

# SURROGATE-ASSISTED PARAMETRIC CALIBRATION USING DESIGN OF EXPERIMENT PLATFORM WITHIN DIGITAL TWINNING

MADHU SUDAN SAPKOTA<sup>1</sup>, EDWARD APEH<sup>1</sup>, MARK HADFIELD<sup>1</sup>, ROYA HARATIAN<sup>1</sup>,  
ROBERT ADEY<sup>2</sup> & JOHN BAYNHAM<sup>2</sup>

<sup>1</sup>Faculty of Science and Technology, Bournemouth University Poole, United Kingdom.

<sup>2</sup>CM BEASY Ltd, Ashurst Lodge, UK.

## ABSTRACT

The process of developing a virtual replica of a physical asset usually involves using the best available values of the material and environment-related parameters essential to run the predictive simulation. The parameter values are further updated as necessary over time in response to the behaviour/conditions of physical assets and/or environment. This parametric calibration of the simulation models is usually made manually with trial-and-error using data obtained from sensors/manual survey readings of designated parts of the physical asset. Digital twinning (DT) has provided a means by which validating data from the physical asset can be obtained in near real time. However, the process of calibration is time-consuming as it is manual, and as with each parameter guess during the trial, a simulation run is required. This is even more so when the running time of a single simulation is high enough, like hours or even days, and the model involves a significantly high number of parameters. To address these shortcomings, an experimental platform implemented with the integration of a simulator and scientific software is proposed. The scientific software within the platform also offers surrogate building support, where surrogates assist in the estimation/update of design parameters as an alternative to time-consuming predictive models. The proposed platform is demonstrated using BEASY, a simulator designed to predict protection provided by a cathodic protection (CP) system to an asset, with MATLAB as the scientific software. The developed setup facilitates the task of model validation and adaptation of the CP model by automating the process within a DT ecosystem and also offers surrogate-assisted optimisation for parameter estimation/updates.

*Keywords:* BEASY, cathodic-protection, digital twin, model adaptation, software integration.

## 1 INTRODUCTION

The robust prediction of the behaviour of the structure/system is essential to plan for efficient usage of it and/or avoidance of future failure/breakdown. Engineers use simulation tools to predict the performance of the system and to determine the risk to the structure from degradation mechanisms such as corrosion and fracture. To make the prediction, the parametric simulation models derived from the simulators, i.e., process simulating software(s) with the best available standard values parameters, are used more often.

The parameters typically represent the model's inputs describing the properties of the materials and the environmental conditions that the structure experiences. These model input variables cannot be directly measured due to structural complexity in most situations. Parameter setting for the model during the realisation of a virtual replica of an existing physical structure thus relies upon the structural response-data obtained from sensors or inspection surveys. During such parametric calibration, the best set of parameters is found by correlating the model output to the available measurements from the physical system [1].

The traditional approach to calibration when performance validating data are available from the structure is based on a trial-and-error procedure involving manual iterative analyses and modifications. This task is normally performed by the engineers using their experience to choose the appropriate values until good correspondence is obtained (Fig. 1). As the parameters change with time, the process is repeated over the life of the structure based upon the requirements determined by experts and the availability of the dataset. Furthermore, relying upon physics-based model simulation data, analytical tasks such as uncertainty analysis, essential for calibration and adaptation, are hindered as multiple simulation runs are required. This results in redundant computational efforts and costs during model calibration/updating for complex models and in a repetitive manner.

The virtual replica of an existing physical counterpart has been recognised as a digital twin (DT) in recent years. Predictive DTs with self-adaptive simulation characteristics have been proposed [2] to provide robust simulation throughout the lifetime despite changes in the real asset and environments. To achieve the adaptive DT, the necessity of data-based analysis is ubiquitous and accepted by a wide range of literature [3, 4]. However, while enabling DT that demands the analytical features encapsulated/incorporated within it, the features are not necessarily available on the same software platform as for simulation.

This work utilises software integration to enable an experimental design platform to automate the DT realisation process exclusively with the analytical support required for adaptation. The platform utilising analytical features such as approximation (surrogate) modelling expedite the parameter estimation process, ultimately supporting to enable the most accepted aspect of DT, i.e. self-adaptation.

2 DIGITAL TWIN CONCEPT WITH SELF ADAPTATION POTENTIAL

The choice of model types to rely upon for the predictive analysis is mostly dependent on the problem domains. Though data-driven models are preferred when enough data is available, physics-based models are useful to provide insight into the complex phenomena that have not yet occurred. Likewise, from the state-of-art regarding DT, we can infer that to yield robust and effective digital twin, statistical approaches are essential together with physics-based modelling [2–4]. Therefore, both physics-based and data-driven models are important from

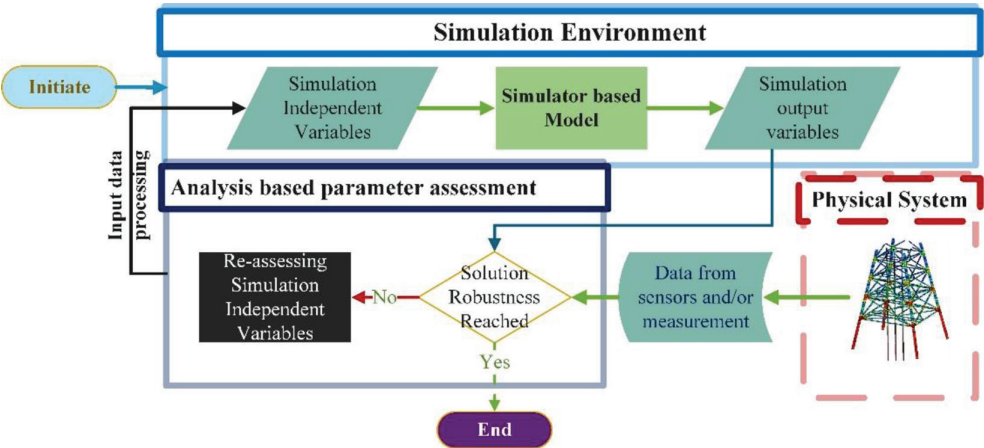


Figure 1: The parametric simulation model parameter updating for performance enhancement with output analysis (trial-and-error method).

the DT perspective. In some cases, a digital twin that incorporates both physics model(s) and data model(s) is referred to as a hybrid digital twin [5]. This section will discuss the digital twin concept, especially focused on two major model calibration-related problems, which are (a) time-consuming manual parametric calibration and (b) the requirement of multiple-time simulator running that induce delay.

### 2.1 Parametric modelling using simulator (physics-based solver)

The use of a previously validated model at the conceptual level assists in the design of not just one simulation model but many within the problem domain [6, 7]. The parametric simulation model building in multiple domains is already facilitated by the availability of commercial process simulators (simulating software). Such simulators are generally based on the current scientific understanding (physics of phenomena), often involving numerical approximation of differential equations, and they are implemented in complex computer programs [8].

When the commercially available simulator(s) meets the primary modelling requirements, the major efforts in realising virtual replica of the physical asset are focused on determining model independent variables (parameters) such as material and environmental-related parameters. The trial-and-error approach (Fig. 1) based on the traditional method during parametric calibration is inefficient and introduces time-delays due to manual performance. It also has the drawback of not having the ability to calibrate multiple parameters simultaneously and does not guarantee the best results. Furthermore, the real-time calibration or adaptation is hindered when multiple inputs with varied simulation results are required for the estimation/ updating process. This is because complex models with discretized PDEs solver have a long running time, and model performance evaluation requirements over spatially and/or temporally varying parameter(s) space as inputs to such complex models. Though enriched with the core process simulation solver, the simulation environment requires analytical support that overcomes the drawbacks of traditional methods while tailoring the parametric model.

Experimental design, sensitivity analysis, and design (parameter) optimisation are widely adopted techniques within the systematic procedures of model performance enhancement [9]. Design optimisation is an approach that combines mathematical optimization algorithms with a parametric simulation model to search the design (parameter) space for the optimal solution [10]. Gradient-based and non-gradient-based algorithms are used during parameter optimisation [11, 12] for reaching the best fit.

However, the rise in the number of the model's input parameters increases the complexity of optimisation assisted physics model calibration, also known as the curse of dimensionality. This leads to the risk of either ending up at the local minima during gradient-based optimisation or consuming significantly more time, i.e. could be a day or even a week while using non-gradient-based optimisation. To address this shortcoming, especially the dependency upon time-consuming simulation runs for model adaptation, surrogate modelling can provide a substitute for the physics-based model. The data-driven (surrogate) models are utilised to assist the calibration of the physics-based model when simulation running time is significantly high, thus demanding approximation model-based optimisation [13].

### 2.2 Overview of modelling and simulation in digital twinning

The DT concept is a new paradigm of interest in the field of modelling and simulation having online simulation as the core functionality of the system employing seamless assistance [14].

The definitions of DT have been evolving rapidly in recent years as more insights are added to its features/potentials. The DT is a virtual model of the physical object with the potential of understanding changes in the status of the physical system through sensing data and adapting itself followed by some analytical procedure [3]. Henceforward, DT is being preferred over the model in the context of adaptive simulation [4]. Furthermore, the aspects of self-parameterisation and self-adaptation that allow DT to resemble its counterpart physical twin are also of high importance to realize the potential of artificial intelligence (AI) within DT [15].

Successful deployment of a digital twin with the potential of self-adaptation is moreover dependent on the data and the approach for calibration of the model in real time. With the development of new sensor technologies providing the opportunity for more real-time data from the physical twin, more focus is required on the latter one. However, the limitations regarding the online calibration or adaptation of a model discussed in Section 2.1 persist as DT at its application level is still in its infancy.

Also, the digital twin concept is not only about a single behavioural simulation, but it also involves multiple process co-simulations representing different behavioural phenomena. This highlights the DT as collaborative models to provide a comprehensive representation of the system's simultaneously occurring multiple dynamic phenomena. Going along with the concepts of DT, while enabling DT as a high-fidelity simulation model provided with self-analysis features or DT with co-simulations, software integration is anticipated.

### 2.3 Surrogate models in digital twin's realisation

The necessity of surrogates in DT realisation has been discussed since the concept of DT was introduced in the field of simulation and modelling [16]. This discussion till today has mostly been limited to the conceptual level with few applications. However, the works of literature have accepted surrogate modelling as an important aspect of the DT paradigm, which makes DT different to other physics-based or data-driven models [4].

Surrogates are the data-driven models built from the simulation data with significantly less computation time than the physics-based model. Surrogate models in this paper refer to black-box models that are typically created by training a model using datasets obtained from a physics-based model simulation run. Datasets are generated using the Design of Experiments (DOEs) methods for choosing a different combination of input variables, such as Latin Hypercube, Box–Behnken, and central composite design (CCD) [17]. Then the data-driven model, i.e. surrogate, is trained using the datasets. The response surface method (RSM) with polynomial regression fit, artificial neural network (ANN), and Kriging are the commonly used methods for constructing a surrogate in the engineering fields [18].

Leveraging the concept of hybrid predictive models in DT, the surrogates here are considered the data-driven models, which serve multiple purposes as a substitute to the physics-based model. Though surrogates are suggested as an alternative when faster prediction is required or to ease the process of parameter estimation, physics-based full-order model(s) will remain as the core and robust predictive tool of the DT.

## 3 PLATFORM AND RESPONSE SURROGATES IN DT REALISATION

Towards, the real-world realisation of the DT concept, experimentation-based adaptation is essential to tailor the physics-based simulator to have a robust predictive ability. For this, a platform with automated data and control signal flow between the major tool (simulator(s)) and the supporting tool(s) for the analytical task is proposed.

### 3.1 Server platform

Sapkota *et al.* (2021) proposed integration of the simulator and scientific computing software as an approach to facilitate the automation of tailoring a process simulator to enable DT. In the proposed platform and approach, the collaborative/comprehensive platform required for DT realisation from a Simulator is accomplished with scientific software as the server. The scientific software should either be enriched with the analytical support providing tool(s) or at least provide an interface to the external tool(s) [19].

Utilising the integrated platform, this paper will discuss the surrogate-assisted timely parameter estimation for a simulator-based DT, therefore, providing a path to enable a self-adapting model that can be ultimately referred to as a Digital Twin. As data-driven surrogate models, for example, polynomial-fit models, are utilised, the data-driven modelling tools available in the scientific software aid in such hybrid DT realisation. Also, different design of experiments (DOE) algorithms like CCD, Latent-hypercube, and BBD required for the approximation modelling are available within scientific software like MATLAB [22].

The server software furthermore provides the opportunity for the visualisation of the data to denote the state and discrepancies in the performance of the physics-based model from the actual system.

### 3.2 Response surrogates and optimisation assisted parameter estimation

The surrogate, a substitute model, offers to provide an approximation solver with the desired less computation time. Response surface (RS) methodology among one of the surrogate building methods is a collection of mathematical and statistical techniques based on the fit of a polynomial equation to the input-output dataset. A regression model (mostly second-order polynomials) is subsequently used to fit the relationship between the output responses and the input variables. A representative second-order polynomial RS equation is presented as eq 1. [20, 21].

$$Y = \beta_0 + \sum_{i=1}^k \beta_i X_i + \sum_{i=1}^k \beta_{ii} X_i^2 + \sum_{i < j} \sum_{j=2}^k \beta_{ij} X_i X_j + \varepsilon \quad (1)$$

where ‘Y’ is the estimated output of the RS model, and  $X_i$  is the input parameters or variable of the total count  $k$ . Then,  $\beta_0, \beta_i$  are the regression coefficients calculated using the method of data-fitting.

The server uses an ‘offline’ phase to create the surrogate model of the system using the simulation input-output dataset from the physics-based model. The surrogate, i.e. data-driven model for performance validation, can be used in the ‘online’ phase. In the ‘online’ phase, the adequately efficient surrogate(s) is used to approximate optimum parameters within the server (Fig. 2). For this, sufficient counts of searches to reach the global minima are attainable in significantly reduced time for the optimisation-based parameter estimation.

## 4 CASE STUDY – SIMULATOR-BASED CATHODIC PROTECTION DIGITAL TWIN REALISATION

### 4.1 Background

Cathodic protection (CP) is most frequently used for the protection of underground or underwater (seawater) metallic infrastructures from corrosion. The performance of the CP system

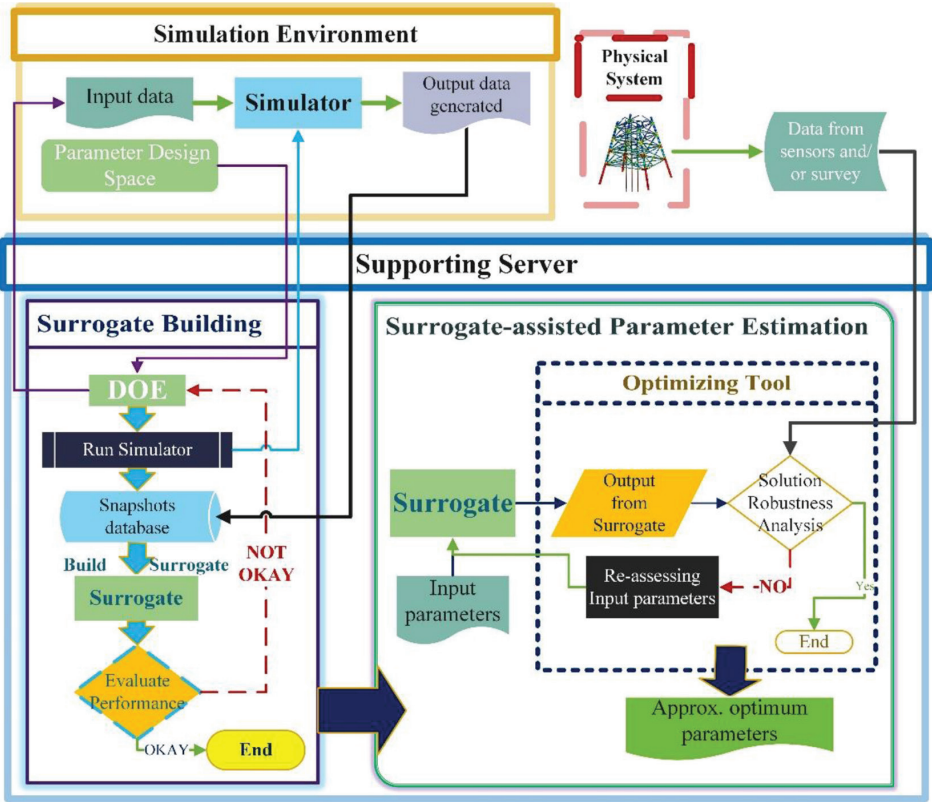


Figure 2: Integration of server to the simulator for surrogate-assisted parameters approximation during the realisation of DT with self-adaptation potential.

can be evaluated and optimised based on the design rules using a CP simulation model which predicts year by year the protection potentials, the depletion of the anodes, and, in the case of Impressed-Current-Cathodic-Protection (ICCP), the current to be required by the system over its life [23].

In reality, the actual performance of the CP system will often be different than the CP model as coatings, for example, often degrade at different rates to that described in the design rules, environmental conditions may vary, the ‘as-built’ structure may be different, and changes and retrofits are made over time. Integrating the CP data collected during the routine inspection surveys with the CP simulation model can be used to calibrate/adapt the model to match the inspection data. This forms the basis for enabling a ‘digital twin’ of the structure [24]. Yet, to maintain the robustness of predictive DT, the process needs repetitive implementation with each new inspection report unless the pattern of adaptation is tracked or/and the process is automated.

#### 4.2 BEASY: simulator for cathodic protection digital twin

The BEASY tool, a commercial parametric simulator designed to simulate the behaviour of galvanic corrosion problems and cathodic protection designs, is adopted to represent a virtual replica of a CP system.



The CP simulation model to represent the CP system replicates the electrochemical behaviour of the system. In the system, the materials of the structure act as the electrodes along with seawater as electrolytes. The basis of the simulation is the estimation of the distribution of electrical potential and protection current density on the surfaces of the electrodes. The numerical computation of electric potential and current density distribution is based on the solution of the following well-known Laplace partial differential equation [23].

With the assumption that electrolyte is homogeneous,

$$-\nabla(k\nabla\varphi)=0 \tag{2}$$

where,  $k$  = electric conductivity,  $\varphi$  = electric potential, and  $\nabla$  is Nabla operator.

BEASY-CP provides a numerical approximation of Laplace’s equation obtained with the boundary element method (BEM) [9, 23] for steady-state corrosion. Also, the commercial BEASY tool facilitates geometrical modelling and meshing of the geometrical model required for numerical approximation. The data about geometry, meshing, and materials and surrounding related parameters can be exported to text files and fed to the solver for numerical approximation (Fig. 3).

The model predicted results such as the protection potential and current density are compared and calibrated with data collected during the routine inspection surveys of the physical CP system. Calibration is then made with repetitive analysis and by updating the parameters until a good agreement is obtained.

A CP simulation model of a marine structure (Fig. 4) protected by sacrificial anodes, i.e. a CP system was built using the BEASY tool. The model parameters required by the Cathodic-Protection model for the CP system of structure (Fig. 4) are the following:

1. Polarisation Behaviour: This is the relationship between potential and current density representing the electrode kinetics of the metal in the seawater. This provides the boundary condition while solving the numerical problem.
  - Polarisation curve for Material 1 of the structure (Fig. 4).
  - Polarisation curve for Material 2 of the structure (Fig. 4).
2. Conductivity/Resistivity: Surrounding Medium/Material related.
  - Seawater-related conductivity (Siemen/m).
  - Seabed-related conductivity (Siemen/m).

The parameters selected for calibration in this case study are the most involved material-related properties, i.e. *Material 1 related polarisation behaviour* and *Seawater conductivity*.

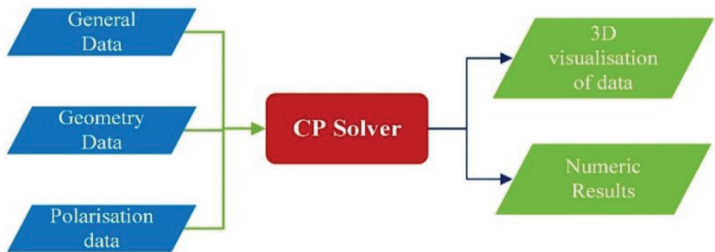


Figure 3: Inputs-Outputs for BEASY tool based Cathodic-Protection Simulation.

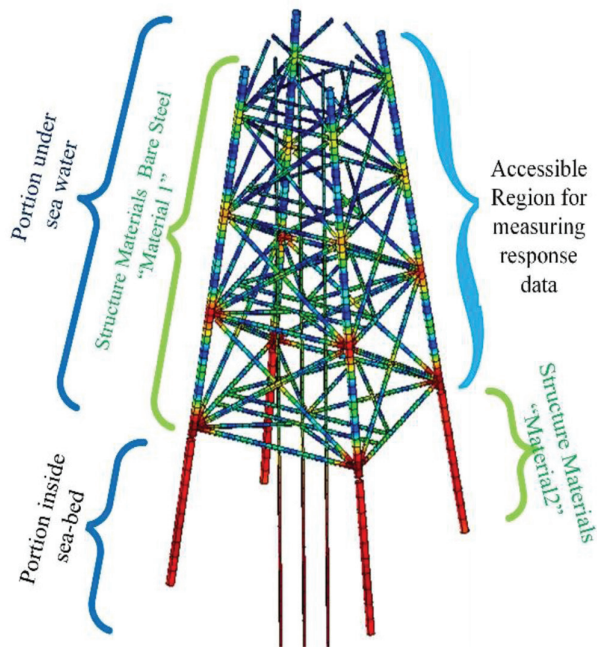


Figure 4: The project adopted the geometrical model of the marine structure protected with sacrificial anodes.

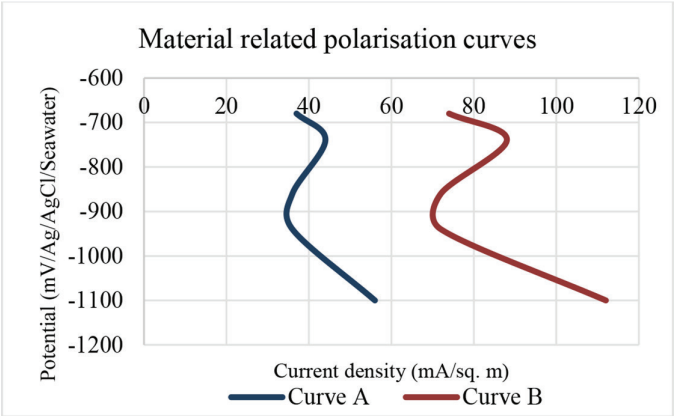


Figure 5: Two different polarisation curves can be transformed from one to the other with some transformation value.

As the polarisation curves for representing the polarisation behaviour are graphical representations (Fig. 5) and dynamics with time, a quantitative representation should be established for re-adjustment of the polarisation data. To address the possible change in polarisation behaviour, the curve transformation value (expanding or squeezing factor) is taken as a variable (parameter) keeping the curve constant as obtained from the design rule. This parameterisation concept can be understood as a modification of the diffusion limiting current (or the barrier properties provided by any coating) in the polarisation behaviour of the materials involved. The transformation vector or parameter is termed as '*p*-value' in this case study.



### 4.3 Software integration with MATLAB as a server

Analytical tools for model performance validation and data-fitting modules are facilitated within scientific software like MATLAB. The extensive data analysis, data-modelling tools availability, plotting capability, and the availability of different optimisation algorithms within MATLAB enable the completion of assess-modify-check loops in reduced computational time. MATLAB, a scientific software granted with these qualities, was selected as the server for the integration platform to enable DT for a CP system.

### 4.4 Surrogate(s)-based parameter estimation and analysis

#### 4.4.1 CP model's performance validating/calibration data and validating criteria

The data types that can be practically obtained from the inspection of a structure and corresponding CP model simulation are (a) surface potential (mV), (b) normal current density (mA/m<sup>2</sup>), and (c) electric field (mV/m) [23]. However, the data dependency for validation depends upon the complexity of the model and data availability from physical assets considering the cost/effort involved.

At this stage, validating data are generated from a simulated virtual reference model provided with different parameter values considering the possible polarisation behaviour in the future. Two types of validating data considered are surface potential (mV) and normal current density (mA/m<sup>2</sup>), and the validating data position counts are 17 and 6, respectively, from the structure's surface (Fig. 6).

#### 4.4.2 Design of experiments for surrogate building

The success of having the best representative model depends upon finding the trade-off between the computational cost associated with simulation runs to generate data suggested by Design of Experiments (DOE), the accuracy of the surrogate and the problem size [25]. Regarding experimental design for Response-Surface-Modelling, the central composite design (CCD) has been frequently discussed [26] and adopted in this case study. The reason that CCD is the most implemented sampling method and prominent in DOE in most engineering problems, made it the first choice in the CP model.

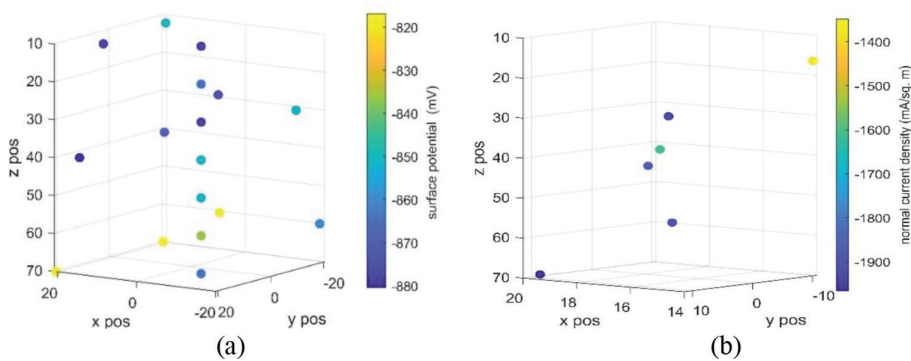


Figure 6: The calibration data was obtained after a simulation run from the reference model (a) surface potential and (b) normal current density, from a selected validating position of the structure's surface (Fig. 4).

The inscribed central composite design (CCD) is selected, which gives 8+1 sample points for 2 variables case. MATLAB based tool *ccdesign* is used for generating the sampling points with CCD for the two independent variables *Material 1 related p-value* and *seawater conductivity*.

The central point is used multiple times to provide high weightage (impact) at the centre of the parameter space.

Total samples = 8 + 4 (center)

#### 4.4.3 Surrogate model for CP system

The response surface method (RSM) with polynomials regression fit (eqn 1) is implemented for surrogate building. As response variables are affected by the interactions of independent variables, the polynomial response surface model (RSM) is an effective tool as a data-fitting model. Now the parameters (independent variables) involved for RSM in the given case are the material-related properties, i.e. polarisation behaviour indicating the *p-value* of the curve and seawater conductivity. The surrogates are built for each of the nodes (17+6) on which the validation/calibration data are accessible/available. The MATLAB based tool '*fit*' is used for the second-order polynomial fit, taking into account two independent variables, which are *Material 1 related p-value* and *seawater conductivity*.

For performance evaluation of the surrogates, the comparative analysis between surrogate output and the full-order model simulation output for different parameter combinations (e.g. Fig. 7b.) is made. The surrogate(s) after performance evaluation on their predictive efficiency within the margin of error provided can now be used as an alternative model for optimisation-based parameter estimation.

#### 4.4.4 Parameter estimation with surrogates

Having the polynomial surrogate now enables performing the exploratory search for the possible combinations of the parameters to obtain the objective function output with less computation time. With two different validating/calibrating data types (4.4.1), the normalised mean square difference between calibrating and model output data with a weightage constant (2:1) is taken as the validating criteria. This ultimately will be used as the objective function during the minimisation-based parameter estimation.

The objective function value is plotted over the possible solution range (Fig. 8) for the respective combination of the parameters. An objective value is calculated using the response

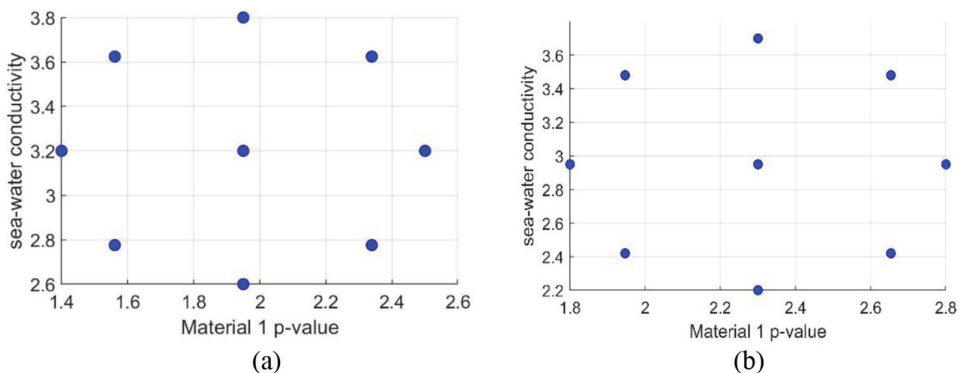


Figure 7: (a) Sample points from CCD for two selected parameters for surrogate building and (b) generated parameter values for surrogate model performance testing.

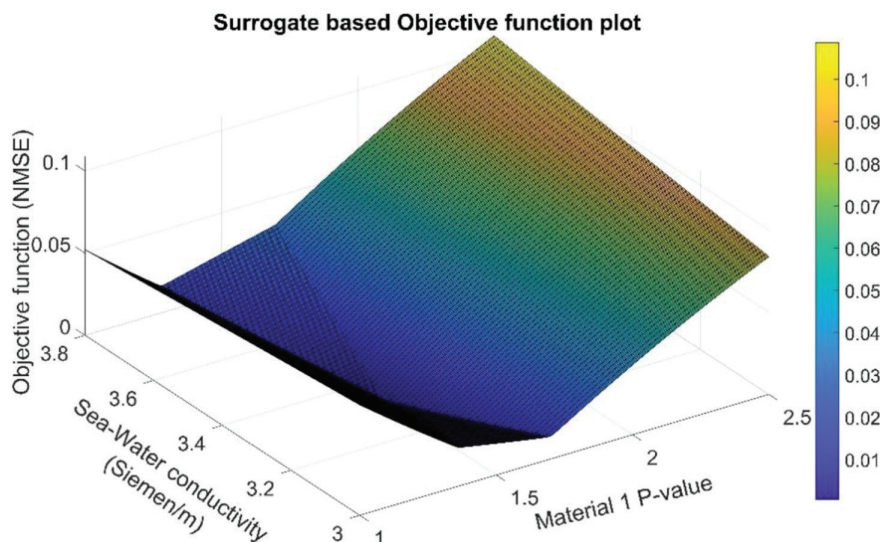


Figure 8: Objective plot for over the parametric space with the fixed validating response data and response data obtained from the surrogate for different parameter inputs.

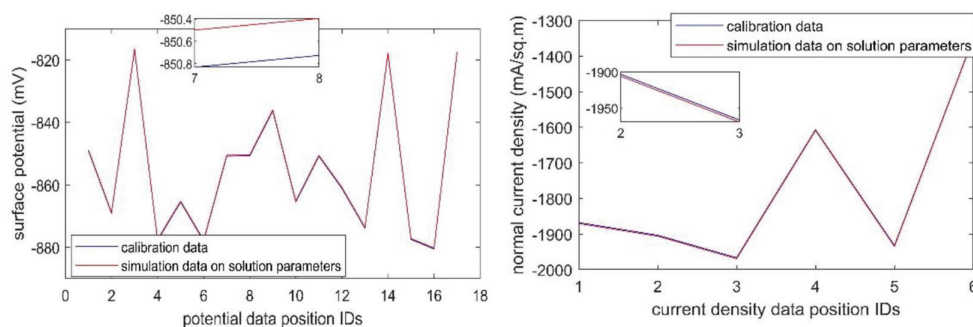


Figure 9: Comparison of (a) surface potential and (b) normal current density data from simulation results with solution parameter achieved and the calibration data.

data predicted from the surrogate with the parameters and the fixed calibration/validation data. The best parameters combination is suggested by the global minima point of the plot.

Taking advantage of the substantially reduced prediction times provided by the surrogate model, the exploratory search could find the global minima in the presented case study. However, in other situations when the exploratory search tends to induce a delay, gradient or non-gradient-based optimisation algorithms/tools are suggested.

Figure 9 demonstrates the performance of the solution model reached with estimated/calibrated parameters, i.e. *polarisation curve for Material 1* and *seawater conductivity* obtained using the calibration data (4.4.1). Likewise, the *polarisation curve for Material 1* obtained with the above-implemented *p*-value estimation approach can be visualised in Fig. 10.

This parameter estimation, which could take many hours or days when performed manually using a physics-based simulation, is reduced to less than a few hours including the offline cost of surrogate building. The importance of the platform is highlighted by the significant

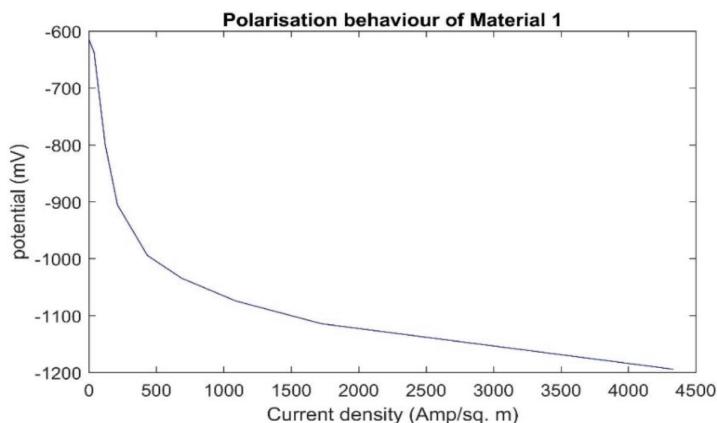


Figure 10: Polarisation behaviour for *Material 1* suggested after surrogate assisted minima finding-based parameter estimation.

reduction in model calibrating time assisted with approximation modelling using the automated experimental platform compared to the manual approach.

## 5 CONCLUSION

This paper discussed the design of experiment platform to establish and maintain the simulation performance during a digital twin realisation of a physical asset. The scientific software-based integrated platform provides the basis of automation of DT enabling from the pre-available simulator(s), utilising supports from the analytical tool(s). This is further supported by the ability of the platform in the surrogate building to ultimately reduce the number of required simulation cases for the optimisation assisted parameter estimation. Surrogates built in the offline phase utilising simulation input-output data are used online as alternative models for timely estimation of the parameters.

The case study demonstrates the application of the integration platform and the importance of surrogates to achieve a cathodic-protection digital twin. Surrogate-assisted optimisation was undertaken with automation to estimate the polarisation data and seawater conductivity of a CP model utilising the given calibration data. The virtual replica of the CP system on calibration provides the predictive capabilities to forecast future levels of protection.

The approach highlights the importance of surrogates in parameter estimation towards establishing the core aspect of DT, i.e. self-adaptation. This suggests and advocates approximation (surrogate) modelling as a crucial aspect within the DT ecosystem in the context that the definition of DT is yet to be standardised. The proposed approach is generic and can be implemented beyond the corrosion and cathodic-protection domains.

## ACKNOWLEDGEMENTS

This work has been undertaken as part of a match-funded PhD research project between Computational Mechanics International Limited and Bournemouth University, UK.

## REFERENCES

- [1] Oliveira, H.L. and Leonel, E.D., Constitutive relation error formalism applied to the solution of inverse problems using the BEM. *Engineering Analysis with Boundary Elements*, **108**, pp. 30–40, 2019.

- [2] Rasheed, A., San, O. & Kvamsdal, T., *Digital twin: Values, challenges and enablers*. *arXiv preprint arXiv:1910.01719*. 2019.
- [3] Barricelli, B.R., Casiraghi, E. & Fogli, D. A survey on digital twin: Definitions, characteristics, applications, and design implications. *IEEE Access*, **7**(MI), pp. 167653–167671, 2019. <https://doi.org/10.1109/ACCESS.2019.2953499>
- [4] Wright, L. & Davidson, S. How to tell the difference between a model and a digital twin. *Advanced Modeling and Simulation in Engineering Sciences*, **7**(1), pp. 1–3, 2020. <https://doi.org/10.1186/s40323-020-00147-4>
- [5] Chinesta, F., Cueto, E., Abisset-Chavanne, E., Duval, J.L. & El Khaldi, F., Virtual, digital and hybrid twins: a new paradigm in data-based engineering and engineered data. *Archives of computational methods in engineering*, 27(1), pp. 105–134, 2020.
- [6] Abdelmegid, M.A., González, V.A., O'Sullivan, M., Walker, C.G., Poshdar, M. & Ying, F., The roles of conceptual modelling in improving construction simulation studies: A comprehensive review. *Advanced Engineering Informatics*, **46**. <https://doi.org/10.1016/j.aei.2020.101175> <https://doi.org/10.1016/j.aei.2020.101175>
- [7] Robinson, S., Arbez, G., Birta, L.G., Tolk, A. & Wagner, G., Conceptual modeling: definition, purpose and benefits. In *2015 Winter Simulation Conference (WSC)* (pp. 2812–2826). IEEE, 2015.
- [8] Brynjarsdóttir, J. & O'Hagan, A. Learning about physical parameters: The importance of model discrepancy. *Inverse Problems*, **30**(11), p. 114007, 2014. <https://doi.org/10.1088/0266-5611/30/11/114007>
- [9] Law, A.M., Kelton, W.D. & Kelton, W.D., *Simulation Modeling and Analysis*. New York: McGraw-Hill; 2000.
- [10] Papalambros, P.Y. & Wilde, D.J., *Principles of Optimal Design: Modeling and Computation*. Cambridge University Press, 2000.
- [11] Venter, G., Review of optimization techniques. *Encyclopedia of Aerospace Engineering*. 2010.
- [12] Whitley, D., Rana, S., Dzuber, J. & Mathias, K.E., Evaluating evolutionary algorithms. *Artificial Intelligence*, **85**(1–2), pp. 245–276, 1996. [https://doi.org/10.1016/0004-3702\(95\)00124-7](https://doi.org/10.1016/0004-3702(95)00124-7)
- [13] Han, Z.H. & Zhang, K.S., Surrogate-based optimization. *Real-World Applications of Genetic Algorithms*, **7**, p. 343, 2012.
- [14] Boschert, S. & Rosen, R., Digital twin—the simulation aspect. In *Mechatronic Futures* (pp. 59–74). Springer, Cham, 2016.
- [15] Schleich, B., Answer, N., Mathieu, L. & Wartzack, S., Shaping the digital twin for design and production engineering. *CIRP Annals*, **66**(1), pp. 141–144, 2017. <https://doi.org/10.1016/j.cirp.2017.04.040>
- [16] Tuegel, E.J., Ingrassia, A.R. Eason, T.G. & Spottswood, S.M., Reengineering aircraft structural life prediction using a digital twin. *International Journal of Aerospace Engineering*, **2011**, 2011.
- [17] Giunta, A., Wojtkiewicz, S. & Eldred, M., Overview of modern design of experiments methods for computational simulations. In *41st Aerospace Sciences Meeting and Exhibit* (p. 649). 2003.
- [18] Rajaram, D., *Methods for Construction of Surrogates For Computationally Expensive High-Dimensional Problems* (Doctoral dissertation, Georgia Institute of Technology) 2020.
- [19] Sapkota, M.S., Apeh, E., Hadfield, M., Adey, R. & Baynham, J., Design of experiments platform for online simulation model validation and parameter updating within digital twinning. *WIT Transactions on Engineering Sciences*, **130**, pp. 3–14, 2020.

- [20] Bezerra, M.A., Santelli, R.E., Oliveira, E.P., Villar, L.S. & Escaleira, L.A., Response surface methodology (RSM) as a tool for optimization in analytical chemistry. *Talanta*, **76**(5), pp. 965–977, 2008.
- [21] Myers, R.H., Montgomery, D.C. & Anderson-Cook, C.M., *Response Surface Methodology: Process and Product Optimization Using Designed Experiments*. John Wiley & Sons; 2016.
- [22] MATLAB and Statistics Toolbox Release, TheMathWorks, Inc., Natick, Massachusetts, United States, 2012b.
- [23] Adey, R.A., *Modelling of Cathodic Protection Systems*. United Kingdom, WIT, 2006.
- [24] Adey, R., Peratta, C. & Baynham, J., Corrosion Data Management Using 3D Visualisation and a Digital Twin. In *NACE International Corrosion Conference Proceedings* (pp. 1–13). NACE International, 2020.
- [25] Alizadeh, R., Allen, J.K. & Mistree, F., Managing computational complexity using surrogate models: a critical review. *Research in Engineering Design*, **31**(3), pp. 275–298, 2020. <https://doi.org/10.1007/s00163-020-00336-7>
- [26] Montgomery, D.C. *Design and Analysis of Experiments*. John Wiley & Sons, 2017.