Fast prediction of aquifer thermal energy storage: a multicyclic metamodelling procedure

Jérémy Rohmer, Antoine Armandine Les Landes, Annick Loschetter, Charles Maragna

Computational Geosciences (2023) 27:223–243
https://doi.org/10.1007/s10596-023-10192-8

ORIGINAL PAPER

Fast prediction of aquifer thermal energy storage: a multicyclic metamodelling procedure

Jeremy Rohmer[1] · Antoine Armandine Les Landes[1] · Annick Loschetter[1] · Charles Maragna[1]

Received: 24 October 2021 / Accepted: 18 January 2023 / Published online: 27 January 2023

Abstract
The metamodel-based approach (also referred to as the surrogate approach) is commonly applied to overcome the computational burden of numerical models that are used to simulate the evolution of reservoir fluids and pressures in response to any production scheme. In this study, we propose an adaptation of this approach for aquifer thermal energy storage (ATES) systems. ATES systems are characterized by cyclic loading/unloading production schemes, which result in a strong similarity in the dynamics of the intercyclic evolution of variables such as the temperature at the producer well. Instead of training several metamodels, i.e., one per cycle ("independent" metamodelling approach), we take advantage of the intercyclic similarity to train a single metamodel within the setting of multifidelity cokriging ("multicyclic" metamodelling approach). To explore the predictive performance of this approach, we applied a random subsampling validation approach multiple times to 300 simulation results of a realistic ATES system in the Paris basin by considering three characteristics, i.e., the minimum and maximum temperature, and the rate of temperature decrease at each cycle. Numerical experiments with varying training dataset sizes (from 33 to 66% of the total number of results) and using 100 test samples show that (1) the predictive error of the multicyclic metamodelling reaches lower levels (by 20–50%) than that of the independent approach; (2) this higher predictive performance is achieved while saving computational time cost because the training phase only needs a few tens of "complete" simulations (run over all cycles) together with a few hundreds of "partial" simulations (stopped at the first cycle); the latter simulations are less expensive to evaluate because of shorter simulated time.

Keywords Kriging metamodel · Multifidelity · Global sensitivity analysis · Aquifer thermal energy storage · Underground thermal energy storage

1 Introduction

Geothermal simulators are key components in the computation of the evolution of fluids and pressures within reservoirs in response to any production scheme (see, for instance, [1] for an overview of the main thermohydraulic geothermal simulators and their modelling capabilities). Despite the progress in numerical modelling, as well as in computational hardware and software developments, the computational time cost of a single model run can reach several hours or days. This computational burden poses severe difficulties when multiple calls to the reservoir model are necessary, for instance, for uncertainty quantification or history matching. To overcome this computational burden, a possible solution consists of the statistical analysis of precalculated high-fidelity simulation results to set up a cheap-to-evaluate metamodel (also referred to as a "proxy" or "surrogate" or "response surface"). The cheap-to-evaluate metamodel is then used in place of the long-running geothermal simulator. The implementation of these statistical methods has a long history in the petroleum industry; see, for instance, the key lessons drawn by Amudo et al. [2] and by Zubarev [3]. An extensive and comprehensive review of the different statistical methods is provided by Razavi et al. [4]. Examples of real case applications are provided by White et al. [5] for uncertainty analysis, Pratama & Supijo [6] for probabilistic geothermal resource capacity assessment, Limbeck et al. [7] for induced seismicity forecasting, Fursov et al. [8] for the Bayesian inverse problem and Manceau & Rohmer [9] for global sensitivity analysis.

In addition to the results of high-fidelity long-running numerical simulations, some advanced metamodelling procedures make the most of alternative information. This is the

✉ Jeremy Rohmer
j.rohmer@brgm.fr

[1] BRGM, 3 av. C. Guillemin, 45060 Orléans Cedex 2, France

case for the multifidelity metamodelling approach, where any results from approximate versions (named low fidelity) of the numerical geothermal simulator are assimilated. For instance, the approximate versions can correspond to numerical models with coarse mesh resolutions (see, e.g., [10, 11]) or simplified physics (see, e.g., [12]). This has positive implications in practice, as low-fidelity models are less expensive to run, which allows performing a larger number of simulations while gathering information that is useful to improve the predictive capability of the metamodel.

The present study builds upon a similar idea to predict the intercyclic evolution of key parameters for aquifer thermal energy storage (ATES). Typically, ATES is associated with a district heating network (DHN), and a large volume of underground water is mobilized in a cyclic manner by alternating the storage phase of low-carbon heat (e.g., waste or solar) and heat recovery (to shave the peaks in the thermal demand that would otherwise be covered by fossil fuels); see, for instance, the comprehensive introduction by Cabeza et al. [13] and the state of the art by Kallesøe et al. [14]. Figure 1(a) presents two examples of temperature evolution at the production well of the ATES system used as an application case in our study (described in full detail in Section 2). We note that there is a strong similarity in the dynamics of the intercyclic temperature evolution (though it differs, the "general" temporal pattern is reproduced at each cycle). This shows that there is a strong intercyclic dependence of the key characteristics, such as the temperature extrema at each cycle (Fig. 1b) or the rate of decrease (defined as the decay parameter of an exponentially decreasing function) at each cycle (Fig. 1c).

In this study, we aim to take advantage of this similarity to improve the prediction over all cycles within a multifidelity metamodelling approach. Our proposal is to link the fidelity level to the cycle index. In this approach, the lower fidelity level then corresponds to the variable of interest computed for the first cycle only (the simulation is a "partial" simulation, and it is stopped after computing the variable at the first cycle), and the higher fidelity level corresponds to the variable of interest computed for all cycles (the simulation is "complete" and run over all 16 cycles, as illustrated in Fig. 1). This means that the "partial" simulations provide results that are viewed as approximate versions of the results of "complete" simulations. The expected result is a reduction in computational burden as more model runs can be performed during the first cycles due to the lower computation time cost (because of shorter simulation times). Our objective is to develop and test the performance of this multifidelity approach, termed "multicyclic metamodelling". This will rely on cokriging metamodels [15] based on the design and analysis of computer experiments [16] and will be applied to an ATES project in the Paris basin.

This paper is organized as follows. Section 2 describes the ATES case, which motivates our developments, namely, a domestic waste incineration plant coupled to an interseasonal heat storage system within the Dogger deep aquifer in the Paris basin. In this section, we present the numerical model used to compute the variables of interest over 16 loading/unloading cycles as well as the input variables that are used to parametrize the setting of the ATES system. Section 3 provides the details of the different statistical methods used to set up the multicyclic metamodel. In Section 4, we apply the procedure and evaluate its capability to predict the variable of interest over the different cycles using an extensive random subsampling validation approach. The performance is compared to a procedure that uses multiple metamodels that are built independently at each cycle.

2 Application case

2.1 Context

The application case is located in the Dogger aquifer of the Paris basin at a depth of 1500–2000 m with rock formation temperatures at the top of the productive layers between 55 and 85 °C. This aquifer has been intensively used since the late 1970s to feed DHNs, mostly in the Paris conurbation [17]. Approximately 50 geothermal plants are currently operating.

Here, we study an ATES system coupled to a conventional plant feeding a DHN, as described by Réveillère et al. [18]. The excess heat (to be stored) originates from a domestic waste incineration plant. The system has two wells that operate alternately as producer and injector to store heat during the summer season and then to recover the heat during the winter season. In summer, the geothermal fluid circulates from the cold to the warm well, and then the circulation is reversed in winter. The loading/unloading cycle of thermal storage is a periodic sequence of 12 months (see an illustration in Fig. 2).

2.2 Setup of the numerical model

To predict the evolution of the interseasonal heat storage system, a numerical model is built based on the "Dogger database" [19] and the scheme of heat storage exploitation. The conceptual model used for hydrothermal modelling is based on a reservoir structure made of one or multiple layers with a constant thickness of 10 m for each structure (Fig. 3b).

In the model, a vertical plane of symmetry (Fig. 3a) in the middle of the aquifer is used (to reduce the computational time cost). Both wells (in the middle of the model) are separated by a distance D. The domain is a 3D model measuring 10 km by 5 km by 90 m in the X, Y and Z directions, respectively. The model is based on a 3D prismatic mesh (Fig. 3c) with 45 layers and a refined mesh close to the wells and comprises approximately 30,000 cells. The mesh is generated using the Computational Geometry Algorithms Library CGAL. The numerical simulations are performed using the ComPASS

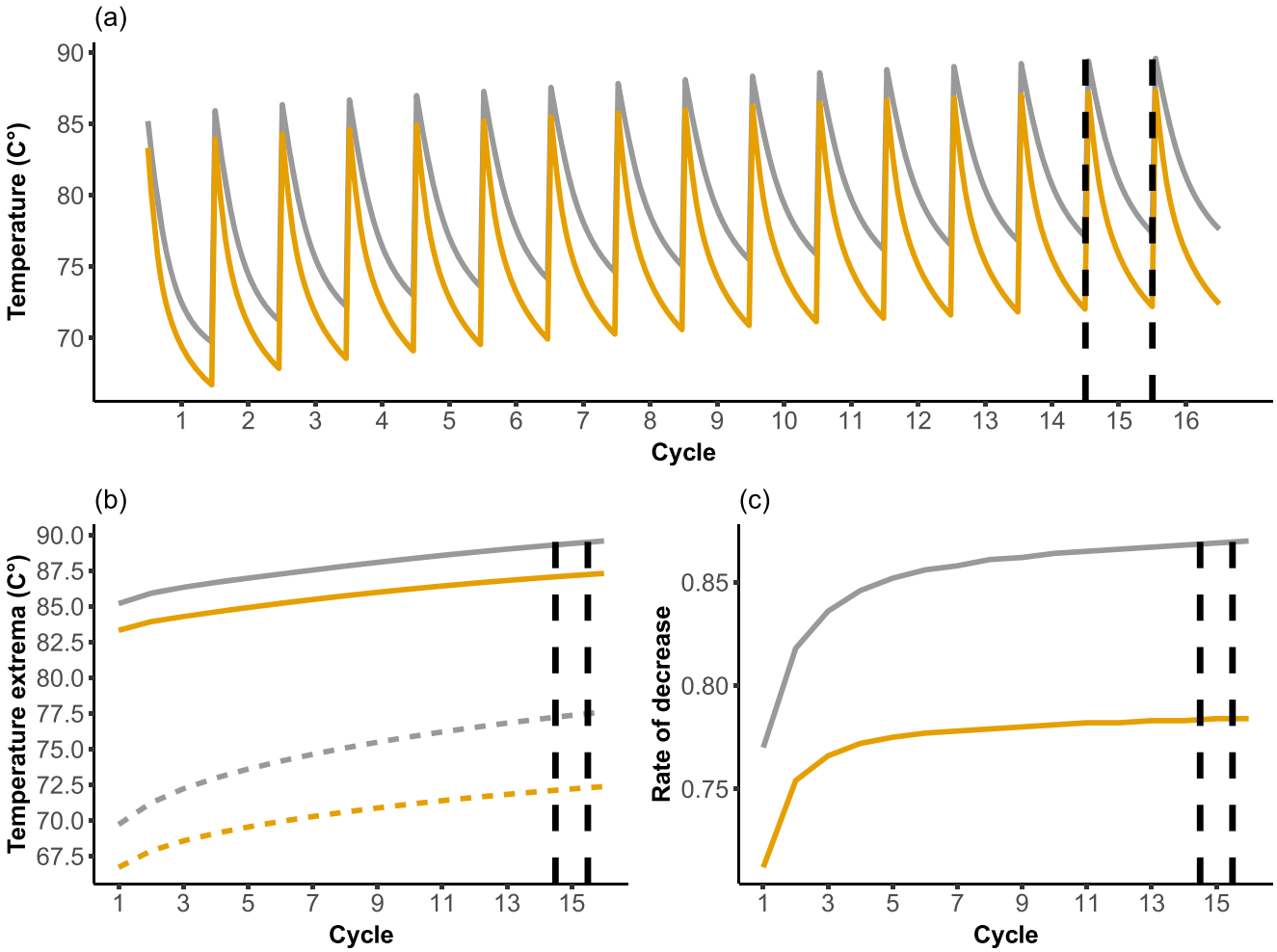

Fig. 1 **a** Examples of temperature evolution over 16 cycles during the unloading phase for two different settings of the ATES system described in Section 2. **b** Evolution of the temperature extrema at each cycle, minimum (dashed lines) and maximum (solid lines) values. **c** Evolution of the rate of decrease (decay parameter of an exponentially decreasing function) for each cycle. The vertical dashed lines indicate a given cycle (cycle N°15)

Platform,[1] which implements state-of-the-art numerical schemes to discretize multiphase Darcean flows on generic unstructured meshes. The reader is referred to [20] (and the references therein) for further theoretical and implementation details. All numerical simulations (described in the next section) are performed using a parallel computation scheme with 4 MPI (message passing interface) processes, and the total computation time for each case ranges from <1 hour to ≈14 hours depending on the setting of the input variables (described in Section 2.3).

2.3 Input variables

The settings of the ATES system are parameterized using a total of 14 different input variables (Table 1). The first category of input variables relates to the hydrogeological factors and corresponds to:

- the reservoir properties (intrinsic permeability, permeability anisotropy, porosity, initial temperature, and reservoir architecture);
- the caprock porosity;
- the hydrogeological setting (regional water flow direction and gradient related to the regional water flow).

Note that most of these uncertain parameters correspond to continuous scalar variables, but some uncertain parameters have been chosen to be mathematically represented with scenario-like variables, i.e., variables that can only take discrete values (named levels). These correspond to the reservoir architecture (three scenarios depicted in Fig. 4), regional water flow direction (two scenarios W–E or E–W), and reservoir

[1] https://charms.gitlabpages.inria.fr/ComPASS/ and http://www.anr-charms.org/page/compass-code

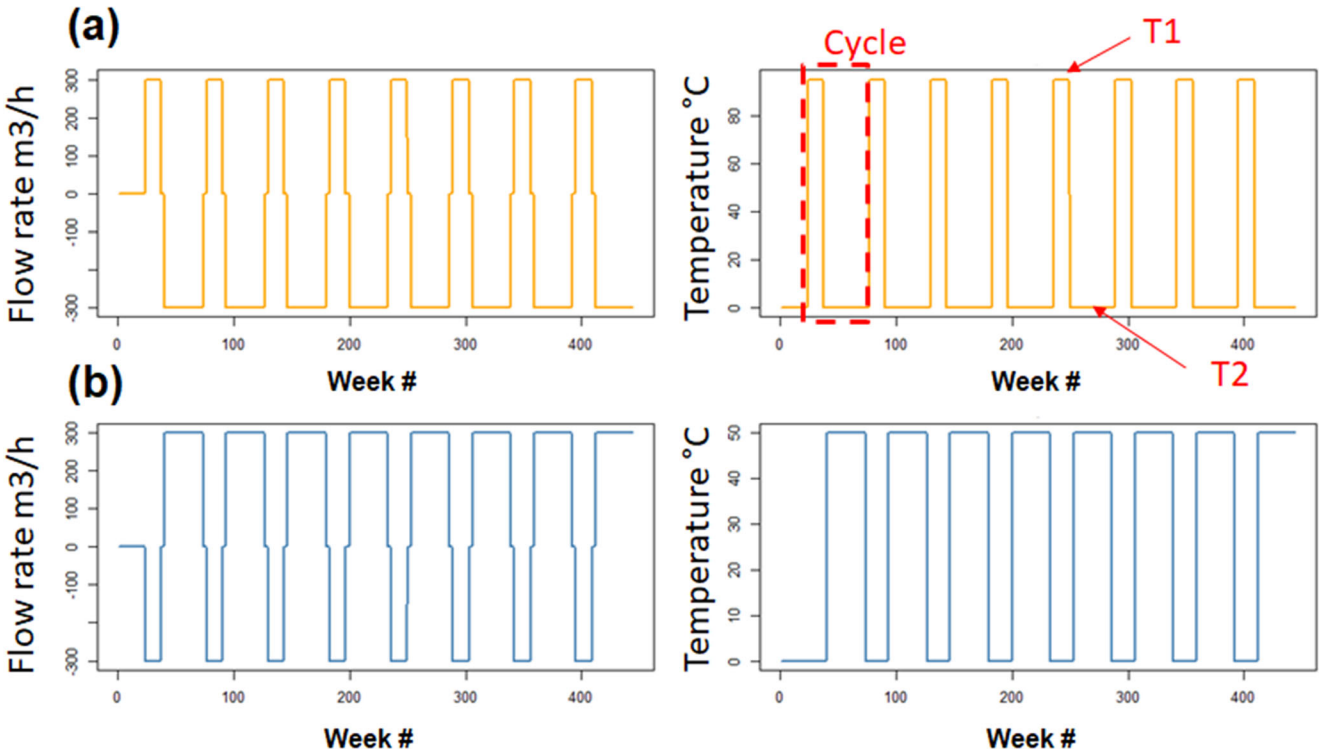

Fig. 2 Description of the loading/unloading cycle (flow rate and temperature) imposed at the hot (**a**) and cold wells (**b**). The horizontal axis provides the time expressed in number of weeks since the start of the production. A cycle is defined by two time durations, T_1 and T_2

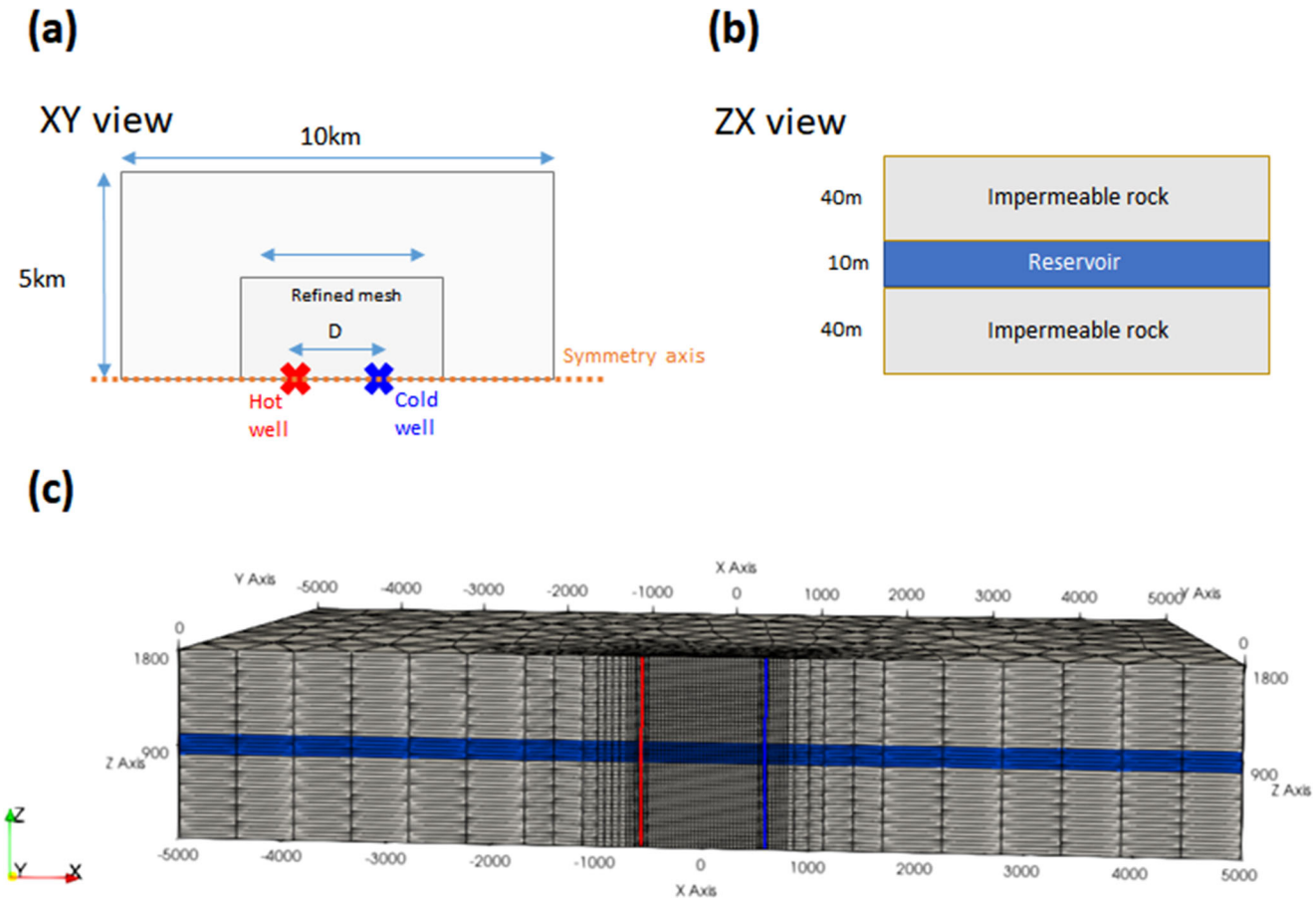

Fig. 3 **a** XY geometric view of the numerical model; **b** ZX view of the reservoir structure; and (**c**) prismatic mesh with a vertical exaggeration of 20

Table 1 Lower and upper bounds of each continuous scalar input variable. For scenario-like input variables, the number of scenarios is indicated

Input variable	Symbol	Lower bound	Upper bound	Unit
Hydrogeological setting				
Gradient	*Grad*	0.01	0.3	bar/km
Direction	*Dir*	2 scenarios		
Reservoir				
Intrinsic permeability*	K_{res}	11	12.3	–
Porosity	P_{res}	10	25	%
Initial temperature	T_{res}	45	85	°C
Anisotropy	*Ani*	3 scenarios		
Architecture	*Res*	3 scenarios (Fig. 4)		
Caprock				
Porosity	P_{cap}	0.1	10	%
Design parameters				
Interwell distance	*D*	800	1600	m
Time duration N°1 of the cycle	T_1	10	15	weeks
Time duration N°2 of the cycle	T_2	30	35	weeks
Maximum flow rate	*Q*	175	375	m^3/h
Temperature at the hot well (HW)	T_{HW}	75	115	°C
Temperature at the cold well (CW)	T_{CW}	30	65	°C

*With transformation -log10()

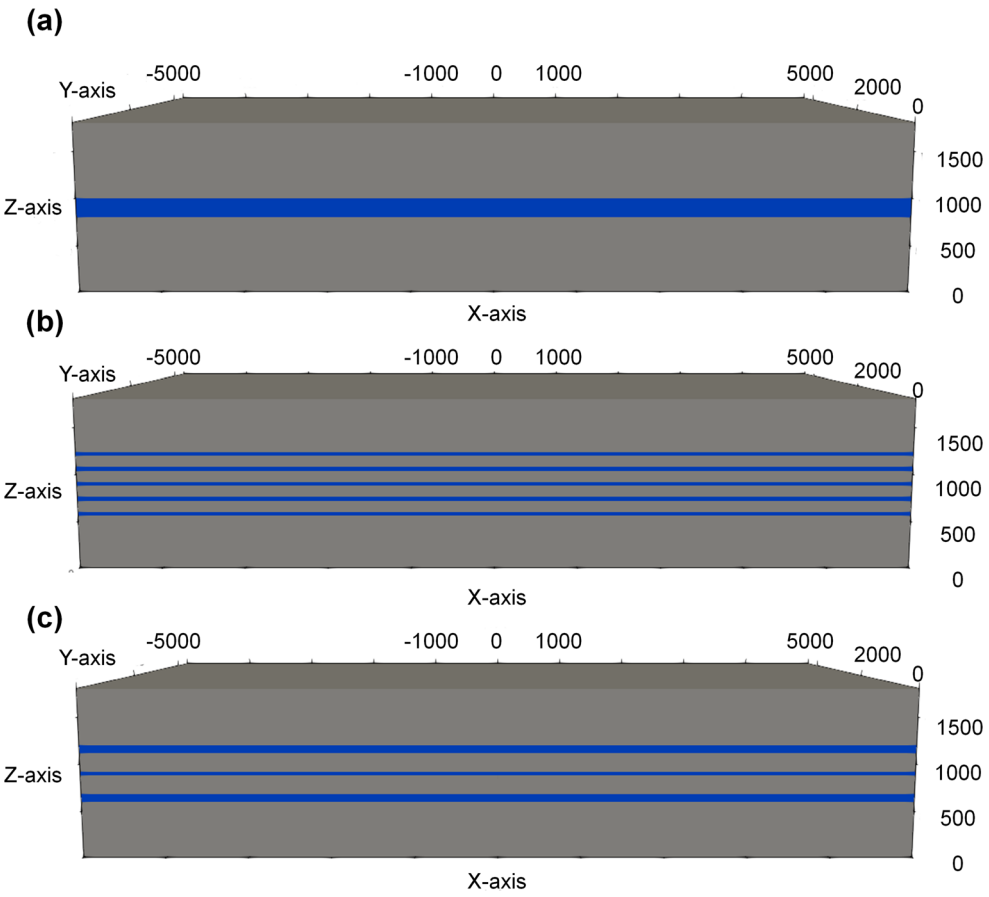

Fig. 4 Three scenarios of reservoir architecture

permeability anisotropy (three scenarios $K_{xx} = K_{yy} > K_{zz}$; $K_{xx} > K_{yy} > K_{zz}$; $K_{yy} > K_{xx} > K_{zz}$).

The second category of input variables corresponds to design parameters, namely, the interwell distance D and the parameters describing the unloading/loading cycles at the hot and cold wells (Table 1). A total of 16 cycles (each of them of approximately one year, depending on the design parameters, see Fig. 2) are simulated. Each cycle consists of (1) $T_1 = 10$ to 15 weeks of hot fluid storage at a constant flow rate Q ranging from 175 to 375 m^3/h and at a temperature T_{HW} ranging from 75 to 115 °C, while water is unloaded from the cold well; (2) $T_2 = 30$ to 35 weeks of hot fluid unloading at a constant flow rate Q; and (3) a 3-week delay time at the end of each period.

2.4 Computer experiments

To cover a broad range of different settings, we perform computer experiments by randomly and uniformly generating 300 different vectors of the 14 input variables using the sliced Latin hypercube design developed by Qian [21]. This method was selected because of its high efficiency in producing space-filling random designs while accounting for a mixture of continuous and scenario-like variables; see also, e.g., the work by Al-Mudhafar et al. [22], for a study of the efficiency of Latin hypercube designs in the context of metamodel-based optimization for oil reservoirs. Prefiltering is applied so that the reservoir initial temperature T_{res} ranges between the cold and hot well temperatures (T_{HW}, T_{CW}). To illustrate the distribution of the training points, the matrix of scatterplots for the group of temperatures (T_{res}, T_{HW}, T_{CW}) as well as the distribution of the scenario-like variables are provided in Fig. A1 and Fig. A2, respectively, in Supplementary Material A.

For each of these random ATES settings, the numerical model is run to compute the time evolution (over 16 cycles) of different variables (pressure, temperature, flow rate, energy amount, etc.). We focus the study on the temperature extracted at the hot and cold wells when they are alternatively producers. A preliminary processing procedure is applied to extract the minimum/maximum temperature T_{min}/T_{max} at each cycle as well as the temporal pattern depicted in Fig. 5. Details on this processing are given in Supplementary Material B.

Figure 5 shows that the "general" temporal pattern at each cycle (denoted c) is reproduced for both the loading (Fig. 5a) and unloading phases (Fig. 5b). This means that only the main characteristics of this pattern vary across different cycles. To model the intracyclic temporal evolution, at each cycle c, we consider an exponentially decreasing model of the form:

$$(T_{max}(c) - T_{min}(c))\exp(-rate(c).t) + T_{min}(c) \quad (1)$$

where $T_{max}(c)/T_{min}(c)$ are the temperature extrema for cycle c and the *rate* parameter measures the decay speed over time (denoted t) for the considered cycle c.

In the following, we aim to predict the intercyclic evolution of the characteristics in Eq. 1 (named variables of interest) given the configuration of the ATES system described by the 14 input variables (Table 1). We restrict the analysis to the unloading phase (Fig. 5b), which presents a more complex (nonlinear) pattern compared to the (quasi)linear intracyclic evolution of the temperature for the loading phase. Figure 6 provides the statistics of the intercyclic evolution for the three variables of interest considering the 16 cycles.

3 Methods

The multicyclic metamodelling approach relies on kriging metamodels (described in Section 3.1) trained within the multifidelity cokriging setting (Section 3.2). The proposed approach is validated by applying the procedure described in Section 3.3. To ease the training phase, a preliminary screening analysis is conducted to identify the variables that influence the most results (Section 3.4), i.e., the variables that are selected as inputs of the multicyclic metamodel.

3.1 Kriging metamodel

Let us consider the vector of d continuous input variables $\mathbf{x} = (x_1, \ldots, x_d)$. The variable of interest y is then computed using the numerical geothermal simulator, which takes $\mathbf{x}$ as input. We denote by $\boldsymbol{y}$ the vector of N known values of $y(\mathbf{x})$ that are computed on a design of experiments $\mathbf{D} = \{\mathbf{x}^{(1)}, \ldots, \mathbf{x}^{(N)}\}$. The set of pairs $\{\mathbf{x}^{(i)}, y^{(i)}\}_{i=1,\ldots,N}$ is the training dataset (comprised of N training samples) that is used to construct the metamodel.

In the context of kriging metamodelling (also named Gaussian process (GP) regression [23]), the function $y(\mathbf{x})$ is assumed to be a realization of a GP ($Y(\mathbf{x})$) with a constant mean μ and a covariance function $k(.,.)$, named the kernel, that can be written as $\forall \mathbf{x}, \mathbf{x}'$, $k(\mathbf{x}, \mathbf{x}') = \mathrm{cov}(Y(\mathbf{x}), Y(\mathbf{x}')) = \sigma^2 . R(\mathbf{x}, \mathbf{x}')$ with σ^2 being the variance parameter and $R(.,.)$ the correlation function related to k(.,.)

The prediction at a new observation $\mathbf{x}^*$ is given by the kriging mean $\widehat{Y}(\mathbf{x}^*)$ using the equations of ordinary kriging (see, e.g., [24]) as follows:

$$\widehat{Y}(\mathbf{x}^*) = \mu + \boldsymbol{r}'(\mathbf{x}^*).\mathbf{R}^{-1}.(\boldsymbol{y} - \mu.\mathbf{I}_N) \quad (2)$$

where $\mathbf{R}$ is the correlation matrix whose element is $R[i,j] = R(\mathbf{x}^{(i)}, \mathbf{x}^{(j)})$; $\boldsymbol{r}$ is the vector composed of the correlation between $\mathbf{x}^*$ and the points of $\mathbf{D}$ written as $\sigma^2 . R(\mathbf{x}^*, \mathbf{D})$, and $\mathbf{I}_N$ is the vector of ones of length N.

For the continuous input variables, the covariance function k_{cont} is assumed to follow the Matérn 5/2 model [23]. As described in Section 2.2, some input variables are scenario-like variables, i.e., categorical variables. There are different ways to

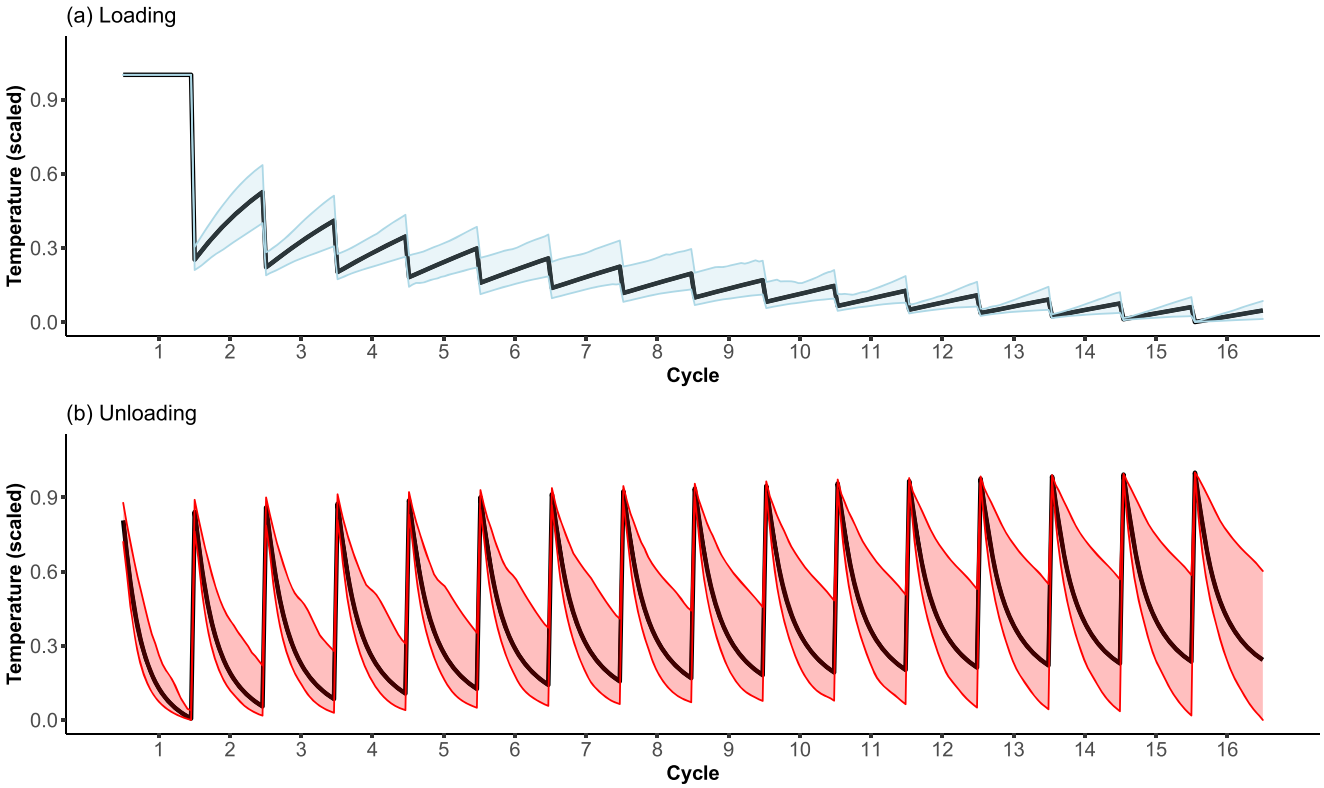

Fig. 5 Evolution of the temperature (scaled between 0 and 1) at the producer (black line) during the loading (**a**) and unloading (**b**) phases. The solid black lines correspond to the mean calculated considering 300 random configurations of the ATES setting. The limits of the coloured envelope are defined by computing the 5th and 95th percentiles

model them in the kriging model [25]. The analysis of the categorical variables in our case reveals that they are implicitly associated with a notion of order (they are ordinal variables). To account for them, we follow the approach described by Roustant et al. [25]: Section 2.2.1, and assume that the corresponding kernel model k_{cat} is described as follows:

$$k_{\text{cat}}\left(u, u^{'}\right) = \widetilde{k}_{\text{cont}}\left(F(u), F\left(\mathrm{u}^{'}\right)\right) \tag{3}$$

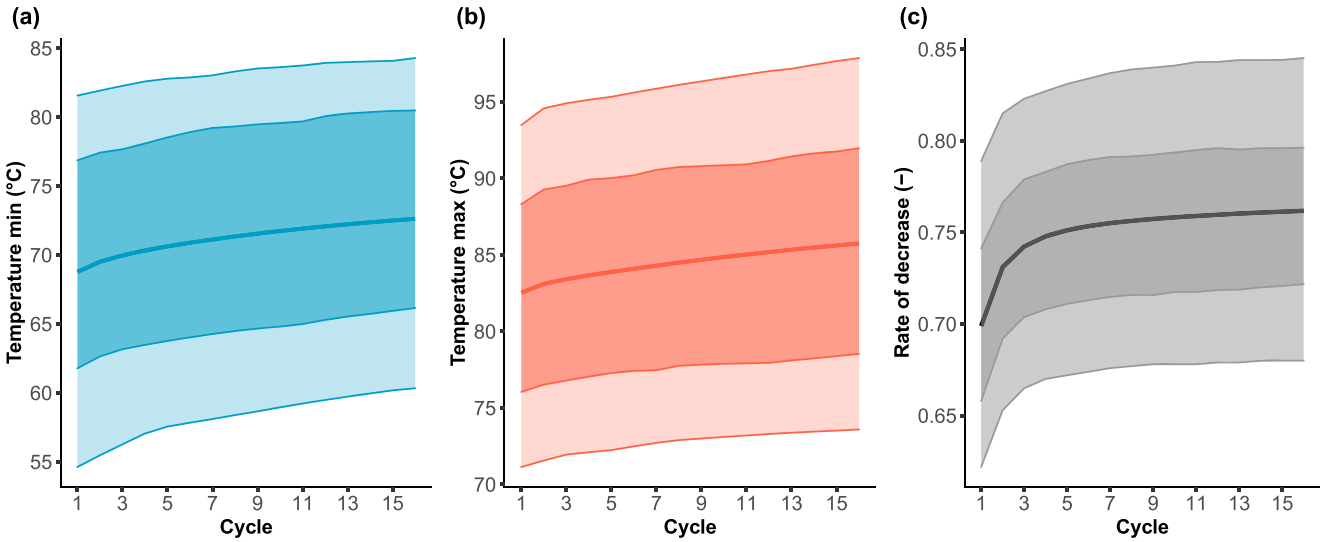

Fig. 6 Intercyclic evolution of the main characteristics of the intracyclic dynamics (described by Eq. 1): minimum (**a**) and maximum (**b**) temperature and rate of decrease (**c**). The solid line corresponds to the median value computed for 300 random settings of the ATES system. The light and dark coloured envelopes are defined by computing the percentiles at 25–75% and 5–95%, respectively

where u is the categorical variable that can take up a finite number of levels L, $\widetilde{k}_{\text{cont}}$ is a one-dimensional continuous kernel (here assumed to be the Matérn 5/2 model), and $F(.)$ is a one-dimensional nondecreasing function (also called warping) $F: \{1, \ldots, L\} \rightarrow \mathbb{R}$. In our case, the warping function is chosen to be modelled by the identity function $F(u) = u$. The covariance function $k(\mathbf{x}, \mathbf{x}')$ for both types of input variables (continuous and categorical) is formed via the tensor product of k_{cont} with k_{cat}.

3.2 Multifidelity cokriging

When running a numerical model, different levels of complexity (for instance, for different spatial resolutions of the mesh) are generally possible. This results in a hierarchy of model levels; the model with the finest mesh is the model with the highest fidelity. Multifidelity metamodelling (e.g., [15, 26]) was introduced to take advantage of these different levels to predict the output at the finest level (high fidelity) using the results provided by the model of a coarser level (low fidelity); the advantage is that the latter model has a lower computation time cost. The analysis of the variables of interest at each cycle (Figs. 5 and 6) indicates a strong hierarchy between the cycles, which suggests making an analogy between the aforementioned levels of complexity and the cycle index, i.e., the higher the cycle index is, the finer the level, as described in the introduction. This means that the results computed for cycle N°1 only (the simulation is said to be partial) are seen as approximate versions of the results computed over all cycles, from cycle N°1 to N°16 in our case (the simulation is said to be complete). The expected result is a reduction in the computational burden because gathering information at a low cycle index (e.g., N°1 or N°2) is less expensive because the model run is stopped earlier. In the following, we present how to set up the multicyclic metamodel within the multifidelity framework.

For simplicity, we restrict the presentation to two levels of fidelity, i.e., cycles N°1 and N°2. We denote by $y_1(\mathbf{x})$ and $y_2(\mathbf{x})$ the output at each respective cycle considering the vector of inputs $\mathbf{x}$, and we denote by $\boldsymbol{y}_1$ the vector of n_1 known values of $y_1(\mathbf{x})$ on a design of experiments $\mathbf{D}_1$ and by $\boldsymbol{y}_2$ the vector of n_2 known values of $y_2(\mathbf{x})$ on $\mathbf{D}_2$ by considering nested designs $\mathbf{D}_2 \subset \mathbf{D}_1$ and $n_2 < n_1$. For our case, this means that we use more results computed at cycle N°1 than results computed from cycle N°1 to N°2; the computation time cost of the former set of simulations is lower because the simulated time is shorter.

We assume that $y_1(\mathbf{x})$ and $y_2(\mathbf{x})$ are both realizations of GP $Y_1(.)$ and $Y_2(.)$. To link them, we follow the approach of [15] by assuming that $\text{cov}(Y_2(\mathbf{x}),\ Y_1(\mathbf{x}')|\ Y_1(\mathbf{x})) = 0, \forall\, \mathbf{x} \neq \mathbf{x}'$, which leads to the following model:

$$Y_2(\mathbf{x}) = \rho Y_1(\mathbf{x}) + Y_d(\mathbf{x}) \tag{4}$$

where ρ is a scale factor and $Y_d(.)$ is a Gaussian process independent of $Y_1(.)$ that represents the difference between $\rho Y_1(.)$ and $Y_2(.)$. The kriging predictors $\widehat{Y}_1$ and $\widehat{Y}_d$ are then constructed by using Eq. 2 with means μ_1 and μ_d and correlation functions $R_1(.,.)$ and $R_d(.,.)$ with variance parameters σ_1^2 and σ_d^2. The kriging predictor $\widehat{Y}_2$ is then defined as follows:

$$\widehat{Y}_2(\mathbf{x}^*) = \mu_2 + \boldsymbol{r}_2'(\mathbf{x}^*).\mathbf{R}_2^{-1}.(\boldsymbol{y} - \mu_2.\mathbf{I}_{n_1+n_2}) \tag{5}$$

where

$\boldsymbol{y} = (\boldsymbol{y_1}, \boldsymbol{y_2})$,

$\mu_2 = \rho.\ \mu_1 + \mu_d$, $\boldsymbol{r}_2 = \begin{pmatrix} \rho.\sigma_1^2.R_1(\mathbf{x}^*, \mathbf{D}_1) \\ \rho^2.\sigma_1^2.R_1(\mathbf{x}^*, \mathbf{D}_2) + \sigma_d^2.R_d(\mathbf{x}^*, \mathbf{D}_2) \end{pmatrix}$,

and the covariance matrix is $\boldsymbol{R}_2 = \begin{pmatrix} \sigma_1^2.R_1(\mathbf{D}_1, \mathbf{D}_1) & \rho.\sigma_1^2.R_1(\mathbf{D}_1, \mathbf{D}_2) \\ \rho.\sigma_1^2.R_1(\mathbf{D}_1, \mathbf{D}_2) & \rho^2\sigma_1^2.R_1(\mathbf{D}_2.\mathbf{D}_2) + \sigma_d^2.R_d(\mathbf{D}_2, \mathbf{D}_2) \end{pmatrix}$.

These equations can be generalized to handle multiple levels of complexity, i.e. multiple cycles in our case, by following [26]. In practice, the training of the multicyclic metamodel implies estimating the different parameters (means μ_1, μ_d and variance parameters σ_1^2, σ_d^2, and length scales of the different correlation functions R_1 and R_d). This is performed by relying on maximum likelihood estimation with the optimized algorithm proposed by Le Gratiet [26].

3.3 Validation procedure

The objective is to assess to what extent the multicyclic metamodel is capable of predicting the variables of interest with high accuracy by considering "yet-unseen" new configurations of the input variables (Table 1). To quantify this predictive performance, we adopt a random subsampling approach (see, e.g., [27]). Given a set of N_{tot} pairs $\left\{\mathbf{x}^{(i)}, \mathrm{y}^{(i)}\right\}_{i=1\ldots N_{\text{tot}}}$, we randomly select N pairs for the training and n_{test} pairs (among those that are not used for the training) for evaluating some performance indicators. This procedure is repeated multiple times. In our case, $n_{\text{test}} = 100$ (i.e., 33% of the total number of available results), and N is varied from 100 to 200 (i.e., 33 to 66% of the total number of available results).

In this study, at each cycle c, we compute three different performance indicators, i.e., the root mean square error *RMSE*, the mean absolute error *MAE* (expressed in units of the variable of interest), and the mean absolute percentage error *MAPE*, which is a normalized criterion expressed in %:

$$RMSE(c) = \sqrt{\frac{1}{n_{\text{test}}} \sum_{i=1}^{n_{\text{test}}} \left(\widehat{Y}_c\left(\mathbf{x}^{(i)}\right) - Y_c\left(\mathbf{x}^{(i)}\right)\right) 2} \tag{6a}$$

$$MAE(c) = \frac{1}{n_{\text{test}}} \sum_{i=1}^{n_{\text{test}}} \left|\widehat{Y}_2\left(\mathbf{x}^{(i)}\right) - Y\left(\mathbf{x}^{(i)}\right)\right| \tag{6b}$$

$$MAPE(c) = \frac{100}{n_{test}} \sum_{i=1}^{n_{test}} \left| \frac{\widehat{Y}_c\left(\mathbf{x}^{(i)}\right) - Y_c\left(\mathbf{x}^{(i)}\right)}{Y_c\left(\mathbf{x}^{(i)}\right)} \right| \quad (6c)$$

where $\mathbf{x}^{(i)}$ is the i^{th} vector of input variables in the test set, Y_c is the true value of the variable of interest at cycle c (i.e., the result of the numerical model at cycle c), and $\widehat{Y}_c$ is the prediction provided by the multicyclic metamodel at cycle c.

In addition to the multicyclic metamodelling approach, we consider an alternative approach named "independent", which consists of predicting Y_c using a kriging metamodel trained by considering only the results at cycle c. This means that the prediction for Y_1, Y_2, …, Y_c are provided by c kriging metamodels trained independently at each cycle using Eq. 2.

3.4 HSIC-based screening analysis

The number of input variables is 14 (see Section 2), which is sufficiently large to pose some difficulties in the training of the kriging metamodel, especially regarding the moderate size of the training dataset, where $N_{tot} = 300$ (see, e.g., [28]). A preliminary screening analysis is then conducted, which consists of filtering out the variables of negligible influence [29]. We adopt here the procedure developed by Da Veiga [30] based on the HSIC (Hilbert–Schmidt independence criterion) measure, which can achieve the screening task with a limited number of numerical simulations (a few hundred). In the following, we recall the main aspects of the procedure; the reader can refer to [30] for full details.

Let us associate X_i with a universal reproducing kernel Hilbert–Schmidt (RKHS) space defined by the characteristic kernel function $k_i(.,.)$. The same transformation is associated with Y by considering a RKHS space with kernel $k(.,.)$. We define the HSIC measure as follows:

$$\begin{aligned} HSIC(X_i, Y) = \mathrm{E}\left(k_i\left(X_i, X_i'\right)k\left(Y, Y'\right)\right) + \mathrm{E}\left(k_i\left(X_i, X_i'\right)\right)\mathrm{E}\left(k\left(Y, Y'\right)\right) \\ -2\mathrm{E}\left(\mathrm{E}\left(k_i\left(X_i, X_i'\right)|X_i\right)\mathrm{E}\left(k_i\left(Y, Y'\right)|Y\right)\right) \end{aligned} \quad (7)$$

where $\left(X_i', Y'\right)$ is an independent and identically distributed copy of (X_i, Y), and E(.) is the expectation operator. As a characteristic kernel for continuous variables, the Gaussian kernel is used and is defined as $\exp(-\lambda\|\mathbf{x} - \mathbf{x}'\|2)$, with λ being the bandwidth parameter chosen as the inverse of the empirical variance of the considered variable. For categorical variables, the identity function is used as a characteristic kernel.

From a sensitivity analysis perspective, using the HSIC measure is useful to identify noninfluential factors because its nullity indicates that X_i does not influence Y. To identify the significant X_i, the null hypothesis $'H_0 : HSIC(X_i, Y) = 0'$ (against the hypothesis $'H_1 : HSIC(X_i, Y) > 0'$) is tested, and the corresponding p value is evaluated (see [31] and references therein). When the p value is below a significance threshold (chosen here as 5%), it indicates that the null hypothesis should be rejected, i.e., this means that if the p value <5%, the considered input X_i has a significant influence on Y. To compute the p value, we rely on the sequential bootstrap-based algorithm developed by El Amri and Marrel [31].

In our case, an additional difficulty is related to the dependence between some input variables, i.e., the hot well and cold well temperatures (T_{HW}, T_{CW}) and the reservoir initial temperature T_{res} cannot be sampled independently because they should fulfil some constraints of inequality (see Section 2.2). In this situation, the corresponding HSIC measures cannot be interpreted as the individual effect of each of these variables. To account for this dependence, we replace these dependent variables with a unique categorical variable. The latter variable can be defined by grouping the three dependent variables by means of clustering algorithms such as k-means [32]. The levels of the categorical variable then correspond to the cluster indices assigned to each triplet of dependent variables. For instance, performing k-means clustering with k = 20 allows the triplets (T_{HW}, T_{CW}, T_{res}) to be grouped among 20 different clusters. The HSIC measure is computed for the newly defined categorical variable using a categorical kernel and reflects the influence of the whole group of dependent variables on the uncertainty of the considered variable of interest.

4 Results

In this section, we first apply the HSIC-based screening analysis (Section 3.4) to identify the input variables of high importance among the 14 variables. Then, in Section 4.2, we train the multicyclic metamodel by considering the input variables identified as important for the three variables of interest (T_{min}, T_{max} and the rate of decrease) and analyse its predictive performance. To extensively explore the predictive performance, the training procedure is conducted 100 times in Section 4.3, and the results are compared with those of the independent metamodelling approach. Finally, we test the robustness of the results to the type of training procedure (Section 4.4).

4.1 Screening analysis

We first apply HSIC-based screening analysis to identify the input variables that are the most important with respect to the three variables of interest, T_{min}, T_{max}, and the rate of decrease over 16 cycles. To do so, we estimate the p values of the statistical test based on the HSIC measures at each cycle (evaluated using the 300 learning samples). As detailed in Section 3.4, we account for the dependence between the three temperatures, T_{HW}, T_{CW} and T_{res}, by replacing this group of variables with a categorical variable named '*TempG*', whose value (level) corresponds to the cluster index provided by a k-means algorithm when k = 20. Preliminary tests have shown that varying k from 15 to 25 does not impact the results of the sensitivity analysis.

Figure 7 provides the boxplots of the p values (across the 16 cycles). This allows us to identify the input variables that are noninfluential with respect to each considered variable of interest, i.e., the ones whose p value is above 5%. This shows that the identified noninfluential variables do not change from one cycle to another. This also shows that the number of variables that should be used as inputs of the metamodel can be reduced, i.e., from 14 to 3 or 4 depending on the variable of interest. The temperature extrema (T_{min} and T_{max}) are mainly influenced by the temperature loading characteristics of the ATES system (*TempG*), which was expected given the loading/unloading of fluids for this ATES system (Section 2.1). This is consistent with the results of our preliminary statistical analysis (Supplementary Material A: Figs. A6 and A7), which indicates that the temperature values vary over several orders of magnitude across the different *TempG* levels. In addition, the regional water flow direction *Dir* impacts the T_{max} value, which appears to be physically coherent because the maximum temperature is related to the fluid propagation from one well to another, which is more or less eased whether the regional flow direction aligns with that of the production fluid flow. Considering the rate of decrease, Fig. 7c shows that it is mainly influenced by the interwell distance D, the reservoir architecture *Res* and both time durations T_1 and T_2 (Fig. 7c). This is consistent with the results of our exploratory statistical analysis (Supplementary Material A: Fig. A5), which provides indications of a linear relation of the rate with D, T_1 and T_2. In addition, this preliminary analysis (Supplementary Material A: Fig. A8) also shows that rate values can vary significantly across the different reservoir architecture scenarios. The identification of these important variables appears to be in agreement with the physical processes acting during the ATES lifecycle because all these parameters are related to the time evolution of the fluid flow during operations. Depending on the reservoir architecture (Fig. 4) and the interwell distance (Fig. 3), the fluid will circulate more or less "easily" between the wells, which is modulated by the time characteristics of the stair-like loading cycles (Fig. 2). In the following, we select only these influential variables as inputs of the metamodels. The implication of this assumption is further discussed in Section 4.2.

4.2 Application of the multicyclic metamodel

We train the multicyclic metamodel for the three variables of interest by randomly selecting $N = 200$ training samples among the $N_{tot} = 300$ available and by retaining the remaining samples ($n_{test} = 100$ samples) to test the performance (Section 3.3). Figure 8 provides the evolution of the performance indicators over the 16 cycles as well as the comparison between the true and the predicted values for the last cycle (cycle N°16). This shows that the prediction error *RMSE*/*MAE* increases as the cycle index increases and reaches small-to-moderate values of a few °C for T_{min}/T_{max} and ~0.03 for the rate, which can be considered satisfactory given the order of magnitude of the considered variables (Fig. 6). In addition, the normalized errors (measured by *MAPE*) reach values up to 1.6, 3.5% and < 4% for T_{min}, T_{max}, and the rate, respectively, which confirms the satisfactory performance. We note, however, that the prediction of the rate is more difficult, as shown by the higher spread of the scatterplot in Fig. 8c as well as the larger bias in the prediction.

The added value of performing a preliminary phase of screening (Section 4.1) is shown in Fig. 9, which corresponds to the same analysis as Fig. 8 but with metamodels trained by considering all 14 input variables. A clear decrease in the predictive performance is outlined for T_{min} and T_{max} (Fig. 9a, b). The increase in the prediction error for the rate's prediction is noted, but it remains smaller than for T_{min} and T_{max}. Figure 9 (c, left) shows increases of 1–1.5 units for *RMSE* and *MAE* and by 0.5–1.0% for *MAPE*. We also note a higher bias for the rate (Fig. 9c, right).

The multicyclic metamodel also informs on the intercyclic dependence of the considered variable of interest. Figure 10 provides the evolution of the scale factor that models this dependence (term denoted ρ in Eq. 4). This shows that the intercyclic dependence decreases over the cycles (as suggested by the analysis of the time series in Figs. 5 and 6) by following a nonlinear trend for T_{min} and a quasilinear trend for T_{max}. Figure 10 also shows that the intercyclic dependence has the largest variation at the beginning and reaches a quasiconstant plateau from cycle N°4/N°5. This suggests that if the first part (before cycle N°4/5) is accurately modelled, the prediction for the second part should be improved (due to small changes).

4.3 Performance analysis

Here, we extensively explore the performance of the multicyclic metamodelling approach by testing the influence of several aspects: (1) the number of training samples ($N = 100$, 150 or 200); (2) the comparison to the independent metamodelling approach; and (3) the influence of different training options:

- Option 1 'C100%'. The same number N of training samples is used regardless of the cycle index;
- Option 2 'Cp%'. The designs of the experiments are nested (as described in Section 3.2) with a decreasing

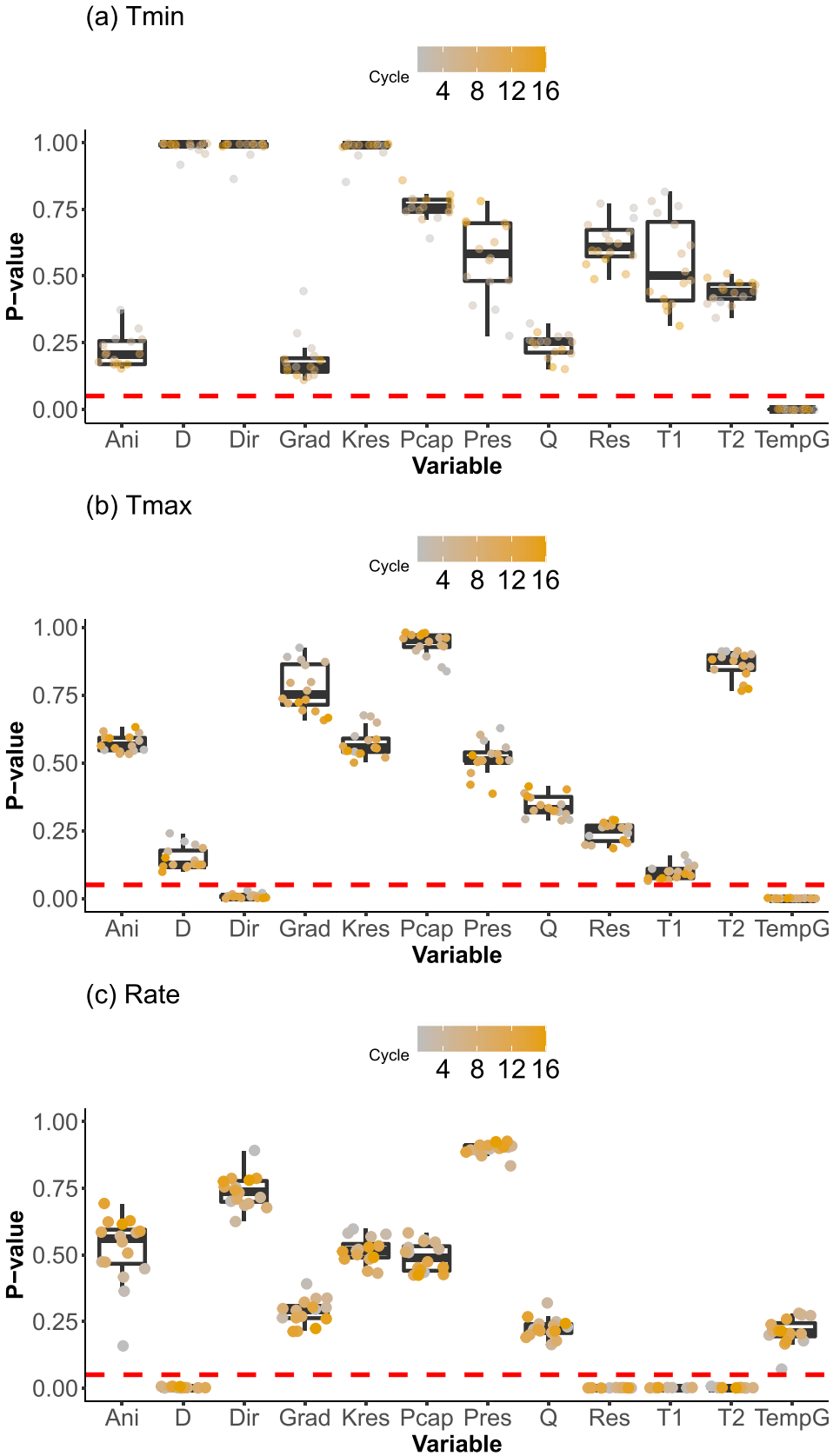

Fig. 7 Boxplot of the p values of the HSIC-based test of independence calculated for the minimum temperature (**a**), the maximum temperature (**b**) and the rate of temperature decrease (**c**) considering all 16 cycles. The horizontal dashed red line indicates the significance level at 5%. The colour indicates the cycle index (from 1 to 16)

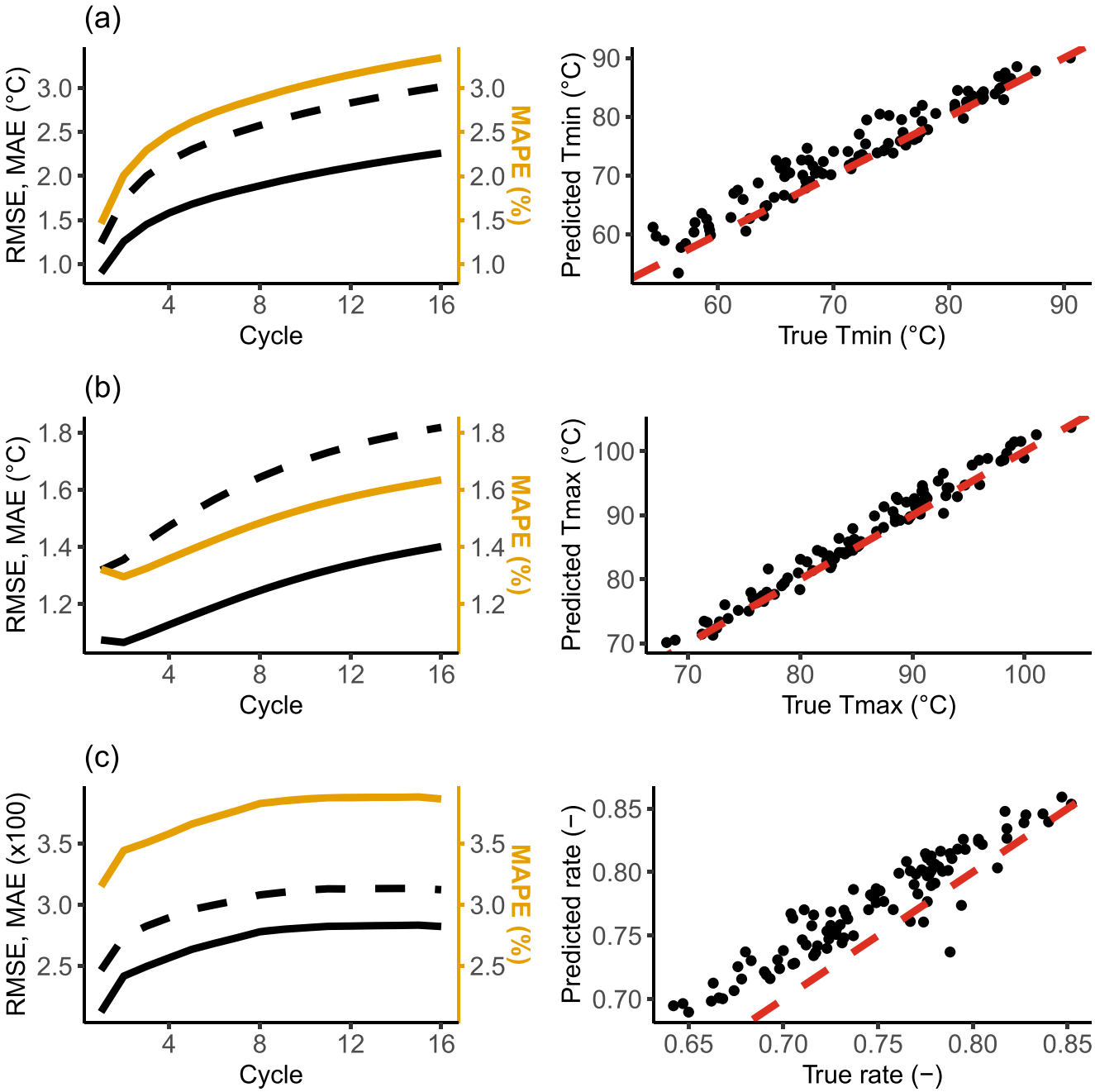

Fig. 8 (Left). Evolution of the performance indicators *RMSE* (sold black line), *MAE* (black dashed line) and *MAPE* (solid orange line) for T_{min} (**a**), T_{max} (**b**) and the rate of decrease (**c**); (Right). Comparison between the true and predicted values for the three variables of interest at the last cycle

number of training samples when the cycle index increases from 2 to 16. In other words, all N training samples are used for cycle N°1 (corresponding to the lower fidelity level), and a lower number of samples are used for cycles N°2 to N°16 by considering a proportion p (among 75%, 50%, and 25%) of N. Note that this option is less expensive to evaluate than option 1 because a smaller number of numerical simulations are run up to cycle N°16. Supplementary Material C: Fig. C1(a) provides an illustration of this training option.

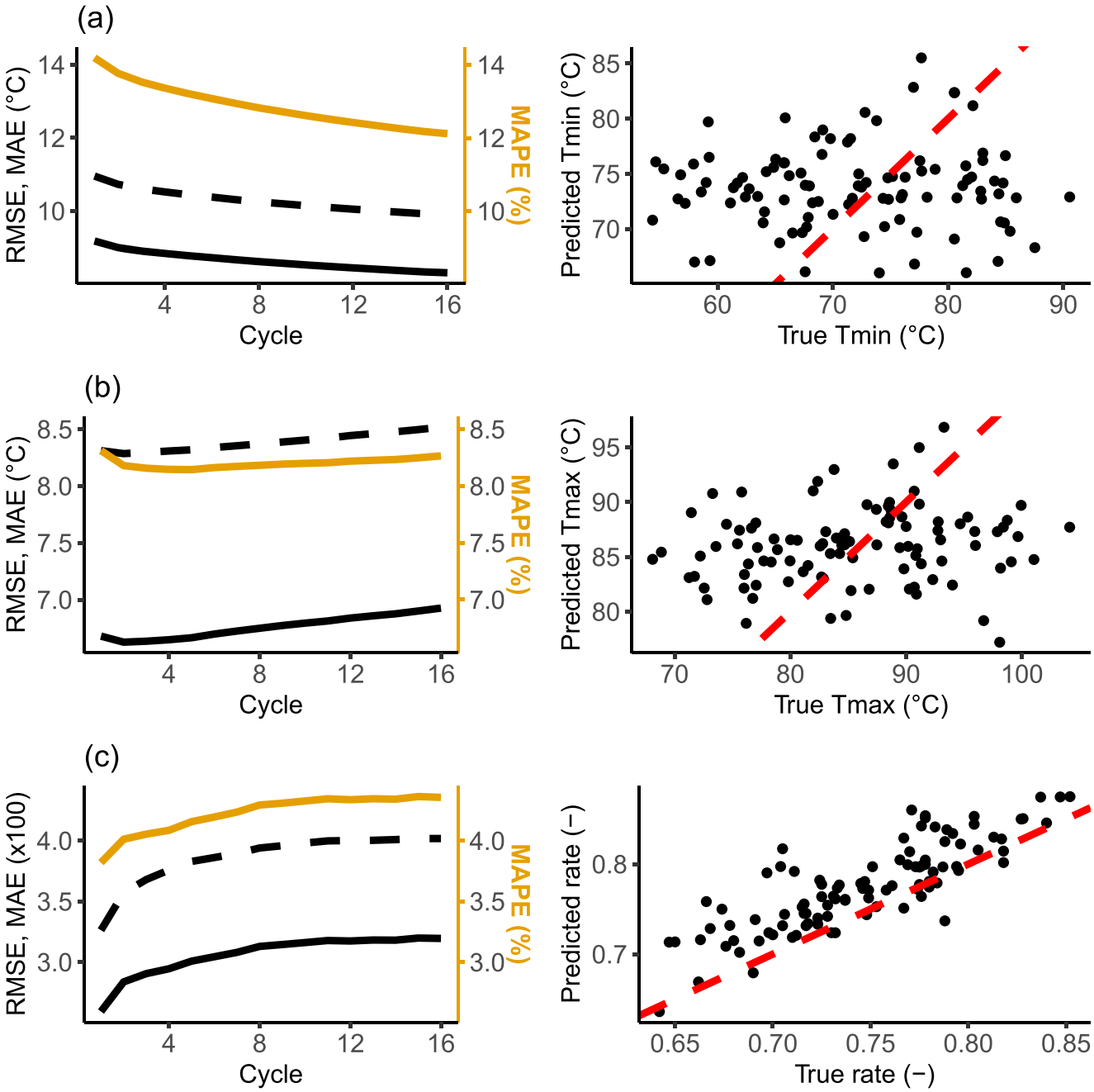

Fig. 9 (Left). Evolution of the performance indicators *RMSE* (solid black line), *MAE* (black dashed line) and *MAPE* (solid orange line) for T_{min} (**a**), T_{max} (**b**) and the rate of decrease (**c**) without the preliminary phase of screening; (Right). Comparison between the true and predicted values for the three variables of interest at the last cycle

The training/validation procedure is conducted multiple times by performing 100 replicates of the random subsampling procedure described in Section 3.3 using $n_{test} = 100$ test samples.

Figure 11 provides the *MAPE* results for the prediction of the rate. Without decreasing the proportion p of training samples, the performances of both metamodelling approaches, multicyclic and independent approaches, are equivalent (Fig. 11a–g). Decreasing p clearly impacts the performance of the independent metamodels, while the multicyclic approach remains almost insensitive to the decrease. At the

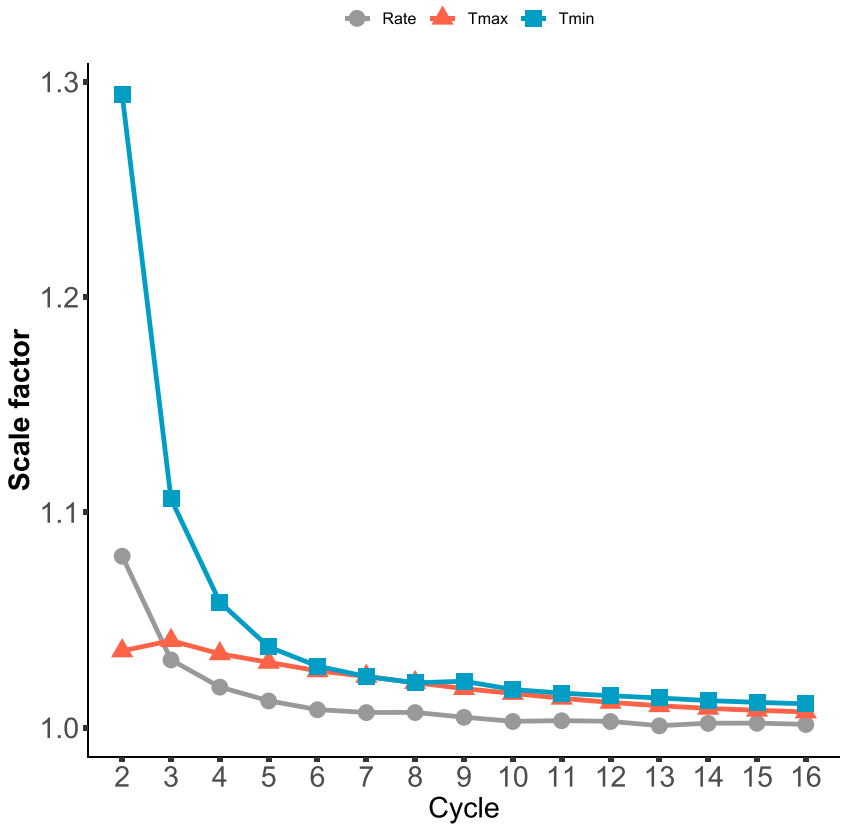

Fig. 10 Evolution of the scale factor over the cycles for the three variables of interest, rate (grey dot), T_{max} (red triangle), and T_{min} (blue square)

last cycle (N°16), the *MAPE* value is decreased by a factor of ~33% when using the multicyclic metamodelling approach. This decrease is also verified for the other performance indicators (Supplementary Material D). This higher performance is also shown for T_{min} (Fig. 12) and T_{max} (Fig. 13) with a decrease in the *MAPE* value by up to 20 and ~50%, respectively. This result clearly outlines the added value of incorporating the information of the intercyclic dependence in the metamodel construction: a higher level of predictive quality can be achieved with a smaller number (down to 25) of high-fidelity numerical simulations, i.e., a smaller number of numerical simulations that are run for all cycles up to cycle N°16. This means that the multicyclic metamodel here requires less computational effort.

4.4 Robustness to type of training procedure

In Section 4.3, we adopt a training procedure based on the 'Cp%' option. We further investigate whether alternative training procedures would impact the performance results. To do so, the predictive performance of the multicyclic metamodel is evaluated by considering an alternative option 'Lp%' in addition to option 1 'C100%' and option 2 'Cp%'. For this training option, the number of training samples at each cycle is defined by progressively and linearly decreasing their proportion from 100% to a given value p (among 75%, 50%, and 25%) as the cycle index increases from N°2 to N°16. This means that the computation time cost of this option is larger than "Cp%" but remains lower than that of option 1. Supplementary Material C: Fig. C1(b) provides an illustration of this training option. Similar to Section 4.3, the different options are investigated by performing the random subsampling procedure multiple times (here 100) with $N = 200$ and $n_{test} = 100$.

Considering the prediction of the rate, Fig. 14 provides the mean value of the *MAPE* performance indicator over the 16 cycles together with the 10th and 90th percentiles for the three different training options. This shows that the performance of the multicyclic metamodel is relatively insensitive to the type of training procedure (compare the top and bottom panels in Fig. 15) regardless of the proportion of samples p that is retained over the cycles. The same conclusion is reached for the two other variables T_{min}/T_{max} (see Supplementary Material E). Figure 14 suggests that (1) very few "complete" model runs (here down to 50) are necessary to reach a high predictive performance (which is also shown in Section 4.3); (2) this overall predictive

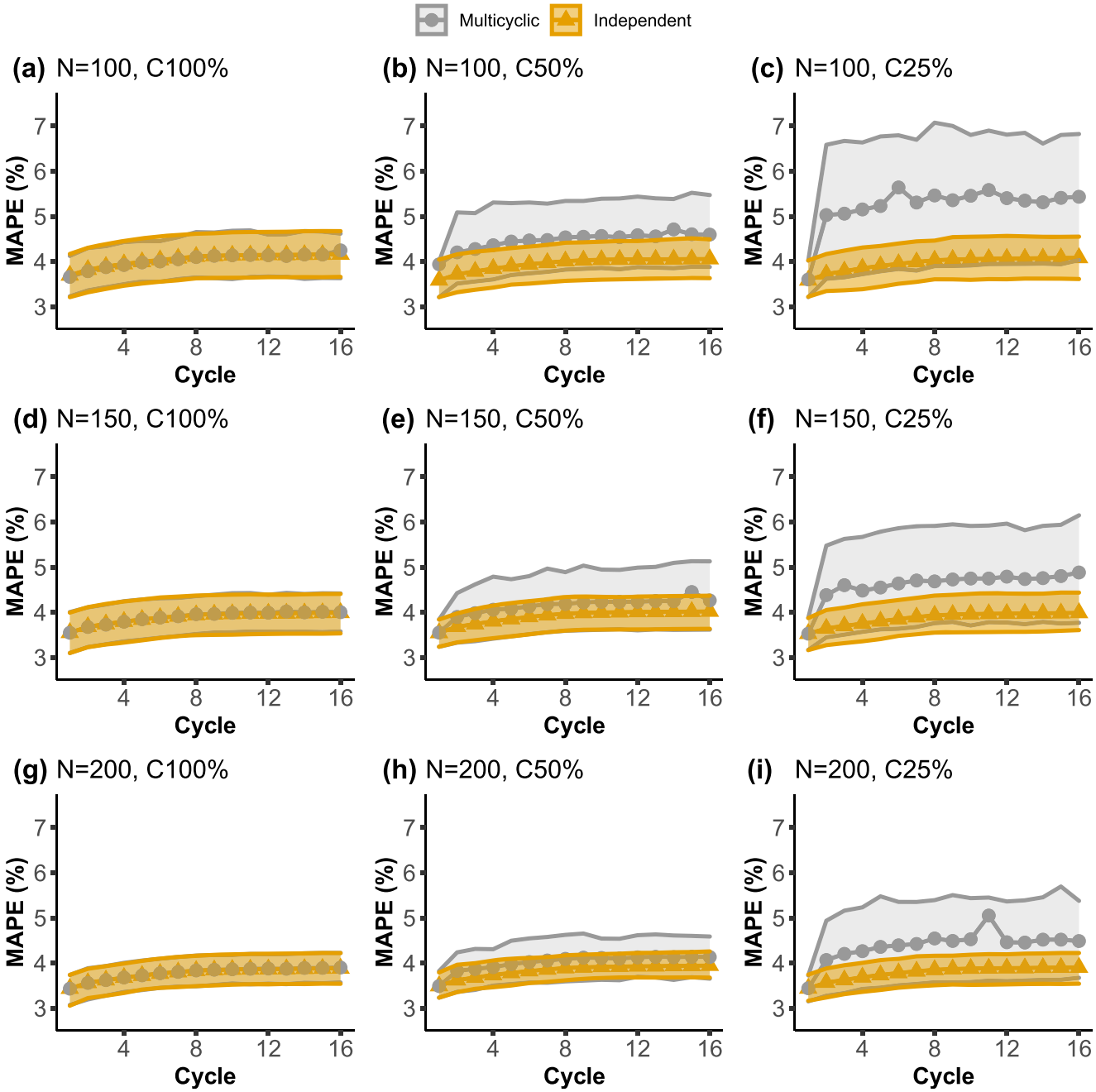

Fig. 11 Evolution of the *MAPE* performance indicator for rate prediction using the independent (grey dot) and multicyclic (orange triangle) metamodelling approaches. Different numbers *N* of training samples are tested as well as different training options (named "Cp%") with *p* indicating the proportion of samples kept among *N* for the training over the cycles N°2 to N°16. The solid line is the mean value calculated from the results of a random subsampling approach conducted 100 times. The envelope limits are computed using the 10th and 90th percentiles

performance mainly depends on the number of training samples for cycle N°1 (training option "Cp%"), and (3) the number of training samples for cycles N°2 to N°16 has little impact on the performance.

We further examine the influence of the training option on the intercyclic dependence (modelled by the scale factor in Eq. 4). Figure 15 shows that the spread in the results (measured by the envelope width) is relatively small for option "Lp%" and

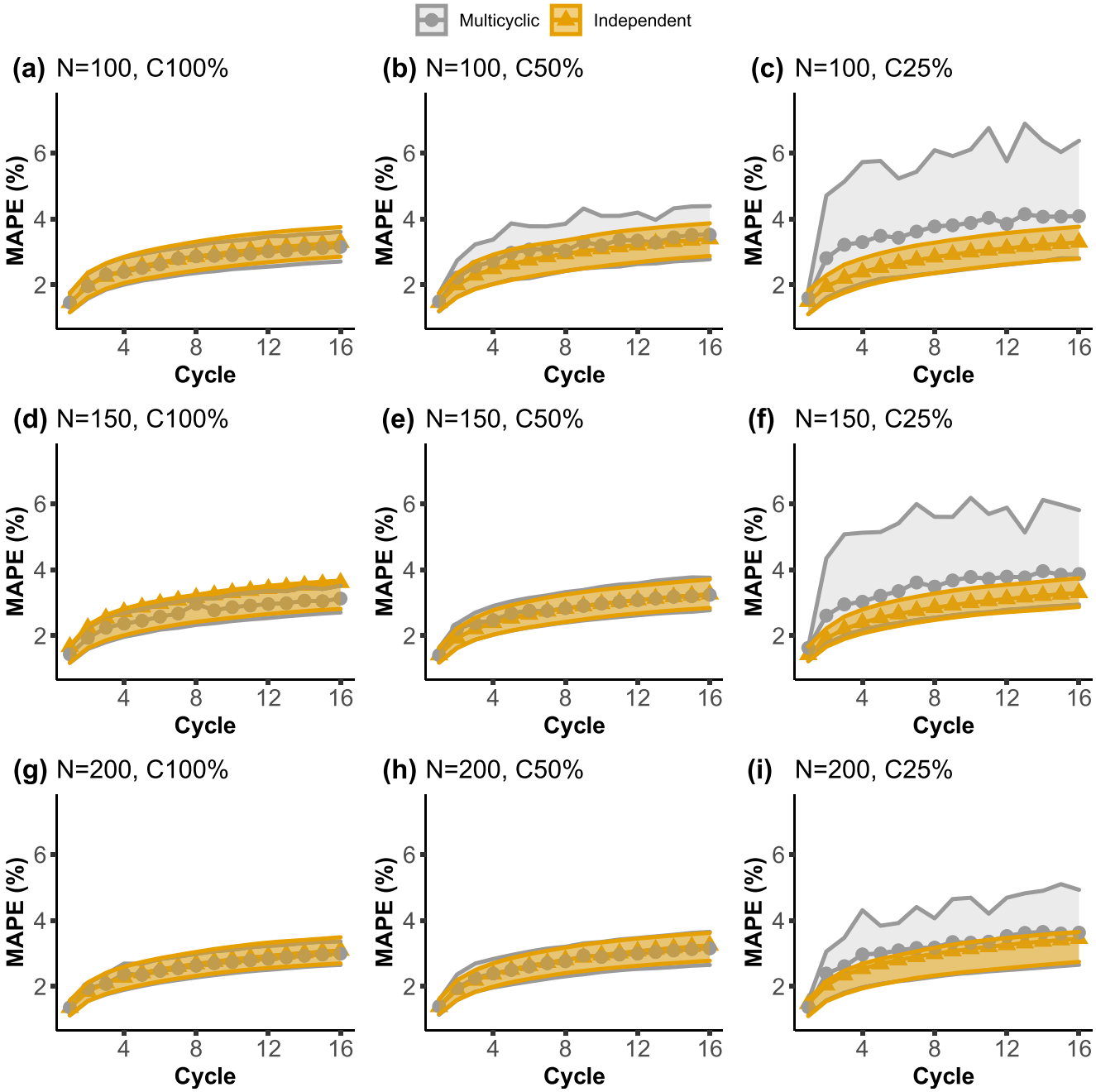

Fig. 12 Evolution of the *MAPE* performance indicator for T_{min} prediction using the independent (grey dot) and multicyclic (orange triangle) metamodelling approaches. Different numbers N of training samples are tested as well as different training options (named "Cp%") with p indicating the proportion of samples retained among N for the training over the cycles N°2 to N°16. The solid line is the mean value calculated from the results of a random subsampling approach conducted 100 times. The envelope limits are computed using the 10th and 90th percentiles

to some extent for option "Cp%". Stable results are obtained for T_{max} and for the rate prediction (grey and red lines in Fig. 15, top, respectively) regardless of the number of training samples, hence confirming our conclusions on the predictive performance. This result is, however, only valid for T_{min} provided that a sufficiently large number of training samples are available ("C75%", Fig. 15a). This is particularly clear in Fig. 15(b,c), where the spread in the results is nonnegligible considering cycles N°2 and N°3. Despite the use of the optimized algorithm developed by Le Gratiet [26], this highlights

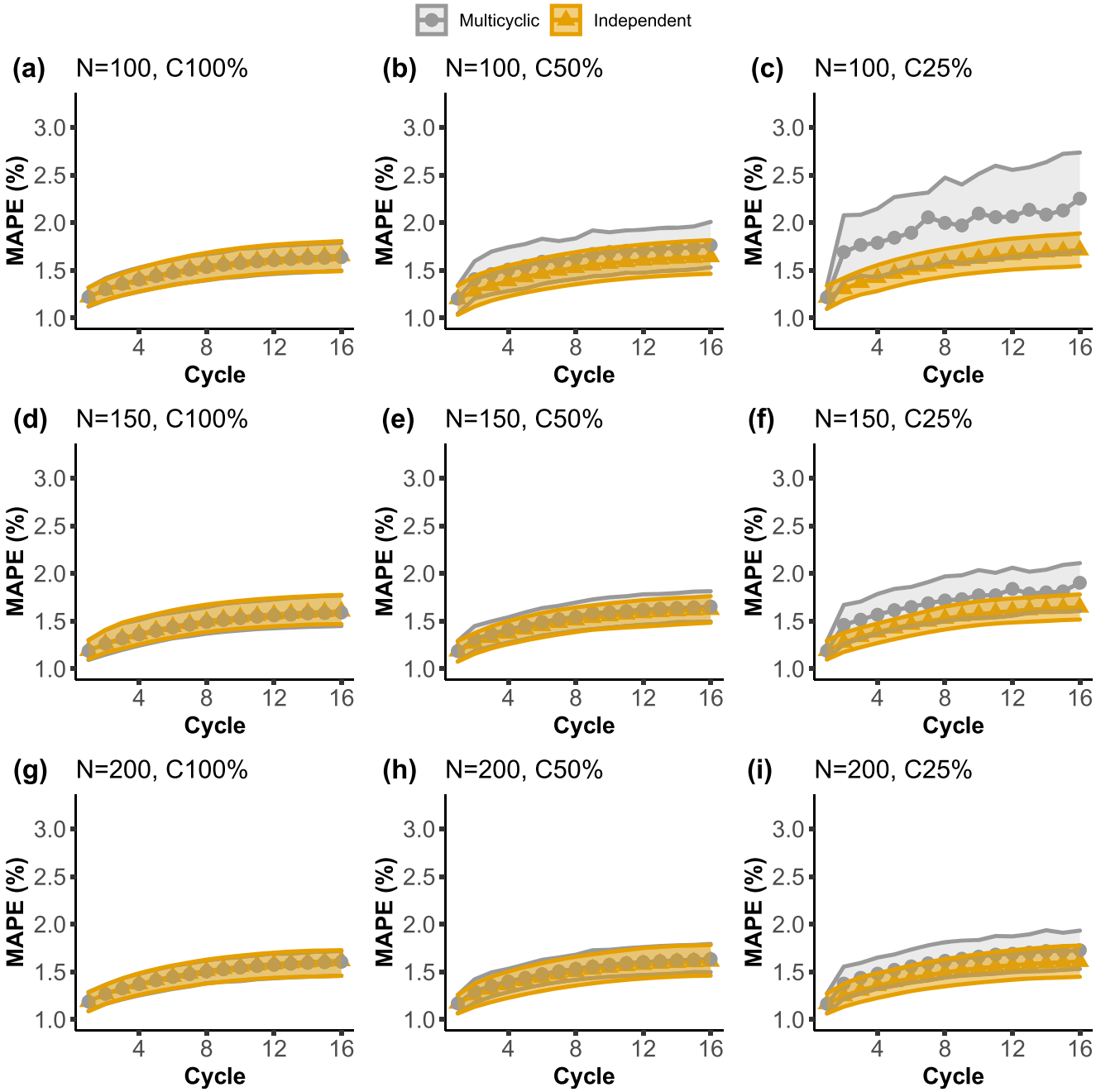

Fig. 13 Evolution of the *MAPE* performance indicator for T_{max} prediction using the independent (grey dot) and multicyclic (orange triangle) metamodelling approaches. Different numbers N of training samples are tested as well as different training options (named "Cp%") with p indicating the proportion of samples retained among N for the training over the cycles N°2 to N°16. The solid line is the mean value calculated from the results of a random subsampling approach conducted 100 times. The envelope limits are computed using the 10th and 90th percentiles

some difficulties in the computation of the parameters of the nested kriging metamodels (termed hyperparameters, see Section 3.2) with a low number of training samples. Some improvements in the stability of our results are observed when tuning the algorithm for hyperparameter optimization (in particular by increasing the number of multistarts of the optimization method for likelihood maximization), but this should be further explored in the future.

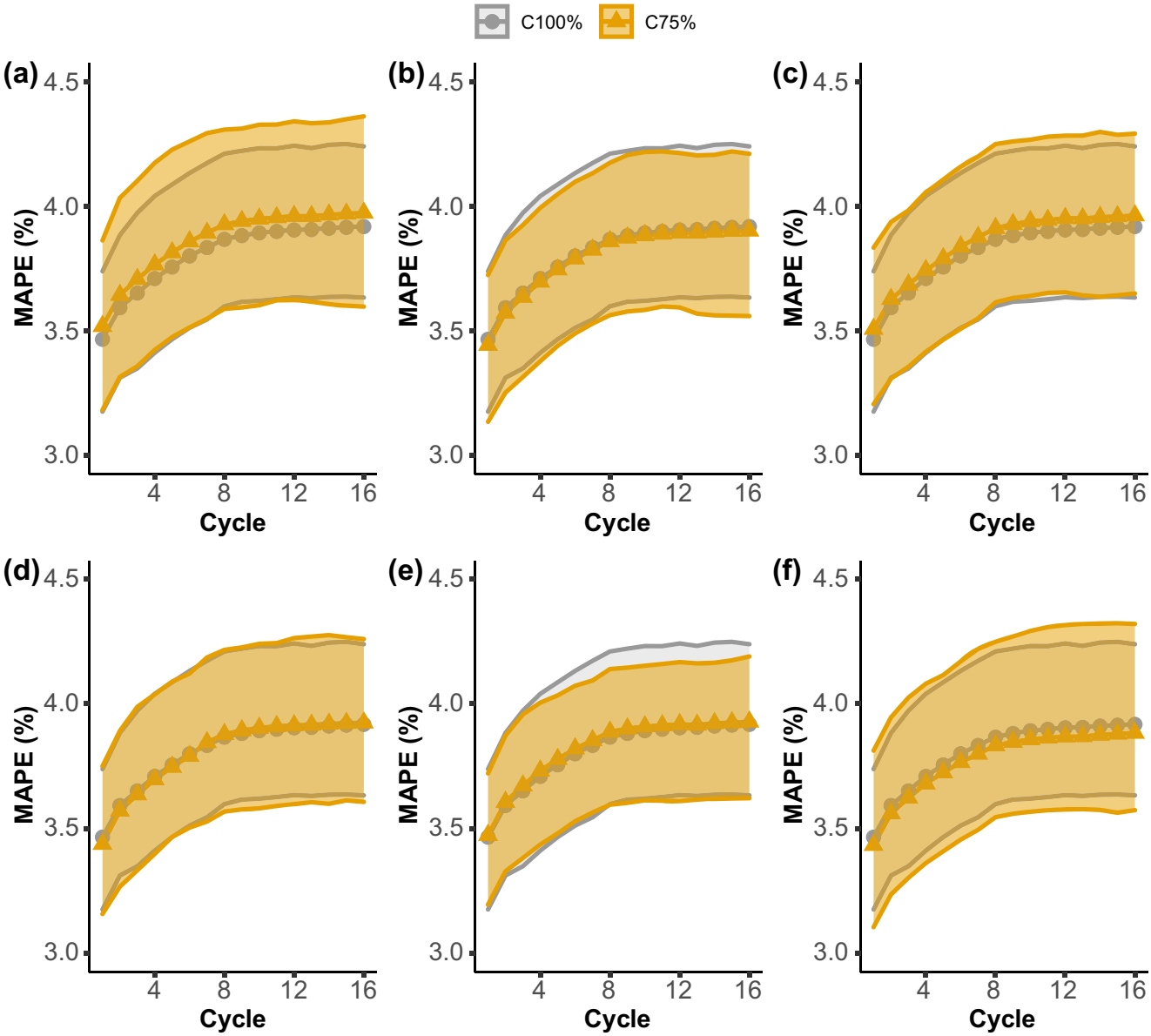

Fig. 14 Evolution of the *MAPE* performance indicator for the prediction of the rate considering different training options (see Section 4.4 for details). The solid line is the mean value calculated from the results of a random subsampling approach conducted 100 times. The envelope limits are computed using the 10th and 90th percentiles

5 Concluding remarks and further work

In this study, we propose a metamodelling approach (named multicyclic metamodelling) based on a multifidelity cokriging model for the prediction of ATES. By using a numerical model of a realistic ATES system in the Paris basin, our tests based on an extensive random subsampling validation approach (by varying the number of training samples from 33 to 66% of the total number of available results) show a higher performance of the multicyclic metamodel (with a decrease in the error indicator by 20–50% on average) compared to the independent metamodelling approach. The major advantage of the proposed approach is the ability to achieve this higher predictive performance while saving computational time cost because the training phase only needs a few tens of complete long-running simulations, i.e., simulations that are run over all

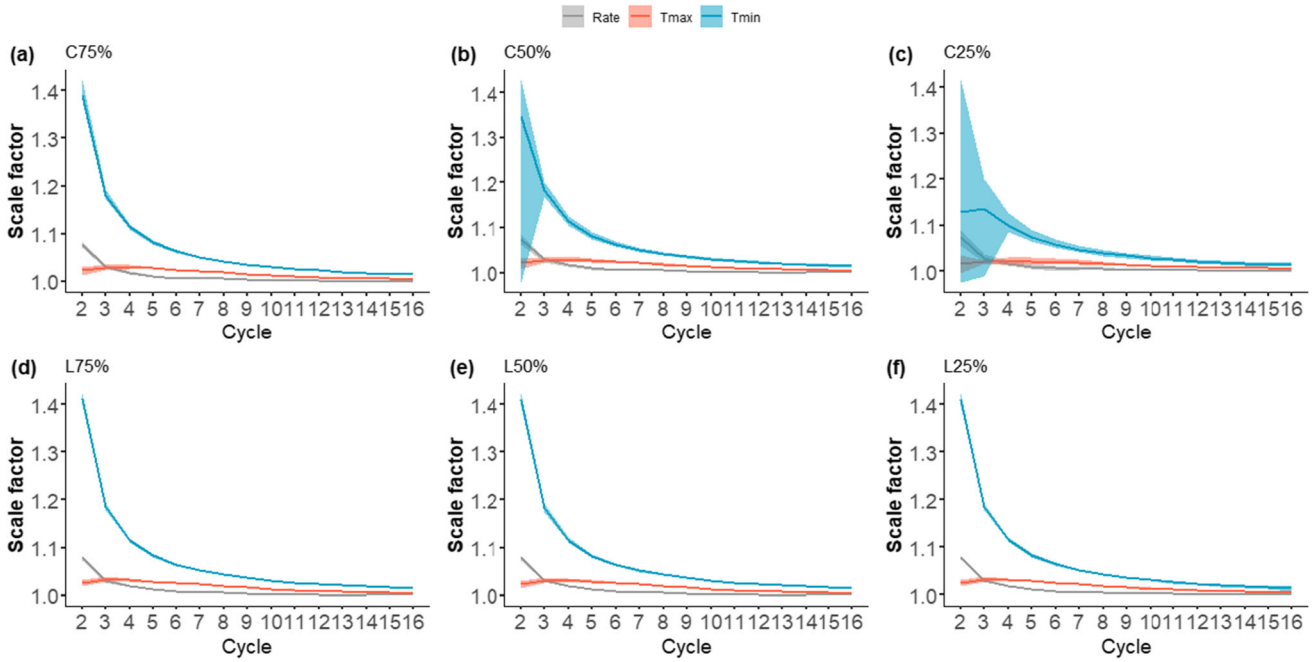

Fig. 15 Mean value of the scale factor considering the different training options detailed in Section 4.4. The envelope is computed from the 10th and 90th percentiles given 100 replicates of the random subsampling procedure

cycles (compared to the hundreds required by the independent metamodelling approach). The influence of the training procedure is tested, and we show that the performance of the multicyclic metamodel is relatively insensitive to it regardless of the proportion of samples that is retained over the cycles (whether a constant or a linear decrease is applied). Finally, we outline that the capability of the multicyclic metamodel to learn the intercyclic dependence is a key aspect for reaching this high predictive capability. In addition, we complement the procedure with a preliminary screening analysis, which enables us to identify the key input variables that influence the most results, hence reducing the number of input variables of the metamodel (here decreased by >70%).

This study should be considered a first feasibility assessment of the multicyclic metamodelling approach. Underground activities other than ATES also involve complex cyclic production/injection phases (such as those in low permeability reservoirs, e.g., CO_2-enhanced gas recovery [33] and storage or production-enhanced cyclic pressure pulsing [34]), and the use of the proposed metamodelling approach should also benefit their prediction. However, to reach a full operative level, several aspects of the approach need to be further improved. First, a simple modelling of the intracyclic evolution was used by fitting an exponentially decreasing function. In other application cases, more complex temporal evolutions at each cycle are expected. A possible approach worth investigating is to combine our approach with a dimension reduction step using, for instance, principal component analysis, as proposed by Thenon et al. [11]. Second, we trained three different multicyclic metamodels, one for each parameter of the exponentially decreasing function (T_{min}/T_{max} and the rate parameter). This means that the interdependence between these three variables is not accounted for here, which might impact the predictive performance. An improvement could rely on a multivariate version of the multifidelity cokriging approach using, for instance, recent developments in this direction, as proposed by Ma et al. [35]. Third, we performed the screening analysis by combining the HSIC measure with a clustering algorithm to account for the dependence of a group of variables. To address this sensitivity analysis with correlated inputs, more advanced methods should be applied in the future by relying, for instance, on the developments of Da Veiga [36] using Shapley effects. Finally, our tests outlined the difficulty of training the multicyclic metamodel when the number of training samples is low (25–50). In addition to the improvements in the optimization of the kriging parameters' estimation, adaptive sampling and sequential design of experiments, as proposed by Liu et al. [37] and Stroh et al. [38], are expected to play a key role.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s10596-023-10192-8.

Acknowledgements This paper has been subsidized by BRGM and the ERANET cofund GEOTHERMICA (grant number 731117) from the European Commission and ADEME, French Agency for ecological transition (grant number 1882C0016). More information can be found on www.heatstore.eu.

Data availability The datasets generated during and/or analysed during the current study are available from the corresponding author on reasonable request. The multicyclic metamodel was implemented in R using the MuFiCokriging package (available at https://cran.r-project.org/web/packages/MuFiCokriging/index.html) and the HSIC-based screening analysis with the sensitivity package (available at https://cran.r-project.org/web/packages/sensitivity/index.html).

Declarations

Conflicts of interests/competing interests The authors have no competing interests to declare that are relevant to the content of this article.

References

1. Mindel, J.E., Alt-Epping, P., Les Landes, A.A., Beernink, S., Birdsell, D.T., Bloemendal, M., et al.: Benchmark study of simulators for thermo-hydraulic modelling of low enthalpy geothermal processes. Geothermics. **96**, 102130 (2021)
2. Amudo, C., Graf, T., Dandekar, R., Randle, J.M.: The Pains and Gains of Experimental Design and Response Surface Applications in Reservoir Simulation Studies. SPE Reservoir Simulation Symposium, The Woodlands, Texas (2009)
3. Zubarev, D.I.: Pros and cons of applying proxy-models as a substitute for full reservoir simulations. Proceedings of the SPE annual technical conference and exhibition (2009)
4. Razavi, S., Tolson, B.A., Burn, D.H.: Review of surrogate modeling in water resources. Water Resour. Res. **48**(7) (2012). https://doi.org/10.1029/2011WR011527
5. White, C.D., Willis, B.J., Narayanan, K., Dutton, S.P.: Identifying and estimating significant geologic parameters with experimental design. SPE J. **6**(03), 311–324 (2001)
6. Pratama, H.B., Supijo, M.C.: Experimental design and response surface method in geothermal energy: a comprehensive study in probabilistic resource assessment. Geothermics. **87**, 101869 (2020)
7. Limbeck, J., Bisdom, K., Lanz, F., Park, T., Barbaro, E., Bourne, S., Kiraly, F., Bierman, S., Harris, C., Nevenzeel, K., den Bezemer, T., van Elk, J.: Using machine learning for model benchmarking and forecasting of depletion-induced seismicity in the Groningen gas field. Comput. Geosci. **25**(1), 529–551 (2021)
8. Fursov, I., Christie, M., Lord, G.: Applying kriging proxies for Markov chain Monte Carlo in reservoir simulation. Comput. Geosci. **24**(4), 1725–1746 (2020)
9. Manceau, J.C., Rohmer, J.: Post-injection trapping of mobile CO2 in deep aquifers: assessing the importance of model and parameter uncertainties. Comput. Geosci. **20**(6), 1251–1267 (2016)
10. Kostakis, F., Mallison, B.T., Durlofsky, L.J.: Multifidelity framework for uncertainty quantification with multiple quantities of interest. Comput. Geosci. **24**(2), 761–773 (2020)
11. Thenon, A., Gervais, V., Le Ravalec, M.: Multi-fidelity metamodeling for reservoir engineering-application to history matching. Comput. Geosci. **20**(6), 1231–1250 (2016)
12. Wilson, K.C., Durlofsky, L.J.: Optimization of shale gas field development using direct search techniques and reduced-physics models. J. Pet. Sci. Eng. **108**, 304–315 (2013)
13. Cabeza, L.F., Martorell, I., Miró, L., Fernández, A.I., Barreneche, C.: Introduction to thermal energy storage systems. In: Advances in Thermal Energy Storage Systems, pp. 1–33. Woodhead Publishing (2021)
14. Kallesøe, A.J., Vangkilde-Pedersen, T.: Underground Thermal Energy Storage (UTES) – state-of-the-art, example cases and lessons learned, HEATSTORE Project report, Geotherm. – ERA NET Cofund Geothermal., (2019). Available at: https://www.heatstore.eu/documents/HEATSTORE_UTES%20State%20of%20the%20Art_WP1_D1.1_Final_2019.04.26.pdf Accessed 22 September 2021
15. Kennedy, M.C., O'Hagan, A.: Predicting the output from a complex computer code when fast approximations are available. Biometrika. **87**(1), 1–13 (2000)
16. Santner, T.J., Williams, B.J., Notz, W.I., Williams, B.J.: The design and analysis of computer experiments, vol. 1. Springer, New York (2003)
17. Lopez, S., Hamm, V., Le Brun, M., Schaper, L., Boissier, F., Cotiche, C., Giuglaris, E.: 40 years of Dogger aquifer management in Ile-de-France, Paris Basin, France. Geothermics. **39**(4), 339–356 (2010)
18. Réveillère, A., Hamm, V., Lesueur, H., Cordier, E., Goblet, P.: Geothermal contribution to the energy mix of a heating network when using aquifer thermal energy storage : modeling and application to the Paris basin. Geothermics. **47**, 69–79 (2013)
19. Hamm, V., Maurel, C.: Low Temperature Deep Geothermal Operations for Direct Use in France: Development of a National Geothermal Database and Last Review. In: Proceedings World Geothermal Congress (2020)
20. Beaude, L.: Numerical simulation of non-isothermal compositional two-phase flows in porous media and its applications to high energy geothermy. PhD dissertation, University of Côte d'Azur (France) (2018)
21. Qian, P.Z.: Sliced Latin hypercube designs. J. Am. Stat. Assoc. **107**(497), 393–399 (2012)
22. Al-Mudhafar, W.J., Rao, D.N., Srinivasan, S., Vo Thanh, H., Al Lawe, E.M.: Rapid evaluation and optimization of carbon dioxide-enhanced oil recovery using reduced-physics proxy models. Energy Sci. Eng. **10**(10), 4112–4135 (2022)
23. Williams, C.K., Rasmussen, C.E.: Gaussian Processes for Machine Learning. MIT press, Cambridge (2006)
24. Roustant, O., Ginsbourger, D., Deville, Y.: DiceKriging, DiceOptim: two R packages for the analysis of computer experiments by kriging-based metamodeling and optimization. J. Stat. Softw. **51**(1), 1–55 (2012)
25. Roustant, O., Padonou, E., Deville, Y., Clément, A., Perrin, G., Giorla, J., Wynn, H.: Group kernels for Gaussian process metamodels with categorical inputs. SIAM/ASA J. Uncertain. Quantif. **8**(2), 775–806 (2020)
26. Le Gratiet, L.: Bayesian analysis of hierarchical multifidelity codes. SIAM/ASA J. Uncertain. Quantif. **1**, 244–269 (2013)
27. Al-Mudhafar, W.J.: Incorporation of bootstrapping and cross-validation for efficient multivariate facies and Petrophysical modeling. Society of Petroleum Engineers Low Perm Symposium (2016)
28. Marrel, A., Iooss, B., Van Dorpe, F., Volkova, E.: An efficient methodology for modeling complex computer codes with Gaussian processes. Comput. Stat. Data Anal. **52**(10), 4731–4744 (2008)
29. De Lozzo, M., Marrel, A.: New improvements in the use of dependence measures for sensitivity analysis and screening. J. Stat. Comput. Simul. **86**(15), 3038–3058 (2016)
30. Da Veiga, S.: Global sensitivity analysis with dependence measures. J. Stat. Comput. Simul. **85**(7), 1283–1305 (2015)

31. El Amri, M.R., Marrel, A.: Optimized HSIC-based tests for sensitivity analysis: application to thermalhydraulic simulation of accidental scenario on nuclear reactor. Qual. Reliab. Eng. Int., in press. **38**, 1386–1403 (2021)
32. Hartigan, J.A., Wong, M.A.: Algorithm AS 136: A k-means clustering algorithm. J. R. Stat. Soc.: Ser. C: Appl. Stat. **28**(1), 100–108 (1979)
33. Kalantari-Dahaghi, A., Mohaghegh, S., Esmaili, S.: Data-driven proxy at hydraulic fracture cluster level: a technique for efficient CO2-enhanced gas recovery and storage assessment in shale reservoir. J. Nat. Gas Sci. Eng. **27**, 515–530 (2015)
34. Artun, E., Khoei, A.A., Köse, K.: Modeling, analysis, and screening of cyclic pressure pulsing with nitrogen in hydraulically fractured wells. Pet. Sci. **13**(3), 532–549 (2016)
35. Ma, P., Karagiannis, G., Konomi, B.A., Asher, T.G., Toro, G.R., Cox, A.T.: Multifidelity computer model emulation with high-dimensional output: An application to storm surge. arXiv preprint arXiv:1909.01836, (2019)
36. Da Veiga, S.: Kernel-based ANOVA decomposition and Shapley effects–Application to global sensitivity analysis. arXiv preprint arXiv:2101.05487, (2021)
37. Liu, H., Ong, Y.S., Cai, J.: A survey of adaptive sampling for global metamodeling in support of simulation-based complex engineering design. Struct. Multidiscip. Optim. **57**(1), 393–416 (2018)
38. Stroh, R., Bect, J., Demeyer, S., Fischer, N., Marquis, D., Vazquez, E.: Sequential design of multi-fidelity computer experiments: maximizing the rate of stepwise uncertainty reduction. Technometrics. **64**, 199–209 (2021)