Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 257 (2025) 826–833

Procedia Computer Science

www.elsevier.com/locate/procedia

The 8th International Conference on Emerging Data and Industry (EDI40)
April 22-24, 2025, Patras, Greece

Surrogate Modeling: Review and Opportunities for Expert Knowledge Integration

Dušan Šturek[a,*], Sanja Lazarova-Molnar[a,b]

[a]*Institute of Applied Informatics and Formal Description Methods, Karlsruhe Institute of Technology, Kaiserstr. 89, 76133 Karlsruhe, Germany*
[b]*Mærsk Mc-Kinney Møller Institute, University of Southern Denmark, Campusvej 55, 5230 Odense, Denmark*

Abstract

Digital Twins (DTs), essential for digitalization and Industry 4.0, demand high-fidelity yet computationally efficient models for real-time replication of physical systems. However, simulating continuous dynamical systems, often stiff and highly granular, poses significant computational challenges. Driven by industry's need for more efficient simulations, this study explores Surrogate Modeling (SMing) as a solution to reduce computational costs while preserving accuracy, enabling real-time performance, iterative design optimization and enhancing feasibility of DT simulations. We review existing traditional and machine learning SMing approaches, analyzing their limitations, accuracy and efficiency. We, furthermore, emphasize the critical role of domain expertise in SMing workflows and explore systematic strategies for incorporating expert knowledge to improve model reliability and applicability. Finally, we identify and discuss challenges and opportunities that emerge by the fusion of expert knowledge with machine learning techniques, highlighting their potential to advance next-generation SMing techniques.

Keywords: Surrogate Modeling; High-fidelity Simulation, Knowledge Integration

1. Introduction

Numerical simulations have been extensively used in engineering to facilitate design optimization and study of physical systems. Simulating such systems is challenging because it often requires interconnecting multiple complex

* Corresponding author. Tel.: +45 71332512.
E-mail address: dusan.sturek@partner.kit.edu

10.1016/j.procs.2025.03.106

models, potentially exported as functional mock-up units (FMUs) [1], which themselves can be composed of numerous subsystems and components exhibiting nonlinear dynamics [2]. As interactions between components increase, both simulation time and fidelity get challenged to closely replicate the real system. In some cases, executing high-fidelity (HF) simulation can take even weeks per design [3]. Yet, the demand for HF simulations increased drastically in recent years [4] due to digitalization, growing adoption of model-based design and automation efforts coupled with Industry 4.0 revolution that is enabled by the significant advancements of Digital Twins (DTs) [5].

Simulation models, however, often suffer from bottlenecks, such as numerical noise, discontinuities and several local optima in the objective function [3], limiting their effectiveness as real time decision support tools in DTs [4,6–10]. In engineering practices, increased computational complexity of simulation models is driven by the iterative nature of optimal design processes [2,7,9-11] to satisfy exploration and optimization needs. HF simulations models require finer meshes and integration of intrusive physics given by high-order differential equations [9]. Recent advancements in simulation software and available computation have allowed simulation models to represent more details of physical systems. Moreover, parallel computing, parametrization and downscaling of critical variables can accelerate simulations and obtain finer resolutions. Despite this progress, conducting HF simulations still raises concerns and often results in an unaffordable task in many design processes [1,4,12–14].

Consequently, surrogate models (SMs) (also known as emulators, meta, approximation, proxy or black-box models [4,8,11]) have gained popularity across disciplines in recent years [4,15], including uncertainty quantification [16], feasibility and sensitivity analysis [15] or simulation enhancement [1]. SMs map input parameters to predict quantity of interest [11,13] by black-boxing first-principle models [6] from finite samples. Emerging in the late 20th century, SMs are constructed in a supervised manner to approximate the system's response surface and mitigate the "curse of dimensionality". Research suggests, SMs effectively support design decisions, reduce the number of required HF simulations [15,17,18], shorten the overall design cycle by accelerating parameter estimation [6] and replacing simulation solvers [4], reduce computational burden [1,7,9], are easier to optimize, hide inaccuracies from stochastic processes, and can supply gradients for approximating objective function [6].

However, existing research on SMs is very domain specific, misses a general overview and its effectiveness and success encourage us to look closer and discover ways to automate and generalize the use of SMs for further enhancement of first-principle physics-based simulation models. We conduct an extensive review on traditional and machine learning-based surrogate modeling (SMing) approaches as enablers for real time DTs, focusing on their limitations, accuracy and efficiency. Acknowledging the importance of domain expertise, we further examine methods for integrating expert knowledge (EK) in SMs. Subsequently, we used the VOSviewer tool to generate a keyword graph in Figure 1 (a) from the collected bibliography on SMs herein. The graph highlights connections between the most encountered SMing terminology, omitting the terms below ten occurrences. Our review further identifies a gap in SM-based simulation enhancement workflows that incorporate EK. We aim to bridge this gap by discussing ways to systematically integrate EK in SMing simulation workflows and presenting related challenges and opportunities.

This paper is structured as follows: We first introduce SMing process in Section 2, complemented by a thorough review of traditional and machine learning SMs. Subsequently, we recommend ways to systematically integrate expert insights in machine learning SMs in Section 3. In Section 4, we describe the opportunities and challenges associated with this integration. Finally, we draw conclusions and provide an outlook in Section 5.

2. Background and Related work

In this section, we establish the foundation for the paper, outlining the SMing process and comparing traditional and machine learning-based surrogate models used for simulation enhancement.

2.1. Surrogate Modeling Process

In engineering designs, fidelity increases as the focus shifts from preliminary observation to accurate system descriptions. Effective SMing thus prioritizes key features from HF models (synthetic) or empirical (lab and field) data [12] to trade computation for accuracy and complexity. SMing is a stepwise process shown in Figure 1 (b) that links design parameters to output responses, enabling efficient navigation in the multidimensional design space to achieve optimal designs at lower computation [11]. In the following, we elaborate on each step of the SMing process.

Strategic sampling, termed *Design of Experiments (DOEs)* [8], is crucial in SMing. It maximizes information gain from limited and expensive *High-fidelity Model Simulations* by distributing HF samples across the design space [17], minimizing redundancy and accelerating convergence of SMs towards optimal parameter values. DOEs is classified into static and adaptive methods. Static (one-shot) DOEs includes Halton and Sobol low discrepancy sequences, Latin hypercube [6] and its extension to continuous variables [19], D-optimal and fractional factorial designs [12], and quasi-Monte Carlo or Hammersley sampling [3]. Whereas classic methods, like fractional factorial designs, may place only a few samples in the design space interior, space-filling techniques, such as Latin Hypercube, typically do not put samples at the corners and edges of the design space. Consequently, static DOEs may inadequately capture corner cases, potentially reducing the prediction accuracy of SMs at and around those points, and may not effectively target critical, high-sensitivity regions or guarantee optimal computational investment for finding the global optimum.

Adaptive (online) DOEs reduces overall HF sampling requirements and addresses limitations of static methods by iteratively exploring adjacent design space regions through HF *sample infill* and stopping criteria to balance exploitation and exploration [6]. Various infill criteria include Expected Improvement [8,15], which decreases with increasing predicted values and standard deviations, Lower-Bounding Confidence [8], which uses a weighted sum of the predicted function value, Searching Surrogate Models [8], which augments sampled datasets and rebuilds SMs until prediction errors fall below threshold, and Bumpiness Function [15], which minimizes system equations. Reviews in [14,20] considered time, computation, accuracy, and relative sample correlation, showing that infill criteria reduced the confidence interval of prediction accuracy with kriging SMs [14], through selection of infill criterion remains subjective to SMs [8,12]. Moreover, initializing sample infill with static DOEs, such as Latin Hypercube maximizing minimum sample distance, improves the SM convergence towards optimal parameter values [9,13,20,21].

Surrogate Model Fitting follows the collection of HF simulation data in SMing process. Training SMs on only a small subset of HF data increases approximation uncertainty, meaning the response surfaces from SMs may deviate from the HF model responses. Note that there are no established guidelines for selecting an initial dataset size, other than relying on application domain expertise, model complexity, and design space dimension.

Lack of these guidelines highlights the need for *Surrogate Model Validation* whose costs scale poorly with higher fidelities as more data is needed to ensure fidelity and extrapolation requirements are met. With multi-dimensional design spaces, visual comparison of the model response surfaces is insufficient. Instead, relative (preferred for their context independence [9]) and absolute errors [17], sum of absolute differences, $R2$ determination coefficient, bootstrap or cross validation error, predictive variance or estimation of model fidelity [14] assess the model accuracy. Jiang [14] showed leave-one-out cross-validation error, despite higher variance, offers superior accuracy, efficiency, and generalizability [3]. Unlike $R2$, it avoids overfitting by iteratively building SMs with equal-sized HF sample bins leaving one validation bin out, unlike bootstrap error which employs random validation bin.

2.2. Surrogate Modeling Methods

Selecting appropriate SMing approach is challenging as computational costs, approximation accuracies and convergences towards optimal parameter values and of each methodology vary [12] with problem complexity and non-linearity [18]. The following SMing approaches emerged as prevalent in surrogate model-based simulation:

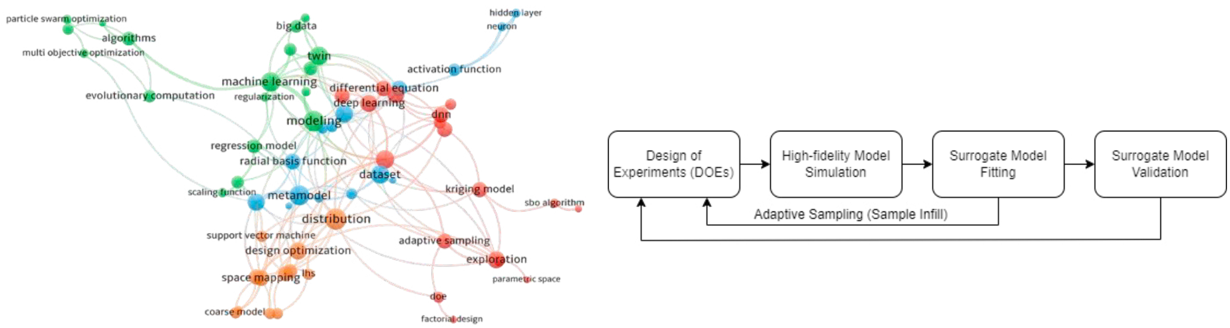

Fig. 1. (a) Keyword graph generated from literature review on surrogate-based simulation optimization; (b) Structure of surrogate modeling process

(1) interpolation-based, such as kriging, and radial basis function utilizing Euclidean distance between a sample point and the origin [8]; *(2)* polynomial-based, like chaos expansion mimicking stochastic variations of the system [3] and polynomial surface regression based on least squares [8,12,13]; *(3)* neural network-based, including support vector regression [15] and artificial neural network (NN) [18,21,24]. Cozad [11] developed ALAMO, a systematic model building method that improves and validates SMs with adaptive sampling. Despite these advancements, tailoring SMs to problem at hand remains essential and needs further investigations [15]. Several comparative analyses for SM-based simulation were presented in [3,12,14,15]. Table 1 extends these analyses, highlighting accuracy, computational cost, features and limitations of prevalent SMing approaches. Among the key observations are: *(1)* Besides polynomial surface regression [23] any approach can be employed even for higher dimensional and non-linear problems with reasonable success in preserving fidelity *(2)* Traditional SMing approaches (without machine learning) can integrate EK as statistical priors and hyperparameters. *(3)* Certain NN-based SMs also embed EK as physical laws directly into the model architecture and training process. *(4)* No single SM approach consistently excels across all scenarios. Designers, thus, pick SMing approaches based on their familiarity level and software availability.

With the increasing number of SMing approaches and lack of prior knowledge, manual selection, is tedious process. Evolutionary algorithms like Particle Swarm Optimization or Genetic Algorithms are commonly utilized to automate SM selection with multiobjective hyperparameter optimization. Despite advantages with numerous variables and high-dimensional spaces, they require significant computation due to their inherent trial-and-error nature. Implementations like COSMOS [25] are accurate and robust but lack interpretability. In contrast, AUTOSM [26] improves interpretability by pre-screening and building a decision tree from problem size, noise, sample and non-linearity to only tune the promising candidates. However, this work is restricted to lower dimensionality, radial basis function, kriging and multiple adaptive regression splines SMs. While automated SM selection streamlines the modeling process, surrogate ensembles and multi-fidelity SMs can further enhance the effectiveness and prediction performance of SMs. For more on surrogate ensembles, we point readers to the recent comprehensive review by Chen et al. [27], and on multi-fidelity SMs, to Hebert et al. [4] and references therein. Nonetheless, traditional SMs, based on smaller datasets and lacking machine learning (ML) capabilities, often fail to generalize across diverse problems [18]. ML, a subset of artificial intelligence, characterized by its ability to discover complex patterns and relationships within data, is well-suited for capturing the often-subtle and ill-posed correlations inherent in SMs. Consequently, the shift towards ML approaches represents a natural evolution in SMing. Subsequently, we present MLSMs in more detail.

Table 1. Performance comparison of prevalent surrogate modeling approaches, highlighting key features and limitations.

Surrogate Model	Accuracy / Comp. cost	Features / Limitations
Polynomial Response Surface	Low / Very Low	Good continuity, conductivity, noise suppression and practical with a priori known or polynomial functions Constrained to simple, linear, unimodal problems and non-mixed variables else instabilities may occur [8] Fidelity loss-rapid execution trade-off [3,12], overfitting in high-dimensions [13], local metamodeling suitability[23]
Polynomial Chaos Expansion	Medium / Medium	Sparse representation saves evaluation and maintains accuracy in high dimensions and polynomial degrees Tedious implementation with unknown orthogonal polynomials, still increasingly popular in last decades [3] Nataf transform, and adaptive basis strategies needed with correlated variables and finding polynomial truncations
Radial Basis Function	High / Medium	Flexible, simple [15], efficient even in high-dimensions [12], similar to neural networks High-frequency oscillations occur, thus use ridge [19] regression hyperparameter with noisy and numerous data Accuracy dependent on radial (correlation) function chosen based on desired global or local estimation
Support Vector Regression	Medium / Medium	Superior generalization, overfitting resistance and local optima avoidance [3] Unlike kriging, avoids numerical instabilities with many low-fidelity samples in multi-fidelity models Complex thus unpopular [15], outperformed by kriging and radial basis functions in complex cases [8,12] Mapping accuracy subjective to kernel function: linear; Sigmoid; Gaussian; (non)homogeneous polynomial
Artificial Neural Network	Very High / Low	Fits numerous parameters in nonlinear systems with cheap model evaluation but expensive training [10] Several architectures and toolboxes available for direct expert knowledge integration, enabling coarser datasets [10] Difficult design due to infinite architectures using large, pre-processed datasets driving convergence [7,10,17,24]
Gaussian Process (Kriging)	Very High / Medium	Less regression parameters and stable, noise-resisting prediction also with deterministic simulations [3] Uncertainty quantification besides prediction [8,13,15] with Ordinary, Static, Gradient-Enhanced options Prevents overfitting with regression term in covariance matrix (nugget effect), that may include prior knowledge [15] Computation scales poorly with problem size [15,19], and inherent smoothness hinders stiff systems

2.3. Machine Learning Surrogate Models

Without ML pattern recognition capabilities, SMs fail to uncover corner cases and identify complex, often hidden or poorly understood, data patterns [7,23], potentially introducing bias due to data paucity [18]. Unlike traditional surrogates, MLSMs require large training datasets to approximate HF responses [10]. ML's inherent non-linearity through activation functions popularizes it in SMing of large-scale problems [13,22]. Moreover, MLSMs are non-intrusive [5,24], protecting intellectual property and enabling simulation democratization by sharing the models with stakeholders. Pre-trained and hyperparameter optimized MLSMs offer a key advantage: forward evaluation of weights and biases takes only a fraction of the initial training time to obtain responses for new inputs [3,9,17]. Numerous studies demonstrate the success of MLSMs in reducing simulation runtime. Sun [9] developed several MLSMs and presented methods for selecting NN architecture. Khan and Green [24] introduced a methodology using TensorFlow for constructing gravitational wave astronomy MLSMs. Balchanos [2] implemented Jordan recurrent NNs for a thermometric power system and highlighted the criticality of time step selection on the fidelity of these MLSMs.

However, ML imposes its own challenges in SMing such as architecture selection and tuning of NNs to achieve favorable performance with simplest architecture. Careful consideration of training data size, diversity, and potential biases is substantial to ensure model stability, mitigate numerical issues and spread the high initial training cost across use cases. Data preprocessing such as min-max and standardization (zero mean, unit variance) scaling is a crucial with MLSMs to preserve data's physical meaning, enhance prediction accuracy, reduce dataset size, and prevent large amplitude differences and exploding gradients during training [9]. The problems with pure MLSMs are: *(1)* models are complete black-box which increases prediction uncertainty; *(2)* extensive, in industry and HF simulations often prohibited, data is required for reliable generalization; and *(3)* EK governing SMs cannot be incorporated, resulting in simulation models of dynamical systems with poorer performance and training convergence. Recognizing the critical role of EK in SMing and DTs, which has been shown to reduce data requirements and improve robustness, interpretability and generalization of the models [28-30], we elaborate on systematic integration of EK in MLSMs in the next section.

3. Expert Knowledge Integration in Surrogate Models

Across the scientific community, there is a clear trend toward hybrid approaches that combine knowledge and data-driven techniques to leverage their respective strengths [18,29,30], referred to as informed ML. From ML perspective, EK is understood as governing laws, theories or any validated information about a physical system with the potential to enhance corresponding models' fidelity and interpretability. However, EK takes many forms and representations as e.g., logic rules, algebraic/differential equations, probabilistic dependencies, graphs, engineering intuition, etc. [18,29]. In SM-based simulation of continuous dynamical systems, these forms typically reduce to ordinary/partial differential equations, expert insights and years of experience with simulation solvers. Current research emphasis on DTs and domain-specific MLSMs underscores the need for more generalizable SMing that systematically fuses MLSMs with EK. To incorporate EK from continuous dynamical systems, biases must be introduced in the ML model to guide it towards physically consistent solutions. Our review shows that "physics-guided", "physics-informed" or "physics-aware" ML can embed EK in SMs using physical laws from model-based design, unlike mainstream ML applications using prior knowledge in feature engineering or post-processing [29]. Compared to pure ML, physics-informed ML enhances prediction stability, reliability, explainability, and sample efficiency. While physics-guided ML may trade broader generalization to other problems for accuracy [6,18,28], recent deep learning advancements have improved extrapolation to unseen data and efficiency of gradient computations through automatic differentiation. We identified the following three key ways to fuse EK in the SMing workflow:

1. Training data: the most straightforward way to embed EK in MLSMs is to augment the synthetic data from numerical simulations with empirical datasets. Numerical simulations are deterministic and the lack of stochastic variability in empirical datasets can cause DTs to deviate from actual system behavior [5]. This "reality gap" challenges the prediction reliability and robustness of MLSMs underlying in DTs in industry. The time-consuming augmentation of training datasets with empirical data enhances fidelity of MLSMs even before integrating explicit EK thanks to observational biases that can address hard-to-simulate corner cases and missing data areas, which may be impossible or expensive to model [18]. For studies augmenting simulation with empirical datasets, we point the readers to [18,29,30] and references therein. Moreover, the design parameter range used for generating synthetic data provides valuable feedback. SMs can alert designers about potential design space regions where the response surfaces approximated by SMs might deviate from HF model responses if SMs encounter parameter values outside this range.

2. Loss function: During training, MLSMs minimize loss functions that embed learning biases by quantifying discrepancies between real outputs and model predictions. These loss functions typically include supervised loss and regularization terms based on model complexity. To comply with physical laws, facilitate learning from unlabeled data, improve generalization and reduce search space, additional loss terms can be derived and incorporated from differential equations [10]. These physics-loss terms are introduced with hyperparameters controlling their overall loss contribution. Physics-informed neural networks (PINNs) exemplify this approach by integrating governing physics equations directly into the training, ensuring adherence to physical principles. Several successful attempts with PINN SMs have been conducted in highly nonlinear fluid flow systems. For instance, Haghighat et al. [7] utilized PINNs with transfer learning to reduce training cost and initial losses, achieving effective, physics-compliant SMs with improved convergence for sensitivity analysis even with unseen data. In another study [9], fast approximations obtained from PINN SMs closely followed HF simulations relying on partial differential equations. Niaki et al. [28] considered PINNs as natural SMs that require coarser datasets, further highlighting their utility and effectiveness.

3. Model architecture: underlying NNs architectures drive fidelity and explainability of trained MLSMs. Incorporating inductive biases by encoding physics into NNs design, followed by hyper-parameter tuning, is an active research area [18,29,30] where certain neurons can: *(1)* encode physical meaning by introducing fixed parameters and intermediate values e.g. conservation of energy in Hamiltonian NNs; *(2)* have fixed weights determined by expert inputs; or *(3)* have solution space restricted through layer designs or connectivity inspired by symmetries and invariances in dynamical systems. Following these approaches, one can encode known a priori features such as conservation laws and symmetries from the differential equation solutions and force NNs to respect corresponding initial or boundary conditions [30]. Similarly, simulation solvers can be combined with NNs to achieve continuous-time depth models with neural ordinary differential equations (NODEs) [21]. Moreover, Rackauskas et al. [1] integrated Continuous-Time Echo State NNs in JuliaSim for automated data-driven training. This tackled instabilities in PINNs and NODEs with stiff models (experiencing high variations in short time) and validated the developed SMs in co-simulations by replacing Air Conditioner FMUs to obtain 340-fold speedup with less than 4% error in dynamics.

To our best knowledge, the work by Rackauskas et al. [1] is the only approach automating ML SMing, based on Modelica HF models, still requiring further testing, FMU export and more explicit EK fusion. Our literature review thus reveals a notable gap in surrogate-based simulation workflows that utilize EK. The gap is especially pronounced in the real time industrial DTs context. In the next section, we discuss the observed relevant challenges and opportunities that need to be considered when integrating EK in machine learning SM-based simulation workflows.

4. Challenges and Opportunities

Our review highlighted potential challenges for systematic integration of expert knowledge in ML and SM-based simulation that we describe in the following. Subsequently, we highlight opportunities evolving from this integration.

4.1. Challenges in Fusing Expert Knowledge with Machine Learning and SM-based Simulation

We identified three key challenges for the systematic fusion of EK in MLSMs that we expect to have a significant impact on the automation of EK integration in SM-based simulation workflow:

1. Strong correlation between expert knowledge and simulation models: In stiff continuous dynamic systems, even the smallest changes in input often result in abrupt output variations. This can cause numerical difficulties and discontinuities in model responses, potentially breaking system symmetries and invariances which may lead to wrong predictions of MLSMs that encode EK inside their architecture relying on those symmetries. Additionally, embedding learning biases through the loss function may introduce training instability due to competition between the individual loss terms [28,30], thus preventing guaranteed convergence to the global minimum. This necessitates adaptive hyperparameter tuning algorithms that balance data fitting and adhering to physical laws [28]. Consolidating these aspects, fusing EK in SMing results in highly tailored approaches where even slightly different systems may necessitate their own ways of integrating EK, which strongly challenges its automation.

2. Limitations of physics-informed surrogate modeling: Although physics-guided ML is continuously and rapidly evolving, it still faces significant challenges, particularly with multiscale and multiphysics problems. Physics-aware MLSMs may experience varying convergence rates or fail to train effectively due to steep gradients arising from high-frequency components in the solutions [30]. These issues underscore the challenge of integrating EK as it requires developing more robust architectures and training algorithms, such as adaptive activation functions [10] and meta-learning, to enhance performance and generalizability across different problems.

3. Addressing multidisciplinary expertise requirements: Integrating EK into MLSMs requires expertise in multiple domains such as ML, simulations, and physics. This multidisciplinary approach creates a steep learning curve for designers, making it challenging to develop effective MLSMs. Potential errors in implementation and longer development times can arise due to the complexity. To address this, designers must be interactively guided through the SMing process and presented with design options to reduce the impact of this need. We aim to bridge this gap with automated implementation of SMing workflow verified on frequency drives, devices that control the speed and torque of an electric motor by varying the frequency and voltage of its power supply.

4.2. Opportunities from Fusing Expert Knowledge with Machine Learning and SM-based Simulation

Next, we derive and discuss the key opportunities arising from the fusion of EK in SM-based simulation workflow. These opportunities focus on different fidelity and interpretability enhancements of MLSMs:

1. Improving surrogate model efficiency: Fusion of expert knowledge reduces the data requirements needed to alleviate the computational burden [18], thus, increasing design space exploration with more interpretable models even further. Substantially more simulations can run with the same resources, empowering engineers to explore a wider range of design parameters, leading to finding superior product designs at reduced costs.

2. Automating synthetic data generation: To effectively automate integration of EK in SMing workflow, we see potential in exploring NNs scaling laws between synthetic dataset size, number of MLSM training parameters and available compute power. For example, size of EK-driven datasets could be set relatively to the amount of synthetic data, simulation model stiffness ratio and design space properties like space-fillingness and dimension.

3. Enhancing robustness and deployment of surrogate models: Despite their potential, SMs face barriers to widespread adoption due to robustness concerns, error metrics not translating well when integrated with HF simulations [18] and lack of automated workflows [1]. Physics-informed learning uses automatic differentiation (eliminating the need for mesh generation) and can systematically integrate EK in MLSMs. This approach can thus be an enabler for DTs in Industry 4.0 applications where real-time, robust, and interpretable simulation models that adhere to physical constraints are necessary.

4. Supporting continuous update of Digital Twins: Random weight initialization is common ML practice but can lead to physically inconsistent initial states or models stuck in local minima early in training [29]. Transfer learning can mitigate this by pre-training more general MLSMs on extensive HF data and then fine-tuning them for specific tasks whose application is related to the base model [7]. One could, for example, use EK to guide the selection of base models for transfer learning, e.g., by evaluating candidate base models depending on their performance on benchmark tasks that are defined by experts to mimic the target system's behavior. This would shift the long initial training to early stages of SMing and speeding up the later creation of more application focused MLSMs. This can benefit continuous update of DTs deployed in industrial applications where many physical systems evolve over time due to aging, changing conditions, or external events. DTs must adapt to these changes with models trained in reasonable time to maintain their accuracy and reliability.

5. Summary and Outlook

We reviewed existing literature on surrogate modeling approaches, considering accuracy, computation, features and limitations. We further analyzed advantages and drawbacks of purely machine-learning-based surrogate models (MLSMs), highlighting the lack of integrated expert knowledge (EK). Next, we provided a high-level review on how the different types of EK can be integrated with MLSMs. Subsequently, we derived and discussed challenges and opportunities arising from this integration. With this paper, we aim to establish a foundation for systematic integration of expert-augmented surrogate-based simulation modeling workflow for stiff continuous dynamical systems. Our future work will focus on developing a conceptional framework, implementing case studies in the frequency drives domain, addressing identified challenges and investigating further opportunities to enhance surrogate modeling effectiveness.

Acknowledgements

This work is partly funded by the Innovation Fund Denmark (IFD) under File No. 3129-00014B. We thank Michael Laursen for the insightful discussions concerning variable frequency drives simulations.

References

[1] Rackauckas C, Gwozdz M, Jain A, Ma Y, Martinuzzi F, Rajput U, Saba E, Shah V. B, Anantharaman R, Edelman A, Gowda S, Pal A, Laughman C. Composing Modeling And Simulation With Machine Learning In Julia. 2022 Annual Modeling and Simulation Conference (ANNSIM), San Diego, CA, USA: IEEE; 2022, p. 1–17.
[2] Balchanos M, Moon K, Weston N. R., Mavris D. N. A Dynamic Surrogate Model Technique for Power Systems Modeling and Simulation, 2008, p. 2008-01–2887.
[3] Koziel S, Pietrenko-Dabrowska A. Fundamentals of Data-Driven Surrogate Modeling. Surrogate Modeling for High-Frequency Design, WORLD SCIENTIFIC (EUROPE); 2022, p. 1–37.
[4] Hebert J. L., Holzer T. H, Eveleigh T. J, Sarkani S. Use of Multifidelity and Surrogate Models in the Design and Development of Physics-Based Systems. Systems Engineering 19 (2016) 375–391.
[5] Rasheed A, San O, Kvamsdal T. Digital Twin: Values, Challenges and Enablers From a Modeling Perspective. IEEE Access 8 (2020) p. 21980–22012.
[6] Bárkányi Á, Chován T, Németh S, Abonyi J. Modelling for Digital Twins-Potential Role of Surrogate Models. Processes 9 (3) (2021) 476.
[7] Haghighat E, Raissi M, Moure A, Gomez H, Juanes R. A physics-informed deep learning framework for inversion and surrogate modeling in solid mechanics. Computer Methods in Applied Mechanics and Engineering 379 (2021) 113741.
[8] Han Z-H, Zhang K-S. Surrogate-Based Optimization. In: Roeva O, editor. Real-World Applications of Genetic Algorithms, InTech; 2012.
[9] Sun G, Wang S. A review of the artificial neural network surrogate modeling in aerodynamic design. Proceedings of the Institution of Mechanical Engineers, Part G: Journal of Aerospace Engineering 233 (2019) 095441001986448.
[10] Sun L, Gao H, Pan S, Wang J-X. Surrogate Modeling for Fluid Flows Based on Physics-Constrained Deep Learning Without Simulation Data. Computer Methods in Applied Mechanics and Engineering 361 (2020) 112732.
[11] Cozad A, Sahinidis N. V, Miller D. C. Learning surrogate models for simulation-based optimization. AIChE Journal 60 (2014) 2211–2227.
[12] Alizadeh R, Allen J. K, Mistree F. Managing computational complexity using surrogate models: a critical review. Research in Engineering Design 31 (2020) 275–298.
[13] Chugh T, Sun C, Wang H, Jin Y. Surrogate-Assisted Evolutionary Optimization of Large Problems. In: Bartz-Beielstein T, Filipič B, Korošec P, Talbi E-G, editors. High-Performance Simulation-Based Optimization, vol. 833, Cham: Springer International Publishing; 2020, p. 165–187.
[14] Jiang P, Zhou Q, Shao X. Surrogate Model-Based Engineering Design and Optimization. Singapore: Springer Singapore; 2020.
[15] Bhosekar A, Ierapetritou M. Advances in surrogate based modeling, feasibility analysis, and optimization: A review. Computers & Chemical Engineering 108 (2018) 250–267.
[16] Sudret B, Marelli S, Wiart J. Surrogate models for uncertainty quantification: An overview. 2017 11th European Conference on Antennas and Propagation (EUCAP), Paris, France: IEEE; 2017, p. 793–797.
[17] Jatnieks J, De Lucia M, Dransch D, Sips M. Data-driven Surrogate Model Approach for Improving the Performance of Reactive Transport Simulations. Energy Procedia 97 (2016) 447–453.
[18] Willard J, Jia X, Xu S, Steinbach M, Kumar V. Integrating Scientific Knowledge with Machine Learning for Engineering and Environmental Systems. ACM Computing Surveys 55 (2022) 66.
[19] Urquhart M, Ljungskog E, Sebben S. Surrogate-based optimisation using adaptively scaled radial basis functions. Applied Soft Computing 88 (2020) 106050.
[20] Fuhg J. N, Fau A, Nackenhorst U. State-of-the-Art and Comparative Review of Adaptive Sampling Methods for Kriging. Archivs of Computational Methods in Engineering 28 (2021) 2689–2747.
[21] Legaard C. M, Schranz T, Schweiger G, Drgoňa J, Falay B, Gomes C, Iosifidis A, Abkar M, Larsen P. G. Constructing Neural Network-Based Models for Simulating Dynamical Systems 2022.
[22] Maulik R, Fukami K, Ramachandra N, Fukagata K, Taira K. Probabilistic neural networks for fluid flow surrogate modeling and data recovery. Physical Review Fluids 5 (2020) 104401.
[23] Do Amaral J.V.S, De Carvalho Miranda R, Montevechi J.A. B, Dos Santos C. H, Gabriel G.T. Metamodeling-based simulation optimization in manufacturing problems: a comparative study. International Journal of Advanced Manufacturing Technology 120 (2022) 5205–5224.
[24] Khan S, Green R. Gravitational-wave surrogate models powered by artificial neural networks. Physical Review D 103 (2021).
[25] Mehmani A, Chowdhury S, Meinrenken C, Messac A. Concurrent surrogate model selection (COSMOS): optimizing model type, kernel function, and hyper-parameters. Structural and Multidisciplinary Optimization 57 (2018) 1093–1114.
[26] Jia L, Alizadeh R, Hao J, Wang G, Allen J. K, Mistree F. A rule-based method for automated surrogate model selection. Advanced Engineering Informatics 45 (2020) 101123.
[27] Chen H, Zhang Z, Li W, Liu Q, Sun K, Fan D, Cui W. Ensemble of surrogates in black-box-type engineering optimization: Recent advances and applications. Expert Systems with Applications 248 (2024) 123427.
[28] Amini Niaki S, Haghighat E, Campbell T, Poursartip A, Vaziri R. Physics-informed neural network for modelling the thermochemical curing process of composite-tool systems during manufacture. Computer Methods in Applied Mechanics and Engineering 384 (2021).
[29] Rueden L von, Mayer S, Beckh K, Georgiev B, Giesselbach S, Heese R, Kirsch B, Pfrommer J, Pick A, Ramamurthy R, Walczak M, Garcke J, Bauckhage C, Schuecker J. Informed Machine Learning - A Taxonomy and Survey of Integrating Knowledge into Learning Systems. IEEE Transactions on Knowledge and Data Engineering 35 (2021) 614-633.
[30] Karniadakis G. E, Kevrekidis I, Lu L, Perdikaris P, Wang S, Yang L. Physics-informed machine learning. Nat Rev Phys 3 (2021) 422–440.