likely due to the more reasonable length scales initialised for each new parameter, obtained from the simulations of lower dimension models. Effectively, only

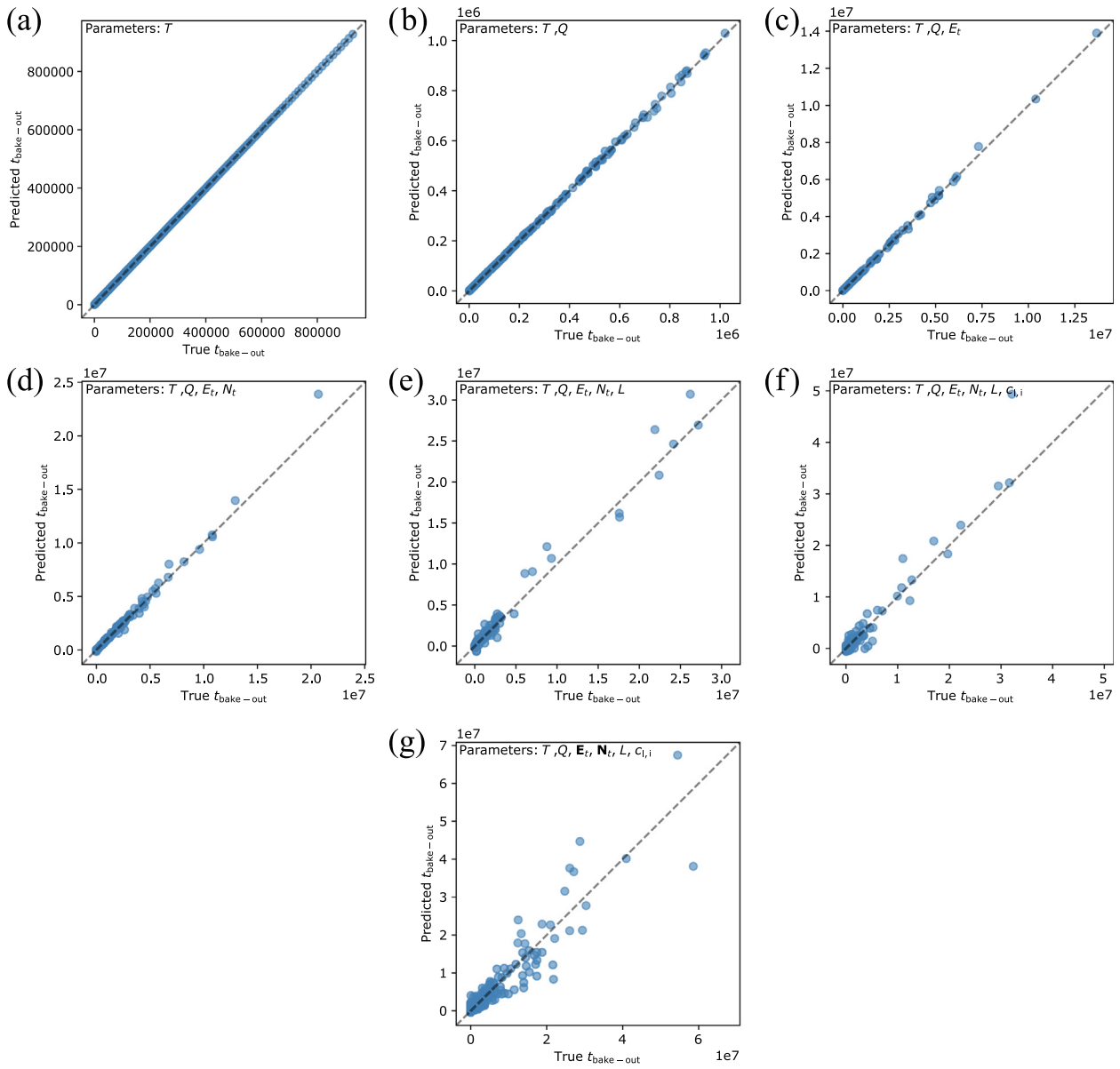


Fig. 4. Comparison between the predicted and true bake-out times for the single output (a-f) single trap models with 1–6 parameters, respectively, and the (g) multi-trap model with 10 parameters. The corresponding values of R^2 and RMSE are shown in Table 3.

Table 4
Numbers of training steps used, performance and computation times of the d-by-d multi-output model.

Param.	n_{train}	R^2	RMSE	Training time/ s		
				Sample	PCA	Optimise
1	500	0.9851	2.29×10^1	4.9	1.3	8.3
2	2000	0.9822	1.89×10^1	4349.1	29.8	249.3
3	2000	0.9273	4.92×10^1	4473.0	28.7	236.5
4	2000	0.9663	3.35×10^1	5080.6	33.6	553.0
5	2000	0.9381	3.52×10^1	5042.9	33.2	664.1
6	2000	0.7838	8.48×10^1	4647.7	36.0	492.1

one length scale is fully optimised from scratch for each dimension added, whilst the other ones are only fine tuned. However, it is not straightforward rationalising why this approach is successful. Whilst a formal proof is beyond the scope of the current work, it is worth scrutinising the diffusion equation used (Eq. (5)). This shows a nonlinear behaviour due to Oriani's equilibrium approximation in Eq. (1). Nonlinearities arise in particular as the concentrations in the traps tend towards their maximum values. Regardless, the effects of all parameters in the model are monotonic, so the output function ($t_{\text{bake-out}}$ or $c_{\text{total}}(t)$ for the single or multi-output models, respectively) has no local extrema. Thus, varying an additional parameter in the equations changes the

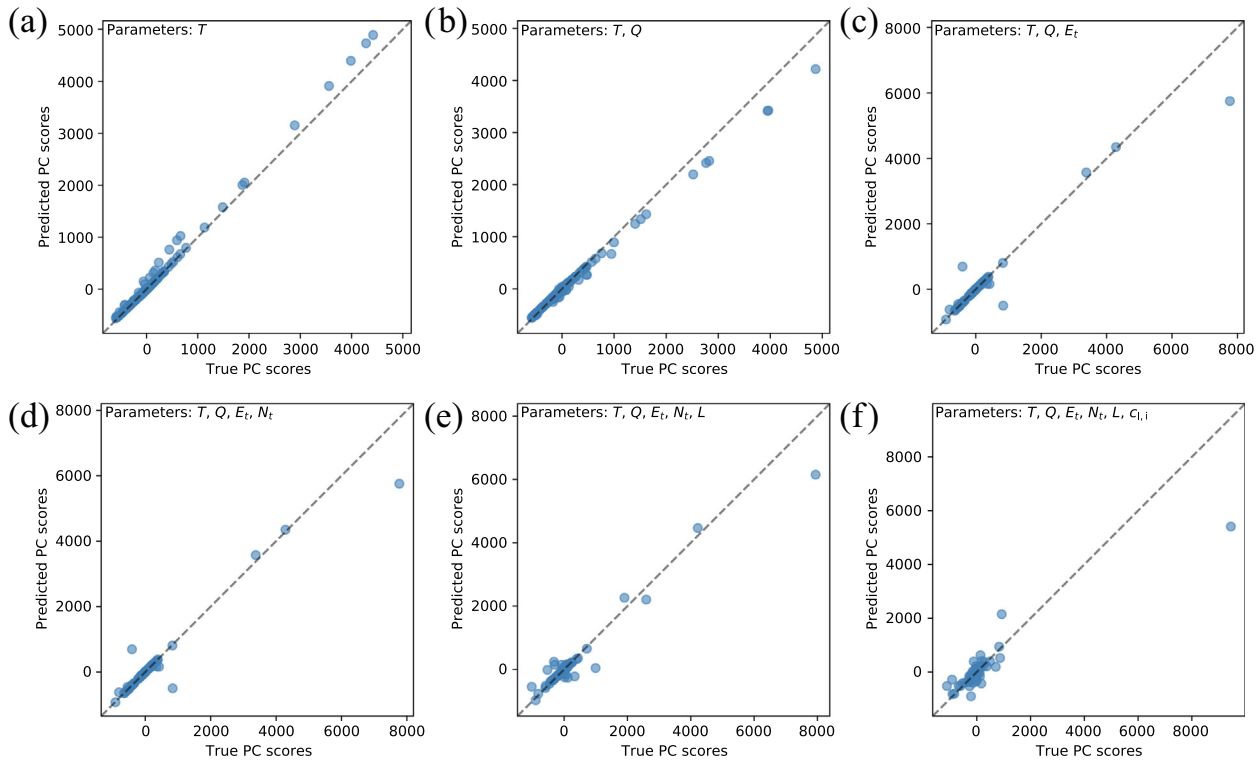


Fig. 5. Comparison between the predicted and true PC scores for the multi-output model with (a-f) 1–6 parameters, respectively. The corresponding values of R^2 and RMSE are shown in Table 4.

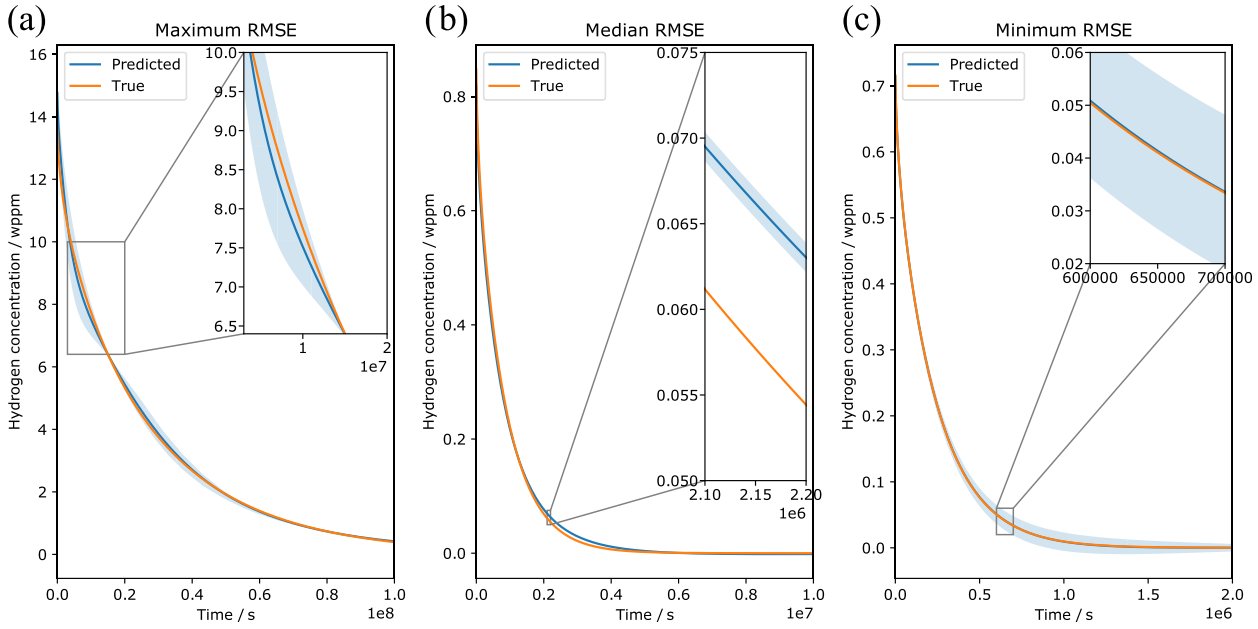


Fig. 6. True vs. predicted hydrogen evolution profiles for multi-output model trained on 5 parameters, showing the samples with (a) minimum (b) median and (c) maximum RMSE. Shaded regions show an estimate of the model variance as $\pm 1\sigma$ and insets show regions with high error.

kinetics of hydrogen diffusion but not the overall behaviour, and length scales learned remain similar as the dimensionality of the problem increases.

Clearly, the R^2 values decrease and the RMSE values increase with the number of parameters, indicating a poorer fit for large numbers of parameters, as shown in Table 3 and Fig. 4. This is expected, as for each additional parameter, the sampled data becomes sparser and the model less accurate in predicting from un-sampled regions. The clustering of points at low values of $t_{\text{bake-out}}$ for all cases can be attributed to the fact that only a small number of samples will have many parameter values at the extremes of their ranges. Only some combinations of the parameters result in very long times (*e.g.* low temperatures, small thicknesses, scarce and weak traps, etc.). Nonetheless, the use of the

acquisition function to obtain the training points makes sure that these regions of the input parameter domain are also sampled.

The multi-trap step is the most computationally expensive of the single output models, as shown in Table 3. This is mainly because of the optimising times, which scales with $\mathcal{O}(N^3)$. Yet, the trap permutation technique enables the implementation of four new dimensions at once, shortening the number of initialisation and training steps needed fourfold (assuming each new dimension uses the same number of training points). More training steps had to be used for the multi-trap model than for the lower dimension ones to obtain an $R^2 \sim 90$ %, but this difference would need to be considerably larger without the permutations implemented. This approach could find applications in other emulation GP models whenever there is a symmetry between two or more of the input parameters involved.

The multi-output model demonstrated its capability to predict the evolution of hydrogen concentrations during bake-out regimes. However, the overall performance for similar training times was worse than that for single output models. This is in large part due to having to simulate the differential equations for a set number of time steps, rather than by interrupting the function. Not having an experimental design stage makes this model considerably faster, but at the expense of not quantifying the variance during the training stage. Relying solely on LHS for sampling of the training points, regions of high variance may not be sampled, explaining the higher RMSE. Moreover, this approach is inherently more complex than the single output one due to having five PCs to train the model on. Upon performing the inverse PCA transformation, the overall shapes of the concentration evolutions are reproduced, but with some deviations between true and predicted values. Sources of error may be the inherent loss of information in the PCA transformation (as the PC score matrix is truncated) and insufficient sampling.

Overall, both single and multi-output models present advantages and disadvantages for their use in the optimisation of the bake-out treatment. The single output model can estimate the optimal bake-out times for a series of input parameters. Alternatively, the multi-output model gives additional information about the effusion kinetics at the expense of either a slower training stage or a lower regression score.

A more detailed analysis of the computing times of the individual routines is required to assess the scalability of the GP models developed. As mentioned before, the current work represents a 'proof-of-concept' of what can be achieved with these methods, and it currently presents little benefit from the direct numerical simulations of hydrogen kinetics with the differential equations. However, from Tables 3 and 4, one can see that the routine that takes the least time is the actual running of the simulations. Running computationally heavier models would increase this time, whilst other routines would see no or only small variations. Such models could be multiphase diffusion simulations in 2D or 3D, which can become considerably slower. The multi-output model has the potential to be extended into other simulation techniques, such as finite element methods. This could be attempted by applying dimension reduction to the concentration values associated with each mesh node, analogous to the work of Stowers et al. [5] in the context of stress–strain relationships.

4.2. Optimisation of the bake-out treatment

Designing and optimising the bake-out treatment involves selecting an adequate time and temperature to allow enough hydrogen to effuse from the sample and avoid HE, accounting for the material and processing parameters involved. In practice, the accuracy that can be achieved in the control or characterisation of such parameters varies to different degrees. For instance, adequate machining can ensure negligible calculation errors due to variations in the sample thickness, but trap parameters are difficult to characterise and may even vary within a single sample due to material heterogeneities. The accuracy of the diffusion equations and the computational efficiency of the GP models can be leveraged to assess the effects of such variations on the predicted bake-out times, especially in thick sections where the microstructure can vary significantly.

A single output d-by-d GP model was trained over three dimensions (T, E_t and N_t, in that order) to investigate the effects of heterogeneities in the trap properties. The training followed the same schedule as that in Section 3.2, with 100 initialisation and training points for the first two dimensions and 1000 training points for the third one. The ranges used for the variable dimensions were those in Table 2, and the constant parameters were set to $Q = 3.85$ kJ mol^{-1} and $L = 100$ mm. Rather than fixing the initial lattice concentration, all simulations started with a total concentration of $c_{\text{total},0} = 10$ mol m^{-3}, from which the lattice concentration was obtained numerically via Eqs. (1)–(4) to account for the variations in the hydrogen redistribution due to temperature. The concentration threshold of 1 wppm was also used here for consistency. The training process took 724.9 s and it resulted in a coefficient of determination of $R^2 = 0.9992$.

Useful information for the design and optimisation of the hydrogen bake-out treatment can be obtained by evaluating the trained GP model. Fig. 7 shows the optimal bake-out times predicted as a function of temperature and trap parameters. Note that the training points are not homogeneously distributed due to the nature of the acquisition function used to select them. As expected, longer treatments are required for lower temperatures, and higher trapping energies and trap site densities. The range of energies plotted in Fig. 7(a) covers many possible traps present in steels, such as dislocations, grain boundaries and some carbides. Alternatively, the trapping energy of $E_t = 20$ kJ mol^{-1} used as an example in Fig. 7(b) corresponds to that reported for dislocations in ferritic steels [2]. The trap density is in this scenario a function of the dislocation density, which varies widely for different forging routes. This contour plot shows that, for instance, if the bake-out time is set to one hour, the annealing temperature must increase in a non-linear form as N_t increases, *e.g.* as consequence of residual dislocations after forging. Plots like these ones can readily inform the manufacturing process of steel components in an easy and effective way.

The methods developed can also be directly applied to specific alloys, showing the robustness of the approach. Optimal bake-out treatments were calculated for two steels with trapping parameters reported in the literature. Firstly, ferritic-martensitic dual phase DP800 steel, ice-water quenched and tempered at 300 °C, where trapping with $E_t = 30$ kJ mol^{-1} and $N_t = 309$ mol m^{-3} occurs at the martensite interfaces [2]. Secondly, a model vanadium-rich ferritic steel annealed at 740 °C, where most hydrogen trapping occurs at the interface of vanadium carbides with $E_t = 26$ kJ mol^{-1} and $N_t = 73$ mol m^{-3} [3]. All other parameters are those used for the GP model above. The solid lines in Fig. 8 show the predicted bake-out times for these steels calculated directly with the diffusion equations. The effects of temperature and trapping parameters are evident, with heat treatment times that differ by orders of magnitude.

Additional data is needed to make sure that the hydrogen bake-out treatment is effective. As mentioned before, there is often uncertainty in the values selected for the material parameters involved in the simulation. The trapping energy is difficult to characterise accurately from experimental setups, and actual variations

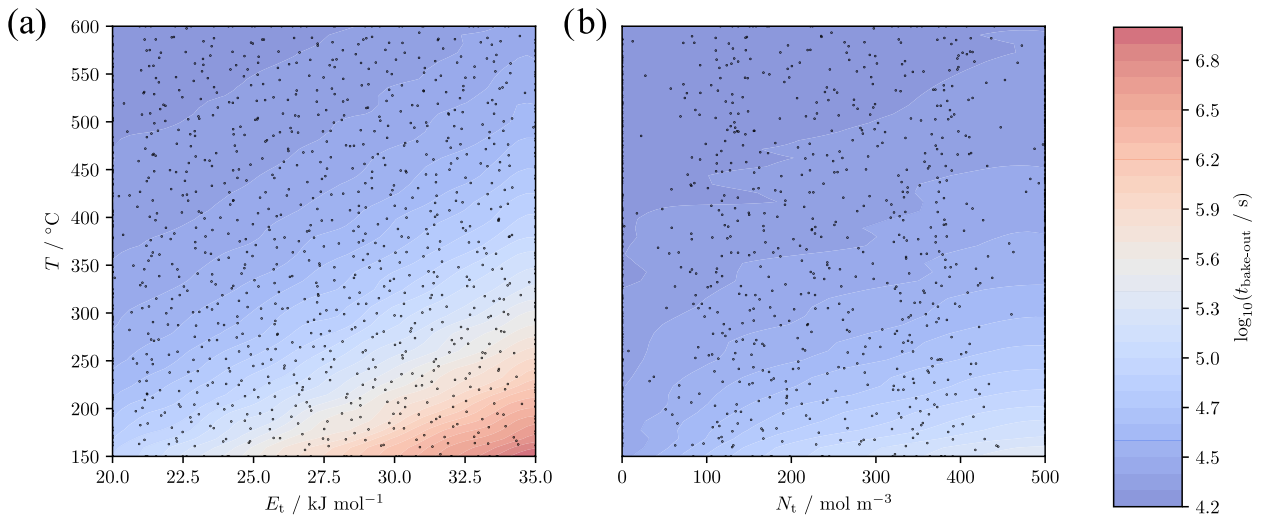


Fig. 7. Predicted bake-out times for the single output single-trap model as a function of temperature, and (a) trapping energy (with $N_t = 170$ mol m^{-3}) and (b) dislocation trap site density (with $E_t = 20$ kJ mol^{-1}). The training points are plotted for reference.

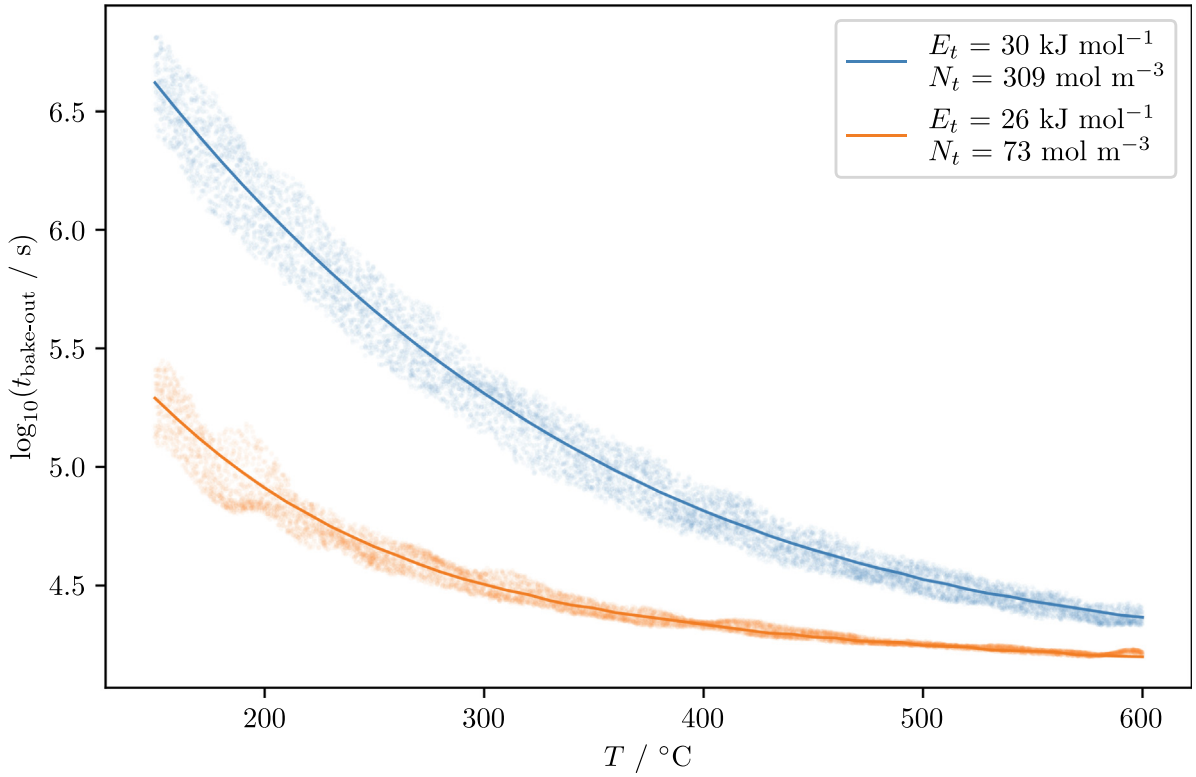


Fig. 8. Predicted bake-out times for two steels with different trapping parameters [2,3] as a function of temperature. The solid lines show the times calculated with the diffusion equations, and the dots are obtained by uniformly sampling from the GP model within a range with ±5% variations of the trap.ping parameters.

in the trap site densities may arise across engineering components due to heterogeneities in the distribution of defects. The above-mentioned GP model was used to examine the optimal bake-out treatments accounting for variations of ±5% of the trapping parameters reported. For each alloy, 10000 points were uniformly sampled within this range of values and across the temperature domain simulated, which resulted in the envelopes observed in Fig. 8. The fast computational times needed to evaluate the GP model, of about 1.3×10^{-4} s per data point (compared to 0.58 s per data point for the FDM) allow these results to be obtained quickly for this whole range, highlighting another advantage of using our data-based approach in heterogeneous materials over standard diffusion modelling methods. This plot shows an uneven spread in the bake-out times required, with the envelopes becoming wider at lower temperatures likely due to the exponential dependence of the trapping term in Eq. (2). Similarly, the spread is larger for the sample with larger trapping parameters (due to the logarithmic scale) for the same reason.

A plot like that in Fig. 8 can guide the design of heat treatments for engineering components to optimise energy consumption and

cost. An adequate characterisation of the hydrogen diffusion behaviour could offer a more accurate control over the confidence intervals for the selection of the heat treatment time and temperature. Note that the uniform sampling over the variations assumed for the trapping parameters was performed due to a lack of statistical data on the actual values. A more detailed characterisation of the mechanical behaviour of the alloy is also desired to determine a viable hydrogen concentration threshold. The framework introduced can also be further expanded to include more complex diffusion simulations. In all cases, the current work will serve as a reference for the computational times and predictive accuracies that can be achieved by applying GPs to the optimisation of the bake-out heat treatment.

5. Conclusions

Simulations of hydrogen diffusion can be used to create computationally inexpensive GP surrogate models that optimise the bake-out heat treatments of steel components. The methods implemented can have a direct impact in the design of more efficient processing routes, accounting for the effects of microstructural heterogeneities, this not being currently considered using standard diffusion/trapping models for hydrogen. Multiple approaches were developed and compared to each other. The main findings are the following:

- The models developed can be employed for a variety of alloys containing different types of hydrogen traps.
- Implementing one dimension at a time (d-by-d) results in better fits than incorporating all at once (a-in-o), likely due to the nature of the nonlinear diffusion equations used.
- A single output model with experimental design can be used to optimise the bake-out times. The incorporation of multiple traps can be sped up by exploiting the symmetries of the diffusion equations with respect to the trapping features using permutations of the training sets.
- A multi-output model that predicts the evolution of hydrogen content in a sample can be performed via principal component analysis, increasing the amount of information supplied by the models at the expense of additional computational time.

Author contributions

FDLC and EIGN conceptualised the research. EIGN acquired the funding. EC developed the methodology and carried out the modelling, under the supervision of FDLC and EIGN. EC and FDLC wrote the manuscript.

Data availability

The article has no raw data. The processed data cannot be shared at this time due to technical or time limitations.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

We gratefully acknowledge the funding received from the EPSRC (Grant No. EP/T008687/1). EIGN acknowledges funding from RAEng for his research fellowship.

References

[1] R.A. Oriani, The diffusion and trapping, Acta Metall. 18 (1) (1970) 147–157.
[2] A. Turk, G.R. Joshi, M. Gintalas, M. Callisti, P.E.J. Rivera-Díaz-del Castillo, E.I. Galindo-Nava, Quantification of hydrogen trapping in multiphase steels: Part I - Point traps in martensite, Acta Mater. 194 (2020) 118–133.
[3] A. Turk, D.S Martín, P.E.J. Rivera-Díaz-del Castillo, E.I. Galindo-Nava, Correlation between vanadium carbide size and hydrogen trapping in ferritic steel, Scripta Mater. 152 (2018) 112–116.
[4] Z. Zheng, B. Chen, X. Yanwen, N. Fritz, Y. Gurumukhi, J. Cook, M.N. Ates, N. Miljkovic, P.V. Braun, P. Wang, A Gaussian Process-Based Crack Pattern Modeling Approach for Battery Anode Materials Design, J. Electrochem. Energy Convers. Storage 18 (1) (2021) 1–9.
[5] C. Stowers, T. Lee, I. Bilionis, A.K. Gosain, A. Tepole, Improving reconstructive surgery design using Gaussian Process surrogates to capture material behavior uncertainty, J. Mech. Behav. Biomed. Mater. (2021) 104340.
[6] R. Saunders, C. Butler, J. Michopoulos, D. Lagoudas, A. Elwany, A. Bagchi, Mechanical behavior predictions of additively manufactured microstructures using functional Gaussian process surrogates, npj Comput. Mater. 7 (1) (2021) 81.
[7] A. Turk, S. Pu, D. Bombač, P.E.J. Rivera-Díaz-del Castillo, E.I. Galindo-Nava, Quantification of hydrogen trapping in multiphase steels: Part II – Effect of austenite morphology, Acta Mater. 194 (2020) 118–133.
[8] H.K.D.H. Bhadeshia, Prevention of Hydrogen Embrittlement in Steels, ISIJ Int. 56 (1) (2016) 24–36.
[9] R. Kirchheim, Lattice discontinuities affecting the generation and annihilation of diffusible hydrogen and vice versa, Philos. Trans. R. Soc. Math. Phys. Eng. Sci. 375 (2098) (2017) 20160403.
[10] F.D. León-Cázares, E.I. Galindo-Nava, General model for the kinetics of solute diffusion at solid-solid interfaces, Phys. Rev. Mater. 5 (12) (2021) 123802.
[11] M. Koyama, C.C. Tasan, E. Akiyama, K. Tsuzaki, D. Raabe, Hydrogen-assisted decohesion and localized plasticity in dual-phase steel, Acta Mater. 70 (2014) 174–187.
[12] J. Svoboda, F.D. Fischer, Modelling for hydrogen diffusion in metals with traps revisited, Acta Mater. 60 (3) (2012) 1211–1220.
[13] F.D. Fischer, J. Svoboda, E. Kozeschnik, Interstitial diffusion in systems with multiple sorts of traps, Modell. Simul. Mater. Sci. Eng. 21 (2) (2013) 025008.
[14] D. Bombac, I.H. Katzarov, D.L. Pashov, A.T. Paxton, Theoretical evaluation of the role of crystal defects on local equilibrium and effective diffusivity of hydrogen in iron, Mater. Sci. Technol. 33 (13) (2017) 1505–1514.
[15] M. Enomoto, D. Hirakami, T. Tarui, Modeling Thermal Desorption Analysis of Hydrogen in Steel, ISIJ Int. 46 (9) (2006) 1381–1387.
[16] S. Jothi, T.N. Croft, L. Wright, A. Turnbull, S.G.R. Brown, Multi-phase modelling of intergranular hydrogen segregation/trapping for hydrogen embrittlement, Int. J. Hydrogen Energy 40 (43) (2015) 15105–15123.
[17] A. Turk, D. Bombač, J.J Rydel, M. Ziętara, P.E.J. Rivera-Díaz-del Castillo, E.I. Galindo-Nava, Grain boundary carbides as hydrogen diffusion barrier in a Fe-Ni alloy: A thermal desorption and modelling study, Mater. Des. 160 (2018) 985–998.
[18] K.W. Morton, D.F. Mayers, Numerical Solution of Partial Differential Equations, Cambridge University Press, 2005.
[19] J. Capelle, J. Gilgert, I. Dmytrakh, G. Pluvinage, Sensitivity of pipelines with steel API X52 to hydrogen embrittlement, Int. J. Hydrogen Energy 33 (24) (2008) 7630–7641.
[20] T. Zakroczymski, A. Glowacka, W. Swiatnicki, Effect of hydrogen concentration on the embrittlement of a duplex stainless steel, Corros. Sci. 47 (6) (2005) 1403–1414.
[21] M. Wang, E. Akiyama, K. Tsuzaki, Effect of hydrogen on the fracture behavior of high strength steel during slow strain rate test, Corros. Sci. 49 (11) (2007) 4081–4097.
[22] D.R. Harries, G.H. Broomfield, Hydrogen embrittlement of steel pressure vessels in pressurised water reactor systems, J. Nucl. Mater. 9 (3) (1963) 327–338.
[23] C.E. Rasmussen, C.K.I. Williams, Gaussian Processes for Machine Learning, The MIT Press, 2005.
[24] The GPyOpt authors, Gpyopt: A bayesian optimization framework in python, 2016.
[25] S. Basak, S. Petit, J. Bect, E. Vazquez, Numerical issues in maximum likelihood parameter estimation for Gaussian process interpolation, in: 7th International Conference on machine Learning, Optimization and Data science (LOD 2021), 2021, p. 16.
[26] G. Tapia, S. Khairallah, M. Matthews, W.E. King, A. Elwany, Gaussian process-based surrogate modeling framework for process planning in laser powder-bed fusion additive manufacturing of 316L stainless steel, Int. J. Adv. Manuf. Technol. 94 (9–12) (2018) 3591–3603.
[27] R.B. Gramacy, Surrogates: Gaussian Process Modeling, Design, and Optimization for the Applied Sciences, Chapman and Hall/CRC, 2020.
[28] S. Seo, M. Wallat, T. Graepel, K. Obermayer, Gaussian process regression: active data selection and test point rejection, Proceedings of the IEEE-INNS-ENNS International Joint Conference on Neural Networks, IJCNN 2000. Neural Computing: New Challenges and Perspectives for the New Millennium, vol. 3, IEEE, 2000.
[29] D.J.C. MacKay, Information-Based Objective Functions for Active Data Selection, Neural Comput. 4 (4) (1992) 590–604.

[30] A. Figueroa, M. Göttsche, Gaussian processes for surrogate modeling of discharged fuel nuclide compositions, Ann. Nucl. Energy 156 (2021) 108085.
[31] M.D. Mckay, R.J. Beckman, W.J. Conover, A Comparison of Three Methods for Selecting Values of Input Variables in the Analysis of Output From a Computer Code, Technometrics 42 (1) (2000) 55–61.
[32] A. Paleyes, M. Pullin, M. Mahsereci, N. Lawrence, J. Gonzalez, Emulation of physical processes with Emukit, NeurIPS, 2019.
[33] S. Shalev-Shwartz, S. Ben-David, Understanding Machine Learning, Cambridge University Press, Cambridge, 2014.
[34] A. Géron, Hands-On Machine Learning with Scikit-Learn and TensorFlow: Concepts, Tools, and Techniques to Build Intelligent Systems, O'Reilly Media Inc, 2017.
[35] X. Li, S. Wang, Y. Cai, Tutorial: Complexity analysis of Singular Value Decomposition and its variants, 2019, pp. 1–12.
[36] K. Kiuchi, R.B. McLellan, The solubility and diffusivity of hydrogen in well-annealed and deformed iron, Acta Metall. 31 (7) (1983) 961–984.