Proceedings of the 2025 Winter Simulation Conference
E. Azar, A. Djanatliev, A. Harper, C. Kogler, V. Ramamohan, A. Anagnostou, and S. J. E. Taylor, eds.

SIMULATION OPTIMIZATION 2050 AND BEYOND

Soumyadip Ghosh[1], Peter J. Haas[2], L. Jeff Hong[3], Jonathan Ozik[4], and Benjamin Thengvall[5]

[1]IBM TJ Watson Research Center, Yorktown Heights, NY, USA
[2]Manning College of Info. & Computer Sci., University of Massachusetts Amherst, Amherst, MA, USA
[3]Department of Industrial and Systems Engineering, University of Minnesota, Minneapolis, MN, USA
[4]Decision and Infrastructure Sciences Division, Argonne National Laboratory, Lemont, IL, USA
[5]OptTek Systems, Inc., Boulder, CO, USA

ABSTRACT

The goal of this panel was to envision the future of simulation optimization research and practice over the next 25 years. The panel was composed of five simulation researchers from academia, industry, and research laboratories who shared their perspectives on the challenges and opportunities facing the field in light of contemporary advances in artificial intelligence, machine learning and computing hardware. The panelists also discussed the role simulation optimization can, should, and will play in supporting future decision-making under uncertainty. This paper serves as a collection of the panelists' prepared statements.

1 INTRODUCTION

The field of simulation optimization (SO) concerns the study of how to make optimal decisions with the aid of stochastic simulation models (Fu 2015). SO methodology combines ideas such as mathematical programming, experiment design, metamodeling, and adaptive sampling to identify promising solutions to complex problems posed under uncertainty. Introductory and advanced tutorials on SO have appeared in past proceedings of the Winter Simulation Conference (WSC), including most recently Shashaani (2024). Other past WSC panels have sought to bridge SO research and practice (Fu et al. 2000; Fu et al. 2014). As for experimental research, Pasupathy and Henderson (2011) introduced a testbed of SO problems and solvers, named SimOpt, that has undergone several significant redesigns (Eckman et al. 2024).

The Simulation Optimization track chairs—David Eckman, Siyang Gao, and Yuwei Zhou—organized a panel consisting of five simulation researchers to imagine what the future of simulation optimization might look like in 2050 and beyond. The panelists—Soumyadip Ghosh, Peter Haas, Jeff Hong, Jonathan Ozik, and Benjamin Thengvall—brought complementary expertise on how SO methods are designed from theory, implemented in software, and deployed in practice. Motivated by the rapid and timely advances being made in artificial intelligence (AI), machine learning (ML), and computing capabilities, the panelists were asked to prognosticate about how these and other trends might shape the evolution of SO and suggest how the SO community might best position itself for what lies ahead.

The remainder of this paper consists of a list of prepared questions and responses that framed the panel's topics of discussion.

2 PROMPTS

The following list of questions was shared with the panelists prior to the conference to prompt their thinking about the future of the field.

1. How have recent advances in AI/ML changed how your organization undertakes the optimization of simulation models or how you go about research in SO?

2. Do you believe that emerging AI/ML techniques such as transformers or physics-informed neural networks have the potential to advance SO research? If so, how?
3. The ubiquity of parallel computing resources has been changing how SO algorithms are designed and redefined what constitutes the state of the art. Do you foresee more recent (or near-future) advances in computing hardware (e.g., GPUs and quantum computing) doing the same?
4. Digital twins have brought greater interest in the use of simulation optimization for near-real-time control. Do you see this shift to more online monitoring and decision making opening up new classes of SO algorithms?
5. What are some things that you believe are limiting, or could ultimately limit, advances in SO research over the next 25 years? E.g., a lack of publicly available industry-scale test problems, a decline in the number of PhD students with interest in the field.
6. Is there anything you worry about as an existential threat to the field of SO?
7. In what aspects do you see the boundaries of SO being pushed in the next 25 years? E.g., the dimension of the problems, the time-scale of decision-making, the classes of problems that fall within the realm of SO.
8. Are there any adjacent research communities that you think could benefit the most from SO researchers having more of a presence in? Or that the SO community should be more aware of for finding useful ideas to borrow?
9. What do you anticipate being the difference between the state of simulation optimization in 2025 versus 2050? Controversial or contrarian opinions welcome.
10. In addition to the classical applications of SO in areas like queueing networks, manufacturing systems, or supply chain logistics, what new application fields do you see as particularly promising for SO?

3 PANELIST STATEMENTS

The panelists all prepared written statements that appear below. Three of the panelists addressed the questions point-by-point, whereas the other two provided more general statements.

SOUMYADIP GHOSH

[**Q1 and Q2**] Simulation modeling (SM) and optimization (SO) are one central approach of the broader goal in probability and statistics of being able to explain and control phenomena that are observed via a (typically finite) set of sample data. Traditionally, a simulation model is built with knowledge of the complex system dynamics involved, where SO techniques are needed because it is hard to analytically derive optimization and control policies. AI/ML models seek to explain black-box phenomena where it is not possible to know or observe the dynamics, e.g., how do we recognize a printed digit as a '0' vs '1'? Broadly, these new paradigms in neural network architectures expand the space of model classes that can be used to understand characteristics, particularly spatial/temporal dynamics, of opaque processes that generate data of interest. Thus, they also expand the ability of the SM practitioner to build better models when the underlying dynamics are not observable.

My organization has enthusiastically embraced AI/ML modeling for complex phenomena where previous approaches like dynamic systems models were expensive yet not as effective. For example, the time-series paradigm of Transformer models have been used to understand the goings-on inside (1) blast furnaces and electric arc furnaces to predict what grade of metal is produced by any given set of inputs; (2) weather models to improve predictions at the local level.

AI/ML optimization algorithms have also provided important advances that will be worthwhile to study in the SO context. Model training (fitting and validation in SM) is still largely done via first-order (gradient) methods. Techniques of interest to SO have gained widespread practical adoption, such as gradient and iterate averaging, and variance reduction methods that compensate for small sample estimates of gradients have also emerged.

The primary paradigm for dynamic control in AI/ML is reinforcement learning (RL), where one simultaneously learns a representation of the value of various input (control) settings while deducing control policies that drive the model towards optimal performance. Ranking and selection is an early formulation from SO in this vein. The SO area of optimal experimental design more broadly seeks to "optimize" a black-box process that is expensive to query. The SO community can (and already does) participate in providing our unique perspective on improving the sample complexity of RL both in theoretical and methodological advances.

[**Q3**] GPU architecture design advances are tied to their applications. AI GPUs target specific numerical operations like matrix-vector multiplications and vector averaging that are amenable to massive block parallelization. These operations are central to computing sample average gradients of massive NN models in applying first-order stochastic optimization. In turn, they may have spurred advances in such methods, particularly making (gradient/iterate) averaging important in practice. Their role in turning other classes of stochastic optimization, e.g., approximate second-order methods like LBFGS, into methods of practical use have not been compelling. I see GPUs as enabling fast computation operations on complex simulation models like stochastic dynamic systems of physical phenomena. The task of orchestrating banks of GPUs and CPUs in a high-performance computing system to improve simulation optimization techniques will continue to be a fertile open area of research.

The basic operand in a quantum computer is a finite-support distribution rather than a real number. Quantum circuits manipulate joint distributions of operands, and we can observe the end state only via samples from the joint distribution. Seen this way, quantum computers look like ideal devices for simulation modeling and optimization. The current state of quantum computer implementations have quite severe limitations: (a) there are no guarantees that general quantum operations are applied correctly, limiting devices to operations over a small number of operands, and (b) even loading a nontrivial classical (continuous or large finite support) distribution onto a quantum device is not efficient. Nevertheless, our collective expertise in stochastic modeling, input and output analysis as well as sample complexity aware SO will be very relevant to the quantum future.

[**Q4**] While I have not paid close attention to "digital twins," online control of systems via techniques like sequential decision processes, reinforcement learning, etc. are indeed becoming more prevalent. Please refer to my answer to Q1&2 above.

[**Q5, Q6 and Q7**] Modeling and validation of stochastic simulation models forms the backbone of the INFORMS simulation community, and our smaller SO sub field has traditionally operated on top of this competency. The SO stream of literature also falls under the broader decision making under uncertainty literature, which will continue to remain a central concern, and so the SO researcher/practitioner will always find relevance in this broader community.

Will traditional simulation modeling remain relevant in the future? This will depend on whether the dominant scientific obsession with end-to-end automation with AI/ML models pans out in terms of providing stable modeling and decision making for diverse applications. They do tend to work well in systems where a majority of control factors and constraints are well known (e.g., games such as Chess, Go, etc.) and so it is reasonable to expect that they will do well in modeling systems where all the factors that are amenable to control, for example in automation of manufacturing systems, are clear and well understood. Consequently, we may cede significant space in modeling factory and supply chain operations. There are, however, interesting algorithmic questions that arise, in for instance, automation in such traditional "hard" engineering applications.

I remain skeptical that AI/ML models will be equally effective in understanding phenomena where all the significant control factors are not obvious or are poorly modeled with strict rules, for instance service systems broadly defining this to be where human behavior is central to the model. I expect the SM and SO expert will continue to find relevance in addressing such hard problems. The field of quantum computing will also be a new fertile field of research once error corrected hardware devices become a reality.

[**Q8**] Most SO researchers are already well versed in and are contributing members to the statistical learning community. This will continue to be of relevance in the near future, and understanding the new paradigms of stochastic decision making such as RL will be key. In the longer run, if error correcting quantum systems become a reality, our unique perspective on computational effort (e.g., sampling) sensitive optimization algorithms will make quantum computing based decision making an important future area.

[**Q9 and Q10**] We may not be building and validating models for systems that are constructed from well-defined rules, even if these are stochastic in nature, e.g., physical processes such as manufacturing. I expect we will still be working on the optimization of such “automatically” built models, especially in gaining key insights from the control policies that methods such as reinforcement learning are able to construct. Despite all the positive press generated in popular media, it is not clear that RL methods are able to avoid overfitting to the massive data available and generate true understanding of the processes that are being studied.

We will still be modeling and optimizing human-centric service applications be it in queueing or complex supply chain systems. By 2050, we will also be have taken the first steps towards adopting quantum computing and would have contributed to the practice of quantum simulation modeling and optimization.

PETER HAAS

I believe that SO will increasingly be viewed as one particular tool within an analytics ecosystem of models and data, coexisting with machine learning, supported by data management technology, and accessed by an ever larger community of data scientists and engineers (Haas and Theodoropoulos 2020). This creates exciting opportunities and challenges for SO research.

The ecosystem provides opportunities for synergies between analytical techniques such as simulation and AI/ML. E.g., simulations are already used to train AI/ML models that can be used for online decision making and reinforcement learning. Conversely, AI/ML methods from machine vision and elsewhere have been used to extract and synthesize the input data needed for SO, e.g., by extracting semantically meaningful event timelines from video frames or conceptualizing social systems (eliciting key aspects, relations, and hypotheses from stakeholders) prior to mathematical formalization (Davis et al. 2022). In my own research I have seen the potential of neural networks to facilitate SO by (i) simplifying simulation input modeling and (ii) facilitating simulation metamodeling to combine the speed of neural network inference using GPUs with the deep domain expertise embodied in simulation models, facilitating creation of digital twins and hence accelerating SO for real-time and near-real-time applications. Future developments, e.g., in transformer models having reduced data requirements, will potentially allow for increasingly effective synergies between analytical techniques. One hazard, however, is that the current heavy allocation of resources to AI/ML research may cause the benefits of simulations based on domain expertise to be overlooked, and hence underfunded.

New applications of SO to highly complex, data-intensive problems in dynamic environments open up many opportunities and challenges, and call for collaboration with the data management community. Applications around internet-scale activities such as recommendation systems and digital marketing lead to data-driven SO problems of unprecedented scale and speed. Moreover, even just within the ecosystem of simulation models, high-level decision-makers dealing with complex issues such as population health and safety have increasingly felt the need to bring together multiple simulation models across a range of disciplines to analyze complex systems-of-systems, raising the issue of SO in composite simulation models. Some examples of challenging problems, encountered in my research and elsewhere, include the following:

- "Stochastic package queries" in probabilistic database systems—which are basically stochastic integer linear programs (ILPs) with simulation-driven inputs—attempt to push SO close to where the data often resides: in a data management system (DMS). Such technologies aim to open up SO to a more general community of data scientists, while inheriting DMS functionality such as access control and data-consistency maintenance (Meliou et al. 2025). However, this setup leads to SO problems involving tens of millions of decision variables, requiring novel approaches.
- Use of graph neural networks for simulation metamodeling leads to an important but understudied class of hybrid discrete-continuous SO problems that aim to jointly optimize both the system structure (e.g., number of queues and the links between them) and continuous operating parameters (e.g., services rates and routing probabilities). Current efforts have involved heuristic methods based on modified Monte Carlo tree search (Cen and Haas 2023) and more principled methods involving MILP solvers, but there is much scope for future work.
- The decentralized nature of global data, combined with evolving privacy-driven restrictions on data sharing, together imply that SO methods may need to move beyond parallel optimization to distributed optimization, in which (possibly sensitive) data is locally processed on a node and the results of such local processing are shared between distributed nodes. These issues have received attention with respect to descriptive analytics, i.e., database query processing, but much less so with respect to prescriptive analytics such as SO.
- Noise is often deliberately added to data to protect privacy, e.g., using the "Laplace mechanism" to ensure differential privacy. Such input uncertainty, as well as increased uncertainty when using ML methods to generate input data, make robust-optimization techniques essential.
- Maintaining an optimal solution as the underlying data is changing, without recomputing the solution from scratch, is an increasingly important problem as data is acquired at increasingly rapid rates from sensors and other sources.
- Work at IBM (Haas et al. 2012) has demonstrated the potential of data-integration techniques to facilitate the creation of composite simulation models for complex systems-of-systems, which bring unique challenges to the practical application of SO methods.
- Users are demanding more from optimization tools. Optimality criteria may include trying notions of solution diversity and fairness, which can be nontrivial to incorporate into a formal optimization specification. To promote user trust of optimization results, there is increasing demand for augmenting the result with high-level explanations, e.g., of which data characteristics or model parameters most influenced the result. As mentioned above, data-privacy requirements add additional complication.

Looking toward 2050, it is interesting to speculate about the degree to which AI systems will eventually be able to autonomously generate simulations and use them to solve SO problems without direct human intervention, thereby automating what-if analysis. Although I am skeptical of artificial general intelligence, it may be possible to make progress in this direction for specialized AI models tuned for specific applications. In such a world, the techniques for robust, explainable SO will be more important than ever.

JEFF HONG

Do you believe that emerging AI/ML techniques such as transformers or physics-informed neural networks have the potential to advance SO research? If so, how?

Yes! In a fundamental sense, generative models are a form of simulation—or simulation can be seen as a kind of generative modeling. Current generative models are more data-driven, whereas simulation models are typically physics- or process-driven. For problems with limited data, I believe these two approaches will inevitably converge. Physics-informed generative models are an early example of this fusion. As simulation models evolve, SO algorithms must also adapt, since new model structures bring new opportunities for optimization methods to exploit.

The ubiquity of parallel computing resources has been changing how SO algorithms are designed and redefined what constitutes the state of the art. Do you foresee more recent (or near-future) advances in computing hardware (e.g., GPUs and quantum computing) doing the same?

Yes! Simulation is inherently computationally intensive, and it would be absurd not to take advantage of the most advanced computing capabilities available. We have already seen some early developments in this area, and I believe much more will come. Advances in hardware, whether through GPUs or quantum computing, will continue to reshape both the types of simulation models we can build and the optimization algorithms we can design to work with them.

Digital twins have brought greater interest in the use of simulation optimization for near-real-time control. Do you see this shift to more online monitoring and decision making opening up new classes of SO algorithms?

Yes! Online simulation optimization will become a major research direction driven by the rise of digital twins, but it will differ from classical online optimization in important ways. In simulation settings, we typically have substantial knowledge about the underlying model, reducing the need for the continual model learning that characterizes traditional online optimization. Moreover, despite continued advances in hardware, achieving true real-time optimization for complex simulations will likely remain infeasible. As a result, contextual SO—or SO with covariates—will likely dominate, with offline-simulation-for-online-application (OSOA) framework becoming the standard approach.

What are some things that you believe are limiting, or could ultimately limit, advances in SO research over the next 25 years? E.g., a lack of publicly available industry-scale test problems, a decline in the number of PhD students with interest in the field.

I agree that a lack of publicly available industry-scale test problems and declining student interest are real concerns, but I see them as symptoms of a deeper issue: we are not solving the problems that industry truly needs solved. Over the past 25 years (2001–2025), advances in SO have been relatively modest, with much of the work focused on refining algorithmic frameworks that were already established by 2000. The growth of AI and ML will bring new SO problems and demand new methods, and the real question is whether the simulation optimization community will embrace these opportunities or be left behind.

Is there anything you worry about as an existential threat to the field of SO?

I am quite optimistic about the future of SO as a discipline, but I am less certain that the current simulation community will drive the next major wave of innovations—or that future developments will even be described as “simulation optimization” in the way we understand it today.

In what aspects do you see the boundaries of SO being pushed in the next 25 years? E.g., the dimension of the problems, the time-scale of decision-making, the classes of problems that fall within the realm of SO.

The simulation community has largely treated SO as a black-box optimization problem, but this viewpoint faces hard limits due to the curse of dimensionality: when treating everything as a black box, scaling up to high-dimensional and large-scale problems becomes extremely difficult, as we have seen over the past 25 years. The real opportunity lies in application-driven problems, where domain-specific structures can be exploited to bypass these limitations and where advances could have meaningful real-world impact.

What do you anticipate being the difference between the state of simulation optimization in 2025 versus 2050? Controversial or contrarian opinions welcome.

Given the largely incremental changes from 2000 to 2025, it is difficult to make bold predictions for the next 25 years. Nevertheless, I will attempt to do so—and remain hopeful. First, discrete-event simulation could lose its popularity if it fails to adapt to massively parallel environments such as GPUs. AI-powered agent-based models (ABMs) could rise in popularity, even though I personally have not worked on ABMs. Similarly, physics-informed generative models are likely to become popular—although their interaction with classical queueing frameworks remains an open question. Second, the boundary between simulation optimization and stochastic optimization will blur. Monte Carlo methods, which go along well with generative AI, will become embedded in many optimization algorithms, some of which may no longer fall neatly under SO. Third, we may move away from a strict black-box mindset and the obsession with general-purpose algorithms. Instead, solving important, high-impact application problems will drive the next wave of advances.

JONATHAN OZIK

How have recent advances in AI/ML changed how your organization undertakes the optimization of simulation models or how you go about research in SO?

The rise of AI/ML has created a virtuous cycle that further facilitated advances in AI/ML-based approaches, such as surrogate-based Bayesian optimization, for the purpose of developing better AI and ML models, e.g., for hyperparameter optimization. As a result, this has created a rich ecosystem of algorithms and software implementations that can be applied to SO. In our group, we have focused on how to leverage these advances and exploit the increasingly ubiquitous access to high-performance computing (HPC) facilities across national laboratories, academic institutions, and industry. This has led to the development of flexible and scalable "HPC-oriented algorithms" for SO (e.g., Fadikar et al. (2023), Binois et al. (2025), Robertson et al. (2025)) and software infrastructure that enables experimentation with and deployment of cutting-edge AI/ML SO algorithms on large-scale HPC resources (e.g., Ozik et al. (2016), Collier et al. (2023), Collier et al. (2024)).

Do you believe that emerging AI/ML techniques such as transformers or physics-informed neural networks have the potential to advance SO research? If so, how?

The simple and not very controversial answer is yes, though I think there are multiple pathways for these advancements to take place. From the point of view of spanning the data availability gap to train neural networks when the simulations are expensive to run, physics-informed neural networks and similar approaches that utilize what we know about a system to reduce the data needed for training will undoubtedly find sweet spots where the flexibility of neural networks can be applied to increasingly complex SO problems. On a more meta-level, there is significant potential for transformer models trained on the *practice* of SO across scientific domains to help find synergies across domains to suggest novel SO integrations or under-examined areas that may be difficult to otherwise discover.

The ubiquity of parallel computing resources has been changing how SO algorithms are designed and redefined what constitutes the state of the art. Do you foresee more recent (or near-future) advances in computing hardware (e.g., GPUs and quantum computing) doing the same?

The largest HPC resources are increasingly hybrid ones, incorporating heterogeneous compute nodes with GPUs and other accelerators co-existing with the more traditional CPUs. This opens up the possibility of optimizing the type of compute resource for each task that needs to be run in an SO workflow. For example, an SO algorithm that can utilize the concurrency of GPUs to efficiently generate candidate simulation parameters can be combined with CPU-optimized simulations. A crucial element in advancing these types of approaches will be infrastructure that lowers the barriers to enable such workflows so that both SO researchers and simulation practitioners can readily benefit.

Digital twins have brought greater interest in the use of simulation optimization for near-real-time control. Do you see this shift to more online monitoring and decision making opening up new classes of SO algorithms?

I certainly hope that this is the case and our experiences in supporting public health decision making has focused our attention on developing fast "time-to-solution" approaches. Compared to a typical timeline for simulation-based scientific inquiry, which can easily expand to months or even years, for simulation-based decision support to be meaningful it has to occur in the timeline of decision making. During the COVID-19 pandemic this timeline was often less than a week. This forced us to determine ways in which we could shorten our simulation-based epidemiological analyses, with algorithmic research to exploit HPC concurrency in new SO algorithms, and also ways to automatically ingest and validate streaming data to be used in event-driven HPC SO workflows.

What are some things that you believe are limiting, or could ultimately limit, advances in SO research over the next 25 years? E.g., a lack of publicly available industry-scale test problems, a decline in the number of PhD students with interest in the field.

I think two of the biggest factors limiting SO research advances are 1) siloed simulation-based scientific and research communities and 2) the lack of accessibility to infrastructure that can facilitate SO development and deployment. In one way, these factors are linked in that the SO infrastructure could help in cross-fertilization, bringing together researchers from a variety of domains to work together to advance SO, which, in turn, would inform a co-evolution of the SO infrastructure. The SO infrastructure could be made to support multi-modal simulation and data integration to enable robust and rapid development of SO approaches that can be run on diverse large-scale computing resources.

Is there anything you worry about as an existential threat to the field of SO?

It's difficult to predict too far into the future, but the current trends I'm seeing across scientific domains seem to point to an increasing interest in simulation-based approaches and the utilization of increasingly capable SO methods. I suppose one possible risk is that with the many advances across disparate fields it could become easier to miss advances relevant to one's application domain.

In what aspects do you see the boundaries of SO being pushed in the next 25 years? E.g., the dimension of the problems, the time-scale of decision-making, the classes of problems that fall within the realm of SO.

An area of recent interest within our group has been in expanding the capabilities of multi-objective robust decision making approaches to increasingly complex problems (Lima et al. 2023). This requires navigating large policy and parameter spaces, while shifting the focus from optimizing for a predicted future to achieving robustness across a potentially large set of unknown futures. These types of approaches are particularly valuable when decisions need to be made despite large uncertainties in our understanding of the dynamics of a system, such as in unfolding public health emergencies or extreme meteorological events. Advances would need to be made in tackling the large problem dimensions and in turning around analyses within decision-relevant timelines.

Are there any adjacent research communities that you think could benefit the most from SO researchers having a presence in? Or that the SO community should be more aware of for finding useful ideas to borrow?

As mentioned a few times above, there are almost too many research communities that fit this description. What I think would benefit both the SO community and the other research communities is to create mechanisms for long term partnership. These could be funding mechanisms that eschew standard and long-standing disciplinary differentiation. They could promote partnerships across national lab, university, government, and industry. One benefit of longer-term engagements is the potential for developing shared understandings and increasing trust among the different partners.

What do you anticipate being the difference between the state of simulation optimization in 2025 versus 2050? Controversial or contrarian opinions welcome.

I think this is an issue of what the algorithmic and computational advances between now and then will allow us to abstract. In a similar way in which LLMs have provided new ways for one to specify *what* problem needs to be solved rather than *how* it should be solved, I imagine that we will see approaches that include the ability to quickly assess the comparative benefits of different algorithms, hardware, and data to apply to a problem rather than requiring the specification of any single SO approach *a priori*. This could also extend to generating the needed data, for example through automated experiments in self-driving labs.

BENJAMIN THENGVALL

As a general preface to my comments, I do not believe that we will see true artificial general intelligence (AGI). There will be powerful general models that will be trained and will answer any question or make any decision. However, relying on these for general automated decision making will lead in too many instances to negative economic consequences and human tragedies. AI/ML models will never be more than very useful and very powerful human decision aids.

Increasingly powerful and useful AI/ML-based models for many different specific use cases will be developed over time and relied on in many domains. These models will be successful where there are complex, repeatable, stable processes or systems and large amounts of historical and/or synthetic training data available related to those processes or systems.

Broadly defined, metaheuristics applied for black box optimization have always been a form of AI/ML. The best optimization approaches have strategies informed by training data, but adapt their search strategies as they proceed based on what they find. Metaheuristics are a class of AI/ML in which additional, specific training data is requested and received in real time.

At OptTek Systems, we continue to refine our SO approaches, seeking to find ever better solutions with ever fewer simulation trials. The recently released OptQuest v10 includes new uses of AI/ML techniques including Gaussian Process Regression (GPR) and Random Forests (RF) for surrogate modeling, improvements to bounded multiobjective Pareto frontier search, and new adaptive sampling capabilities. The new adaptive sampling capabilities are driven by dynamic surrogate modeling approaches and yield the highest quality response surface estimations with the fewest simulation trials. This adaptive sampling is much more efficient than static sampling based on parametric sweeps or design of experiment approaches and can be employed for overall response surface modeling or to quickly identify specific boundary conditions or areas of greatest change for one or multiple model responses.

Emerging AI/ML techniques, such as transformers or physics-informed neural networks, can be thought of as fast-running surrogate models. If a surrogate model is sufficiently accurate for a desired study, then it can be used in place of a simulation model in a traditional SO process to complete an optimization study more quickly or to explore a larger input domain. A very fast surrogate model can enable a huge number of function evaluations, enabling brute-force approaches and making a highly efficient SO process less important.

The drawbacks of transformers or physics-informed neural networks are the large amounts of data required to train them and their inability to answer questions for which they have not been trained. The enduring strength of simulation modeling is the ability to explore new and different systems and processes of variable complexity directly through changes to the model.

Real-time control of repeated, predictable processes is an ideal place to use reinforcement learning (RL) techniques or custom heuristics rather than a general black-box SO approach. For these kinds of problems, historical data is available and additional synthetic data from simulation models can be generated to provide sufficient training data for RL approaches.

Rather than harm our field, AI/ML and computational advances will allow more efficient building and optimization of ever more complex models. These advances will not replace the analyst in the process, just make them more efficient. The size and scope of models being created will increase as our ability to execute and analyze them increases. This has been true since the dawn of the computer age. Vastly more powerful machines and capabilities have led to larger, higher-fidelity models and analysis. This pattern will continue. Experts in many domains will be required to make the best use of these new, more powerful technologies.

At OptTek Systems, one of our top OptQuest researchers has a Ph.D. in computational physics and chemistry. Many of the developers and analysts we hire come from engineering fields rather than pure mathematics or computer science backgrounds. A deep understanding of problems and solutions in diverse areas of the physical sciences often yields general insights that can be applied to SO approaches.

In the future, there will be new opportunities for SO approaches. Models will continue to grow in complexity. Many will have embedded AI/ML decision agents that are being updated and re-trained with updated data over time. As models deployed for real-time use are being updated automatically, based on new data and trained agents, there will be a need for guardrails. There is an opportunity to embed SO in model continuous integration/continuous deployment processes for continuous verification and validation of models. An SO process can run a model many times with varied inputs, exposing hard breaks in a model, and the outputs of these runs can be automatically checked for expected and acceptable behaviors. Automated optimization of model outcomes (maximizing and minimizing) can stress test models to ensure that they stay within physical or desired boundaries.

SO and adaptive sampling techniques paired with appropriate simulation models can also be useful in developing synthetic training data for AI/ML models. “Garbage in, garbage out” will be as true for future AI/ML models as it has been for all previous modeling paradigms. SO approaches can ensure that the best and worst outcomes are found and accounted for in training data. An AI/ML model trained on data that does not include the best solutions will not provide the best value in production. Adaptive sampling approaches can be used to ensure that interesting behaviors are represented in training data and that balanced training data is collected.

SO research and practice will look much the same in 2050 as it does today. The best work will still be done by skilled, experienced specialists in different domains building models to faithfully represent complex systems and processes and then applying state of the art optimizers and analysis to find the best solutions for their problem. In 2050 the problems being studied will be larger, the computational horsepower will be greater, and models will be of higher fidelity, yielding greater benefit to organizations and society. Similarly to the period from 2000 to 2025, the next 25 years will bring evolutionary changes rather than a fundamental shift in our relationship with computers.

ACKNOWLEDGMENTS

Peter J. Haas gratefully acknowledges support from U.S. Army DEVCOM Analysis Center under Contract #W911QX-23-D-009/W911QX-23-F-0115 and from the National Science Foundation under grants 1943971 and 2211918.

REFERENCES

Binois, M., N. Collier, and J. Ozik. 2025. “A Portfolio Approach to Massively Parallel Bayesian Optimization”. *Journal of Artificial Intelligence Research* 82:137–167.

Cen, W., and P. J. Haas. 2023. "Efficient Hybrid Simulation Optimization via Graph Neural Network Metamodeling". In *2023 Winter Simulation Conference (WSC)*, 3541–3552 https://doi.org/10.1109/WSC60868.2023.10408474.

Collier, N., J. M. Wozniak, A. Fadikar, A. Stevens, and J. Ozik. 2024. "Distributed Model Exploration with EMEWS". In *2024 Winter Simulation Conference (WSC)*, 72–86 https://doi.org/10.1109/WSC63780.2024.10838848.

Collier, N., J. M. Wozniak, A. Stevens, Y. Babuji, M. Binois, A. Fadikar, *et al*. 2023. "Developing Distributed High-performance Computing Capabilities of an Open Science Platform for Robust Epidemic Analysis". In *2023 IEEE International Parallel and Distributed Processing Symposium Workshops (IPDPSW)*, 868–877. St. Petersburg, FL, USA: IEEE.

Davis, C. W., A. J. Jetter, and P. J. Giabbanelli. 2022. "Automatically Generating Scenarios from a Text Corpus: A Case Study on Electric Vehicles". *Sustainability* 14(13):7938.

Eckman, D. J., S. G. Henderson, S. Shashaani, and R. Pasupathy 2024. "Simulation Optimization Library". http://github.com/simopt-admin/simopt. Accessed 18th June 2025.

Fadikar, A., N. Collier, A. Stevens, J. Ozik, M. Binois, and K. B. Toh. 2023. "Trajectory-Oriented Optimization of Stochastic Epidemiological Models". In *2023 Winter Simulation Conference (WSC)*, 1244–1255 https://doi.org/10.1109/WSC60868.2023.10408258.

Fu, M. C. 2015. *Handbook of Simulation Optimization*, Volume 216. Springer.

Fu, M. C., S. Andradóttir, J. S. Carson, F. Glover, C. R. Harrell, Y.-C. Ho, *et al*. 2000. "Integrating Optimization and Simulation: Research and Practice". In *2000 Winter Simulation Conference (WSC)*, 610–616 https://doi.org/10.1109/WSC.2000.899770.

Fu, M. C., G. Bayraksan, S. G. Henderson, B. L. Nelson, W. B. Powell, I. O. Ryzhov *et al*. 2014. "Simulation optimization: A Panel on the State of the Art in Research and Practice". In *2014 Winter Simulation Conference (WSC)*, 3696–3706. IEEE https://doi.org/10.1109/WSC.2014.7020198.

Haas, P. J., N. C. Barberis, P. Phoungphol, I. G. Terrizzano, W.-C. Tan, P. G. Selinger *et al*. 2012. "Splash: Simulation Optimization in Complex Systems of Systems". In *2012 50th Annual Allerton Conference on Communication, Control, and Computing (Allerton)*, 414–425.

Haas, P. J., and G. Theodoropoulos. 2020. "Introduction to the Special Issue for Towards an Ecosystem of Simulation Models and Data". *ACM Transactions on Modeling and Computer Simulation* 30(4):1–3.

Lima, P. N. d., A. Stevens, R. Vardavas, J. Ozik, and R. J. Lempert. 2023. *Co-Designing Capabilities for a Robust Pandemic Response: Stakeholder Engagement for Visioning, Backcasting, and Evaluating New Decision-Support Capabilities*. Santa Monica, CA: RAND Corporation.

Meliou, A., A. Abouzied, P. J. Haas, R. R. Haque, A. L. Mai, and V. Vittis. 2025. "Data Management Perspectives on Prescriptive Analytics (Invited Talk)". In *28th International Conference on Database Theory, ICDT 2025, March 25-28, 2025, Barcelona, Spain*, edited by S. Roy and A. Kara, Volume 328 of *LIPIcs*, 2:1–2:12: Schloss Dagstuhl - Leibniz-Zentrum für Informatik.

Ozik, J., N. T. Collier, J. M. Wozniak, and C. Spagnuolo. 2016. "From Desktop to Large-Scale Model Exploration with Swift/T". In *2016 Winter Simulation Conference (WSC)*, 206–220 https://doi.org/10.1109/WSC.2016.7822090.

Pasupathy, R., and S. G. Henderson. 2011. "SimOpt: A library of simulation optimization problems". In *2011 Winter Simulation Conference (WSC)*, 4075–4085. IEEE https://doi.org/10.1109/WSC.2011.6148097.

Robertson, C., C. Safta, N. Collier, J. Ozik, and J. Ray. 2025. "Bayesian Calibration of Stochastic Agent Based Model via Random Forest". *Statistics in Medicine* 44(6):e70029.

Shashaani, S. 2024. "Simulation Optimization: An Introductory Tutorial on Methodology". In *2024 Winter Simulation Conference (WSC)*, 1308–1322. IEEE https://doi.org/10.1109/WSC63780.2024.10838809.

AUTHOR BIOGRAPHIES

SOUMYADIP GHOSH is a member of IBM Research in the mathematics of computation department at Yorktown Heights. His research focuses on stochastic optimization and decision making under uncertainty, with applications in training large statistical learning models with improved generalization in speech recognition and natural language processing, and previously on Smarter Grid and energy analytics, cloud infrastructure, and supply chain management. He is an associate editor of Stochastic Models and has served as a committee member for INFORMS / WSC. He holds a PhD from Cornell University (Ithaca, NY, USA), an MS from Univ of Michigan (Ann Arbor, MI, USA) and a B Tech from Indian Institute of Technology (Chennai, TN, INDIA). His email address is ghoshs@us.ibm.com.

PETER J. HAAS is a Professor of Computer Science and Adjunct Professor of Industrial Engineering at the University of Massachusetts Amherst. His research interests lie at the interface of stochastic simulation, data management for decision-making, and machine learning. He was a Principal Research Staff Member at the IBM Almaden Research Center from 1987-2017 and was also a Consulting Professor in the Department of Management Science and Engineering at Stanford University from 1992-2017. He is a Fellow of INFORMS and ACM, has authored over 160 conference publications, journal articles, and books, and has been granted over 30 patents. His email address is phaas@cs.umass.edu and his web page is

https://www.cics.umass.edu/people/haas-peter.

L. JEFF HONG is a Professor of Industrial and Systems Engineering at the University of Minnesota. His research interests include stochastic simulation, stochastic optimization, risk management and supply chain management. He is currently an Associate Editor of *Management Science* and *ACM Transactions on Modeling and Computer Simulation*, and the Simulation Department Editor of *Naval Research Logistics*. He was the Simulation Area Editor of *Operations Research* (2018-2023) and President of INFORMS Simulation Society (2020-2022). His email address is lhong@umn.edu.

JONATHAN OZIK is Principal Computational Scientist at Argonne National Laboratory, Senior Scientist with Department of Public Health Sciences affiliation in the Consortium for Advanced Science and Engineering (CASE) at the University of Chicago, and Senior Institute Fellow in the Northwestern Argonne Institute of Science and Engineering (NAISE) at Northwestern University. He applies large-scale agent-based models and large-scale model exploration across research domains and computational modeling methods. He is the lead of the Repast project for agent-based modeling toolkits and the Extreme-scale Model Exploration with Swift (EMEWS) framework for large-scale model exploration on high-performance computing resources. His e-mail address is jozik@anl.gov.

BENJAMIN THENGVALL is Chief Executive Officer of OptTek Systems, Inc. based in Boulder, Colorado. OptTek Systems provides the OptQuest simulation optimization software that is available in many commercial simulation tools. At OptTek Systems he leads the team as they combine professional software development with world class optimization and analytics. He holds a Ph.D. in Operations Research and Industrial Engineering from the University of Texas at Austin. He taught at the Peter Kiewit Institute at the University of Nebraska and at the McCombs School of Business at the University of Texas. He is the author of numerous journal articles and book chapters and holds multiple patents related to software and optimization. His e-mail address is thengvall@opttek.com and OptTek's website is https://www.opttek.com.