

We are IntechOpen, the world's leading publisher of Open Access books Built by scientists, for scientists

6,900

Open access books available

185,000

International authors and editors

200M

Downloads

Our authors are among the

154

Countries delivered to

TOP 1%

most cited scientists

12.2%

Contributors from top 500 universities



WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com



Chapter

Artificial Intelligence and Its Application in Optimization under Uncertainty

*Saeid Sadeghi, Maghsoud Amiri
and Farzaneh Mansoori Mooseloo*

Abstract

Nowadays, the increase in data acquisition and availability and complexity around optimization make it imperative to jointly use artificial intelligence (AI) and optimization for devising data-driven and intelligent decision support systems (DSS). A DSS can be successful if large amounts of interactive data proceed fast and robustly and extract useful information and knowledge to help decision-making. In this context, the data-driven approach has gained prominence due to its provision of insights for decision-making and easy implementation. The data-driven approach can discover various database patterns without relying on prior knowledge while also handling flexible objectives and multiple scenarios. This chapter reviews recent advances in data-driven optimization, highlighting the promise of data-driven optimization that integrates mathematical programming and machine learning (ML) for decision-making under uncertainty and identifies potential research opportunities. This chapter provides guidelines and implications for researchers, managers, and practitioners in operations research who want to advance their decision-making capabilities under uncertainty concerning data-driven optimization. Then, a comprehensive review and classification of the relevant publications on the data-driven stochastic program, data-driven robust optimization, and data-driven chance-constrained are presented. This chapter also identifies fertile avenues for future research that focus on deep-data-driven optimization, deep data-driven models, as well as online learning-based data-driven optimization. Perspectives on reinforcement learning (RL)-based data-driven optimization and deep RL for solving NP-hard problems are discussed. We investigate the application of data-driven optimization in different case studies to demonstrate improvements in operational performance over conventional optimization methodology. Finally, some managerial implications and some future directions are provided.

Keywords: Data-driven optimization, Decision making under uncertainty, Mathematical optimization, Machine learning, Deep learning, Reinforcement learning

1. Introduction

Optimization is applied in many engineering and science fields, including manufacturing, inventory control, transportation, finance, economics [1, 2]. Some parameters involved in optimization problems are subject to uncertainty in real practice due to various reasons, including measurement errors and uncontrollable disturbances [3]. Such uncertain parameters can be product demand and price, raw material supply chain cost, production cost. Disregarding uncertainty could, unfortunately, render the solution of a deterministic optimization problem suboptimal or even infeasible. In the era of big data and deep learning (DL), intelligent use of data and knowledge extraction from them have great benefits for organizations. Besides, in today's complex world, uncertainty on the lack of enough data has been replaced by too much data, which creates numerous opportunities for academicians and practitioners [4]. A large amount of interactive data is routinely created, collected, and archived in different industries; these data are becoming an important asset in process operation, control, and design. Explosive growth in volume and different sorts of data in organizations has created the need to develop technologies that can intelligently and rapidly analyze large volumes of data [4]. The traditional optimization methods cannot face big data satisfactorily. Nowadays, a wide array of emerging machine learning (ML) techniques can be leveraged to analyze data and extract relevant, accurate, and useful information and knowledge for smart decision-making. More recently, the dramatic progress of ML, especially DL over the past decade, coupled with recent advances in mathematical programming, sparks a flurry of interest in data-driven optimization [5, 6]. The uncertainty model is formulated based on a data-driven optimization paradigm, allowing uncertainty data to speak for themselves in the optimization algorithm. In this way, rich knowledge underlying uncertainty data set can be extracted and harnessed automatically for smart and data-driven decision making. In such situations, the effectiveness and efficiency of traditional operational research methods are questionable. In recent years, the inefficiency of traditional methods in facing the uncertainty caused by big data has led researchers to integrate *artificial intelligence* (AI) with optimization methods. Integrating AI and optimization methods play a crucial role in solving problems in dynamic and uncertain environments. Nowadays, a wide range of ML tools has emerged that can be leveraged to analyze data automatically and extract relevant, accurate, and useful information for smart and data-driven decision-making. DL is one of the most rapidly growing sub-fields of the ML technique that demonstrates remarkable power in processing and deciphering a large volume of data through a complex architecture. Reinforcement learning (RL) is another ML sub-field that recently is applied to tackle complex sequential decision problems. This branch of ML epitomizes a step toward building autonomous systems by understanding the visual world.

The objective of this study is to provide an overview of the use of data-driven optimization in academia and practice from the following perspectives:

1. How can integrate artificial intelligence techniques with mathematical programming models to develop the intelligent and data-driven *decision support systems* (DSS) in uncertain conditions caused by big data?
2. We demonstrate the use of data-driven optimization across three case studies from operations research.

In this regard, this chapter reviews recent advances in data-driven optimization that highlight the integration of mathematical programming and ML for decision-making under uncertainty and identifies potential research opportunities. We compare data-driven optimization performance to conventional models from optimization methodology. We summarize the existing research papers on data-driven optimization under uncertainty and classify them into three categories: Data-driven stochastic program, Data-driven robust optimization, and Data-driven chance-constrained, according to their unique approach to uncertainty modeling distinct optimization structures. Based on the literature survey, we identify five promising future research directions on optimization under uncertainty in the era of big data and DL, (i) Employment of DL in the field of data-driven optimization under uncertainty, (ii) Deep data-driven models, (iii) Online learning-based data-driven optimization, (iv) Leveraging RL techniques for optimization, and (v) Deep RL for solving NP-hard problems and highlight respective research challenges and potential methodologies. We conducted an extensive literature review on recent papers published across the premier journals between 2002 and 2020 in our field, namely, the European Journal of Operational Research, Operations Research, Journal of Cleaner Production, Production and Operations Management, Journal of Operations Management, Computers in Industry, and Decision Sciences. We specifically searched for papers containing “big data”, “data-driven optimization”, “artificial intelligence”, “machine learning”, “deep learning”, and “Reinforcement learning”. However, our research into the existing literature reveals a scarcity of research works utilizing DL and RL in these disciplines.

The remainder of this paper is organized as follows: Section 2 provides an introduction to the mathematical optimization method. In Section 3, a brief review of AI methods such as ML, DL, and RL is provided. In sections 4–6, applying different ML, DL, and RL techniques in data-driven optimization under uncertainty are presented. Finally, the book chapter ends with the conclusion, some managerial implications, and future research recommendations.

2. Mathematical optimization under uncertainty

In recent years, mathematical programming techniques for decision-making under uncertainty have been applied in many science and engineering areas, including process design, production scheduling and planning, design, control, and supply chain optimization.

Optimization under uncertainty has been motivated because parameters involved in optimization models for design, planning, scheduling, and supply chains are often uncertain parameters such as product demands, prices of raw material, product, and yields.

A major modeling decision in optimization under uncertainty is whether the decision-maker should rely on robust optimization to use stochastic programming [7]. The robust optimization basis idea is to guarantee feasibility over a specified uncertainty set. In contrast, in the stochastic programming approach, a subset of decisions is set by anticipating that recourse actions can be taken once the uncertainties are revealed over a pre-specified scenario with discrete probabilities of uncertainties. The robust optimization basis idea is to guarantee feasibility over a specified uncertainty set. In contrast, in the stochastic programming approach, a subset of decisions is set by anticipating that recourse actions can be taken once the uncertainties are revealed over a pre-specified scenario with discrete probabilities of the uncertainties.

In general, the optimization approach tends to be more appropriate for short-term scheduling problems in which feasibility over a specified set of uncertain parameters is a major concern and when there is not much scope for recourse decisions. On the other hand, the stochastic programming approach tends to be more appropriate for long-term production planning and strategic design decisions.

In this section, the authors briefly explain three leading modeling paradigms for optimization under uncertainty, namely stochastic programming, robust optimization, and chance-constrained programming.

2.1 Stochastic programming

Under uncertainty, a common decision-making approach is stochastic programming, aiming to optimize the expected objective value across all the uncertainty realizations [8]. The stochastic programming key idea is to model the randomness in uncertain parameters with probability distributions. In this approach, the first stage, all the decisions must be made without knowing precisely the uncertainty realizations. The decision-maker then waits for resolving the uncertainty and knowing the actual value of the uncertain parameters. In the second stage, the decision-maker takes corrective actions after uncertainty is revealed. The stochastic programming approach has demonstrated various applications, such as inventory routing problems [9], supply chain network modeling [10], distributed energy systems design [11], optimal tactical planning [12], and energy management [13].

2.2 Robust optimization

Robust optimization is a promising alternative paradigm to optimization under uncertainty that does not require accurate knowledge on probability distributions of uncertain parameters. The key idea of robust optimization is to construct a convex uncertainty set of possible realizations of the uncertain parameters and then optimize against worst case realization within this set [14]. A robust optimization framework aims to hedge against the worst-case within the uncertainty set. The robust optimization approach has demonstrated various applications, such as supply chain planning [15], supply chain management [16], inventory management [17].

2.3 Chance constrained programming

Chance constrained programming is another common paradigm for optimization under uncertainty with soft probabilistic constraints on the decision variable in place of the hard ones present in robust optimization. Specifically, chance-constrained programming aims to compute a solution that satisfies the constraint with high probability in an uncertain environment. In the chance-constrained optimization paradigm, the probability distribution of uncertain parameters should be known to capture the randomness of uncertain parameters. Chance constrained programs are increasingly used in many applications, such as robotics [18], stochastic model predictive control [19], energy systems [20], and autonomous driving [21].

All mathematical optimization methods are inefficient and ineffective in facing uncertainty caused by the large volume of data. In the following section, three AI areas as tools for compensating the weaknesses of mathematical optimizing methods are introduced. The term “AI” is often used to describe machines (or computers) that

mimic “cognitive” functions that humans associate with the human mind, such as “learning” and problem-solving” [22]. A brief description of the three main areas of AI, including ML, DL, and RL, is provided in the following.

3. Machine learning (ML)

ML is a sub-area of AI that can automatically extract artificial information and knowledge from diverse data types with high speed. The advancement in computational power and the emergence of big data have led to ML optimization and simulation methods. Analysis of big data by ML offers considerable advantages for integrating and evaluating large amounts of complex data [23]. ML solutions have scalability and flexibility compared with traditional statistical methods, making them deployable for many tasks, such as clustering, classification, and prediction. ML models have demonstrated outstanding ability for learning intricate patterns that enable them to make predictions about unobserved data. In addition to using models for prediction, it can accurately interpret what a model has learned.

ML techniques use large sets of data inputs and outputs to recognize patterns and effectively “learn” to make autonomous recommendations or decisions [24]. These algorithms attempt to minimize their errors and maximize the likelihood of their predictions being true [25]. The predictive abilities of ML models are increasingly applied in various fields such as healthcare, genetic, finance, education, and production.

3.1 Deep learning (DL)

In real applications, uncertainty data exhibit highly complex and nonlinear characteristics. DL is an ML technique and includes algorithms and computational models that imitate the architecture of the biological neural networks in the brain [artificial neural networks (ANNs)] [25]. The DL technology consists of numerous layers responsible for extracting important abstract features from the data [26]. It can process a large volume of data through a complex architecture [27]. DL algorithms can uncover useful uncertainty data patterns for mathematical programming [28]. Recently, the DL technique has been used in optimization under uncertainty.

3.2 Reinforcement learning (RL)

In particular, RL has gained tremendous attraction recently in different research areas. In RL, an agent gains experience from directly interacting with the environment and selecting an optimal action. RL is concerned with how a software agent should choose an action to maximize a cumulative reward. Combining DL with the RL technique creates the concept of deep RL, which enables RL to tackle the previously intractable decision-making problems. Inspired by the recent advances of deep RL in video games, robotics, and cyber-security, it has been used in optimization problems.

After introducing mathematical optimization methods and three main AI areas, it is time to pay to apply ML, DL, and RL methods in data-driven optimization. They are discussed in turn in the following sections.

4. Leveraging ML techniques for hedging against uncertainty in data-driven optimization

In the big data and ML era, a large amount of interactive data are routinely generated and collected in different industries. Intelligence and data-driven analysis and decision-making have a critical role in process operations, design, and control. The success of the DSS depends primarily on the ability to process and analyze large amounts of data and extract relevant and useful knowledge and information from them. In this context, the data-driven approach has gained prominence due to its provision of insights for decision-making and easy implementation. The data-driven optimization framework is a hybrid system that integrates AI and optimization methods for devising a data-driven and intelligent DSS. The data-driven system applied ML techniques for uncertainty modeling. The data-driven approach can discover various database patterns without relying on prior knowledge while also can handle multiple scenarios and flexible objectives. It can also extract information and knowledge from data without speed [29, 30].

The framework of data-driven optimization under uncertainty could be considered a hybrid system that integrates the data-driven system based on ML to extract useful and relevant information from data. The model-based system is based on mathematical programming to derive the optimal decisions from the information [28]. The inability of traditional optimization methods to analyze big data, as well as recent advances in ML techniques, made data-driven optimization a promising way to hedge against uncertainty in the era of big data and ML. Therefore, these promises create the need for organic integration and effective interaction between ML and mathematical programming. In existing data-driven optimization frameworks, data serve as input to a data-driven system. After that, useful, accurate, and relevant uncertainty information is extracted through the data-driven system and further passed along to the model-based system based on mathematical programming for rigorous and systematic optimization under uncertainty, using paradigms such as robust optimization and stochastic programming.

The various ML techniques and their potentials applications in data-driven optimization under uncertainty are presented in the following.

4.1 Distributionally robust optimization

The stochastic programs are used where the distribution of the uncertain parameters is only observable through a finite training dataset [31]. As the primary assumption in the stochastic programming approach, the probability distribution of uncertain parameters should be clear. However, such complete knowledge of parameters probability distribution is rarely available in practice. In practice, instead of knowing the actual distribution of an uncertainty parameter, what the decision-maker has is a set of historical/ or real-time uncertainty data and possibly some prior structure knowledge of the probability. Also, the assumed possibility distribution of uncertain parameters may deviate from their actual distribution. Moreover, relying on a single probability distribution could lead to sub-optimal solutions or even lead to the deterioration in out-of-sample performance [32]. Motivated by these stochastic programming weaknesses, DRO emerges as a new data-driven optimization paradigm that hedges against the worst-case distribution in an ambiguity set [28]. DRO paradigm integrates data-driven systems and model-based systems. A data-driven approach is applied in the DRO model to construct an uncertainty set of probability

distributions from uncertainty data through statistical inference and big data analytics [28]. In data-driven stochastic modeling, the uncertainty is modeled via a family of probability distributions that well capture uncertainty data on hand [28]. This set of probability distributions is referred to as an ambiguity set. With this ambiguity set, a model is then proposed for problem design. Finally, a solution strategy is applied for solving the optimization problem. For example in the literature, the Wasserstein metric has been used, to construct a ball in the space of (multivariate and non-discrete) probability distributions centered at the uniform distribution on the training samples, to seek decisions that perform best in view of the worst-case distribution within this Wasserstein ball [31]. Different practical approaches, such as the moment-based, and the adopted distance metric, were employed for uncertainty constructing [33, 34], and [31]. DRO is an effective method to address the inexactness of probability distributions of uncertain parameters in decision-making under uncertainty that can be applied for optimizing supply chain activities, for planning and scheduling under uncertainty. This way reduces the modeling difficulty for uncertain parameters. Wang & Chen [35] proposed a two-stage DRO model considering scarce data of disasters. A moment-based fuzzy set describes uncertain distributions of blood demand to optimize blood inventory prepositioning and relief activities together. Chiou [36], to regulate the risk associated with hazardous material transportation and minimize total travel cost on the interested area under stochasticity, presented a multi-objective data-driven stochastic optimization model to determine generalized travel cost for hazmat carriers. Gao et al. [37] proposed a two-stage DRO model for better decision making in optimal design and shale gas supply chains under uncertainty. They applied a data-driven approach to construct the ambiguity set based on principal component analysis and first-order deviation functions. In the other study, Ning & You [28] proposed a novel data-driven Wasserstein DRO model for biomass with agricultural waste-to-energy network design under uncertainty. They proposed a data-driven approach to construct the Wasserstein ambiguity set for the feedstock price uncertainty, which is utilized to quantify their distances from the data-based empirical distribution.

4.2 Data-driven robust optimization

A robust optimization is a popular approach for optimization under uncertainty. It defines an uncertainty set of possible realizations of the uncertain parameters and then optimizes against worst-case realizations within this set [5, 6]. In real-world applications, the underlying distribution of uncertainties may be intrinsically complicated and vary under different circumstances [38]. Choosing the accurate underlying distribution of uncertainties and the uncertainty sets by prior knowledge is somewhat challenging in practice. In robust optimization, the uncertainty is formed as an uncertainty set in which any point is a possible scenario [39]. Since the uncertainty set includes the worst case, robust optimization may be over-conservative. It is essential to apply the appropriate approach to construct the uncertainty set and adjust the conservatism level simultaneously [39]. As an essential ingredient in robust optimization, uncertainty sets endogenously determine robust optimal solutions and, therefore, should be devised with special care [28]. However, uncertainty sets in the conventional robust optimization methodology are typically set a priori using a fixed shape and model without providing sufficient flexibility to capture the structure and complexity of uncertainty data [28]. For instance, the geometric shapes of uncertainty set in the conventional robust optimization methodology do not change with

the intrinsic structure and complexity of uncertainty data. Furthermore, these uncertainty sets are specified by a finite number of parameters, thereby limiting modeling flexibility. Motivated by this knowledge gap, data-driven robust optimization emerges as a powerful paradigm for addressing uncertainty in decision making.

Choosing a good uncertainty set enables robust optimization models to provide better solutions than other approaches solutions [5, 6]. Poor choice of the uncertainty set makes robust optimization model overly conservative or computationally intractable. In the era of big data, many data are routinely generated and collected containing abundant information about the distribution of uncertainties; thereby, ML tools can construct the uncertainty sets based upon these data. Data-driven robust optimization is a new paradigm for hedging against uncertainty in the era of big data. The ML tools can be applied to estimate data densities with sufficient accuracy and construct an appropriate uncertainty set based upon intelligent analysis and the use of uncertainty data for modeling robust optimization problems. A desirable uncertainty set shall have enough flexibility to adapt to the intrinsic structure behind data, thereby characterizing the underlying distribution and facilitating the solutions.

Data-driven robust optimization could be considered a “hybrid” system that integrates the data-driven system based on ML to construct the uncertainty set from historical uncertainty data. The model-based system is based on the robust programming model to derive the optimal decisions from the information. More specifically, data serves as input to a data-driven system. **Figure 1** presents the data-driven optimization paradigm framework. After that, the data-driven method constructs the uncertainty set to extract information from historical data fully. Constructing the uncertainty sets based upon historical data can be considered as an unsupervised learning problem from an ML perspective. So, data-driven robust optimization is a hybrid system that utilizes ML techniques to design data-driven uncertainty sets and develops a robust optimization problem from the data-driven set. Different effective unsupervised learning models such as the Dirichlet process mixture model, maximum likelihood estimation, principal component analysis, regular and conservative support vector clustering, Bayesian ML, and kernel density estimation were employed for uncertainty constructing, which could provide powerful representations of data distributions [38, 40, 41]. Uncertainty set is the set that can offer robust solutions with a conservatism level. Furthermore, this uncertainty set is finally given to the model-based system based on robust optimization to obtain robust solutions under uncertainty.

ML methods of support vector clustering-based uncertainty set (SVCU) and conservative support vector clustering-based uncertainty set (CSVCU) have been applied to finding an enclosed hypersphere with minimum volume which is able to cover all data samples as tightly as possible as uncertainty sets. Conservative support vector clustering is the most suitable choice for obtaining robust solutions in cases with sufficient data to construct an uncertainty set enclosing future data with a high confidence level [42]. Furthermore, it is the most effective choice for obtaining lower

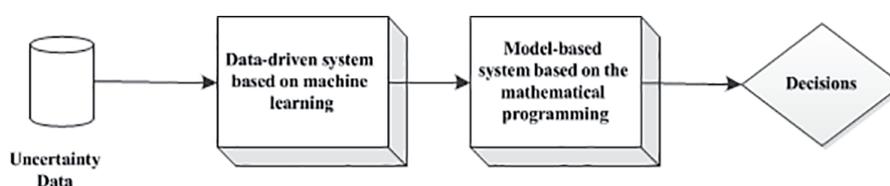


Figure 1.
The schematic of the data-driven optimization paradigm framework.

conservative solutions. On the other hand, CSVCU is suitable for highly conservative decision-makers since it is the only set that can offer robust solutions with a high conservatism level, particularly when there is limited data [42]. A data-driven robust optimization under correlated uncertainty was proposed to hedge against the fluctuations generated from continuous production processes in an ethylene plant [43]. For capturing and enrich the valid information of uncertainties, a copula-based method is introduced to estimate the joint probability distribution and simulate mutual scenarios for uncertainties. A deterministic and data-driven robust optimization framework was proposed for energy systems optimization under uncertainty. The uncertainty set is constructed by support vector clustering based on real industrial data [39]. A data-driven robust optimization was applied to design and optimize the entire wastewater sludge to-biodiesel supply chain [42]. They develop a conservative support vector clustering (CSV) method to construct an uncertainty set from limited data. The developed uncertainty set encloses the fuzzy support neighborhood of data samples, making it practical even when the available data is limited.

4.3 Data-driven chance-constrained program

Chance constrained programming is a practical and convenient approach to control risk in decision-making under uncertainty. However, due to unknown probability distributions of uncertainty parameters, the solution obtained from a chance-constrained optimization problem can be biased. In practice, instead of knowing the actual distribution of an uncertainty parameter, only a set of historical/ or real-time uncertainty data, which can be considered as samples taken from the actual (while ambiguous) distribution, can be observed and stored. On the other hand, even if the probability distribution of an uncertainty parameter is available, the chance-constrained program is computationally cumbersome. Motivated by Chance constrained programming weaknesses, data-driven chance-constrained optimization emerges as a new data-driven optimization paradigm. The data-driven stochastic programming approach is a data-driven risk-averse strategy to handle uncertainties in the era of big data effectively.

In contrast to the data-driven stochastic programming approach, data-driven chance-constrained programming is another paradigm focusing on chance constraint satisfaction under the worst-case probability instead of optimizing the worst-case expected objective. Although both data-driven chance-constrained programs and DRO adopt ambiguity sets in the uncertainty models, they have distinct model structures. Specifically, the data-driven chance-constrained program features constraints subject to uncertainty in probability distributions. Simultaneously, DRO typically only involves the worst-case expectation of an objective function concerning a family of probability distributions [28]. In the data-driven stochastic programming approach, historical data is utilized to learn the uncertain parameters' distributions.

Data-driven chance-constrained programs with moment-based ambiguity sets, distance-based ambiguity set, Prohorov metric-based ambiguity sets [44], ϕ -divergence based ambiguity set [45], kernel smoothing method [46], Wasserstein ambiguity set [47].

Ghosal and Wiesemann [48] applied for Data-driven chance-constrained programs in the capacitated vehicle routing problem (CVRP), which asks for the cost-optimal delivery of a single product geographically dispersed customers through a fleet of capacity-constrained vehicles. They model the customer demands as a random vector whose distribution is only known to belong to an ambiguity set.

4.4 Leveraging DL techniques in the data-driven optimization

The recent development in the data science field, AI, and ML techniques have enabled intelligent and automated DSS and real-time analytics coupled with computing power improvements. Thus, AI techniques are applied to big data sources to extract the knowledge-based rules or identify the underlying rules and patterns by ML techniques, to drive the systems toward set objectives. DL is an ML technique that can extract high levels of information and knowledge from massive data volumes. DL algorithms consist of multiple processing layers to learn representations of data with multiple abstraction levels [26]. For example, recently, DL techniques have been used to accurately forecasting customer demand, price, and inventory leading to optimization of supply chain performance. An intelligent forecasting system leads to optimize performance, reduce costs, and increase sales and profit. DL techniques can apply deep neural network architectures to solve various complex problems. The DL paradigm requires high computing power and a large amount of data for training. The recent advances in parallel architectures and GUP (Graphical Processing Unit) enabled the necessary computing power required in deep neural networks (DNN). The emergence of advanced IoT and blockchain technologies has also solved the need for a large amount of data to learn. IoT and blockchain result in massive amounts of streaming real-time data often referred to as “big data,” which brings new opportunities to control and manage supply chains [49]. Optimizing the parameters in DNN is a challenging undertaking. Several optimization algorithms such as Adam, Adagrad, RMSprop, have been proposed to optimize the network parameters in DNN and improve generalizability. This technique, which stabilizes the optimization, paved the way for learning deeper networks [50]. In real applications, uncertainty data exhibit very complex and highly nonlinear characteristics. DNN can be used to uncover useful patterns of uncertainty data for optimizing under uncertainty [28]. Deep data-driven optimization could be considered a “hybrid” system that integrates the deep data-driven system based on DL to forecast the uncertainty parameters. The model-based system is based on mathematical programming to drive the optimal decisions from predicted parameters (the deep data-driven system). In the DL-based system, DNN has been applied to analyze features, complex interactions, and relationships among features of a problem from samples of the dataset and learn model, which can be used for demand, inventory, and price forecasting. Kilimci et al. [51] developed an intelligent demand forecasting system based on the analysis and interpretation of the historical data using different forecasting methods, including support vector regression algorithm, time series analysis techniques, and DL models. In a study, the Auto-Regressive Integrated the backpropagation (BP) network method, recurrent neural network (RNN) method, and Moving Average (ARIMA) model were tested to forecast the price of agricultural products [52]. Yu et al. [53] developed an online big-data-driven forecasting model of Google trends to improve oil consumption prediction. Their proposed forecasting model considers traditional econometric models (LogR and LR) and typical AI techniques (BPNN, SVM, DT, and ELM).

Accurate automatic optimization heuristics are necessary for dealing with the complexity and diversity of modern hardware and software. ML is a proven technique for learning such heuristics, but its success is bound by the quality of the features used. Developers must handcraft these features through a combination of expert domain knowledge and trial and error. This makes the quality of the final model directly dependent on the skill and available time of the system architect. DL techniques are a better way to build heuristics. A deep neural network can learn

heuristics over raw code entirely without using code features. The neural network simultaneously constructs appropriate representations of the code and learns how best to optimize, removing the need for manual feature creation. DNN can improve the accuracy of models without the help of human experts. Generally, this approach is a fundamental way to integrate forecast approaches into mathematical optimization models. First, a probabilistic forecast approach for future uncertainties is given by exploiting the advanced DL structures. Second, a model-based system based on mathematical programming is applied to derive the optimal decisions from the forecasting data. Comparison and evaluation of the forecasting models are significant since DL models can have different performances depending on the properties of the data [54, 55]. The performances of DL models differ according to the forecasting time, training duration, target data, and simple or ensemble structure [56, 57].

In a study, Nam et al. [54, 55] applied DL-based models to forecast fluctuating electricity demand and generation in renewable energy systems. This study compares and evaluates DL models and conventional statistical models. The DL models include DNN, long short-term memory, gated recurrent unit, and the disadvantages of conventional statistical models such as multiple linear regression and seasonal autoregressive integrated moving average. In another study, the operation of a cryogenic NGL recovery unit for the extraction of NGL has been optimized by implementing data-driven techniques [58]. The proposed approach is based on an optimization framework that integrates dynamic process simulations with two DL-based surrogate models using a long short-term memory (LSTM) layout with a bidirectional recurrent neural network (RNN) structure. Kilimci et al. [51] developed an intelligent demand forecasting system. This improved model is based on analyzing and interpreting the historical data using different forecasting methods, including time series analysis techniques, support vector regression algorithm, and DL models.

Accessing a sufficient amount of data for some optimization models is a practical challenge. For example, the quality of scenario-based optimization frameworks strongly depends on access to a sufficient amount of uncertain data. However, in practice, the amount of uncertainty data sampled from the underlying distribution is limited. On the other hand, acquiring a sufficient amount of uncertainty data is extremely time-consuming and expensive in some cases, which leads to the limited application of some approaches [59]. To deal with the practical challenge of requiring an insufficient amount of data, deep generative models emerge as a new paradigm to generate synthetic uncertainty data with the aim of better decisions with insufficient uncertainty data. DL techniques could be applied to learn the useful intrinsic patterns from the available uncertainty data and generate synthetic uncertainty data. More specifically, in deep generative models, the correct data distribution is mimicked either implicitly or explicitly by the DL techniques. Then the learned distribution is used to generate new data points referred to as synthetic data [28]. After that, these synthetic data serve as input to an optimizing model to derive the optimal decisions. Some of the most commonly used deep generative models are variational autoencoders generative and adversarial networks [26]. These synthetic uncertainty data generated by the DL techniques can be potentially useful in the scenario-based optimization model.

4.5 Deep data-driven models

DL models are a class of approximate models proven to have strong predictive capabilities for representing complex phenomena [60]. Approximate models are

currently experiencing a radical shift due to the advent of DL. However, our research into the existing literature reveals a scarcity of research utilizing DL in approximate modeling. The introduction of DL models into an optimization formulation provides a means to reduce the problem complexity and maintain model accuracy [60]. Recently it has been shown that DL models in the form of neural networks with rectified linear units can be exactly recast as a mixed-integer linear programming formulation. DL is a method to approximate complex systems and tasks by exploiting large amounts of data to develop rigorous mathematical models [60].

Using DNN to model real-world problems is a powerful tool, as they provide an efficient abstraction that can be used to analyze the structure of the task at hand. The rigorous mathematical model is developed based on neural networks modeling complex systems and optimizing their operations in the deep data-driven model framework. This approximate model is developed by exploiting large amounts of data using DL techniques. Then the solving method is applied to obtain the optimal solutions of the developed optimization model. Developing an optimal solution to the approximate model remains challenging [60].

Pfrommer et al. [61] utilized a stochastic genetic algorithm to optimize a composite textile draping process where a neural network was utilized as a surrogate model. Marino et al. [62] presented an approach for modeling and planning under uncertainty using deep Bayesian neural networks (DBNNs). They use DBNNs to learn a stochastic model of the system dynamics. Planning is addressed as an open-loop trajectory optimization problem. In the study, DL-based surrogate modeling and optimization were proposed for microalgal biofuel production and photobioreactor design [63]. This surrogate model is built upon a few simulated results from the physical model to learn the sophisticated hydrodynamic and biochemical kinetic mechanisms; then adopts a hybrid stochastic optimization algorithm to explore untested processes and find optimal solutions. Tang & Zhang [64] developed a deep data-driven framework for modeling combustion systems and optimizing their operations. First, they developed a deep belief network to model the combustion systems. Next, they developed a multi-objective optimization model by integrating the deep belief network-based models, the considered operational constraints, and the control variable constraints.

4.6 Online learning-based data-driven optimization

In conventional data-driven optimization frameworks, a set of uncertainty data serves as input to the data-driven system, in which learning typically takes place once by using learning techniques. This approach fails to account for real-time uncertainty data [28]. For example, in the DRO method, the uncertainty set of probability distributions is constructed from uncertainty data. Once the uncertainty sets of probability distributions are obtained, they remain fixed for the model-based system based on mathematical programming and are not updated or refined. However, in real practice, a vast number of uncertainty data are generated and collected sequentially in an online fashion; therefore, data-driven systems should be developed to analyze the real-time data. An online-learning-based data-driven optimization framework emerges as a new data-driven optimization paradigm. Learning takes place iteratively to account for real-time data, and the data-driven system is updated in an online fashion. The framework of online-learning-based data-driven optimization could be considered a hybrid system that integrates the online data-driven and model-based systems. In the online data-driven system, the real-time uncertainty data should

be saved and analyzed sequentially based on ML to extract sequentially useful and relevant information from the real-time data. The online data-driven system (such as the uncertainty sets, probability distributions sets, and forecasting data) that serve as input to a model-based system should be updated in an online fashion. Then in the model-based system, the optimal decisions are made sequentially from the real-time information based on mathematical programming. There is a “feedback” channel for information flow returning from the model-based system to the data-driven system in this framework. The information flow is fed into the mathematical programming problem from the ML results. Using the feedback control strategy delivers amazingly superior system performance (e.g., stability, robustness to disturbances, and safety) [28]. **Figure 2** presents the potential schematic of the online learning-based data-driven optimization system.

The online-learning-based data-driven optimization framework, updating the data-driven systems, and developing efficient algorithms to solve online learning-based mathematical programming problems have become challenging.

4.7 Leveraging RL techniques for optimization

RL has transformed AI, especially after the success of Google DeepMind. This branch of ML epitomizes a step toward building autonomous systems by understanding the visual world. Deep RL is currently applied to different sorts of problems that were previously obstinate. In this subsection, the authors will analyze Deep RL and its applications in optimization.

RL is one of the ML areas recently applied to tackle complex sequential decision problems. RL is concerned with how a software agent should choose an action to maximize a cumulative reward. RL is considered an optimal solution in addressing challenges where many factors must be taken into account, like supply chain management. For example, Q-learning is a type of RL algorithm that is applied to tackle simple optimization problems. In this approach, the Q-value has been applied to any state of the system. Although the classical RL algorithms guarantee optimal policy, these algorithms cannot promptly solve large states or actions. Many problems in the real world have large and action spaces. Applying RL algorithms for solving large problems would be nearly impossible, as these models would be costly to train. Therefore, deep RL emerges as a new method in which DNN is used to approximate any of the following RL components. Recently, deep Q-network (DQN) algorithms have been used in different areas. For example, deep Q-network (DQN) algorithms have been applied to solve supply chain optimization problems. These DQNs operate as the decision-maker of each agent. That results in a competitive game in which each DQN agent plays independently to minimize its own cost. Instead, recently a unified

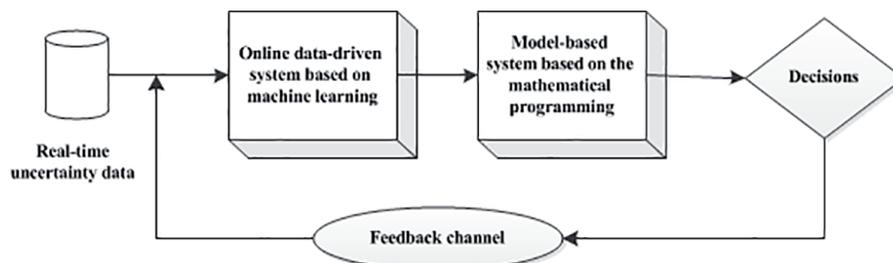


Figure 2. The schematic of the “closed-loop” online learning-based data-driven optimization framework.

framework has been proposed in which the agents still play independently from one another. Still, in the training phase, this model uses a feedback scheme so that the DQN agent learns the total cost for the whole network and, over time, learns to minimize it.

Like other types of reinforcement ML technique, multi-agent RL is a system of agents (e.g., robots, machines, and cars) interacting within a common environment. Each agent decides each time-step and works along with the other agent(s) to achieve a given goal. The agents are learnable units that want to learn policy on the fly to maximize the long-term reward through the interaction with the environment. Recently the multi-agent RL techniques have been applied to develop the supply chain management (SCM) systems that perform optimally for each entity in the chain. A supply chain can be defined as a network of autonomous business entities collectively responsible for procurement, manufacturing, storing, and distribution [65]. Entities in a supply chain have different sets of environmental constraints and objectives.

One of the biggest challenges of the development of MAS based supply chain is designing agent policies. To address designing agent policies, recently, automatic policy designing by RL has drawn attention. RL is considered an optimal solution in addressing challenges where a huge number of factors must be taken into account, like SCM. RL technique does not require datasets covering all environments, constraints, operations, and entity operation results. A multi-agent RL (MARL)-based SCM system can enable agents to learn automatically policies that optimize the supply chain performance using RL concerning certain constraints, environments, and objectives to optimize the performance. More specifically, the RL technique enables an agent to learn a policy by correcting necessary data itself during trial-and-error on the content of operations [66]. All agents also simultaneously cooperate to optimize the performances of the entire supply chain. RL technique can be applied for a certain problem when all processes concerning the problem satisfy a Markov property. Environmental change for a certain agent depends on the previous state of the environment and the agent's action. It is impossible to assume the Markov property because an agent's environmental change depends on the previous state for the agent and the other agent's actions.

There are two problems in developing a MARL technique for SCM: Building Markov decision processes for a supply chain and then avoiding learning stagnation among agents in learning processes. For solving these problems, a learning management method with deep neural network (DNN)-weight evolution (LM-DWE) has been applied [67]. Fuji et al. [67] developed a multi-agent RL technique to develop a supply chain management (SCM) system that enables agents to learn policies that optimize SC performance. They applied a learning management method with deep-neural-network (DNN)-weight evolution (LM-DWE) in the MARL for SCM. An RL framework-FeedRec has been used in a study to optimize long-term user engagement [68]. They used hierarchical LSTM to design the Q-Network to model the complex user behaviors; they also used Q Network to simulate the environment. Zhang et al. [69] proposed a multi-agent learning (MAL) algorithm and applied it for optimizing online resource allocation in cluster networks.

4.8 Deep RL for solving NP-hard problems

Optimization in current DSS has a highly interdisciplinary nature related to integrating different techniques and paradigms for solving complex real-world problems. The design of efficient NP-hard combinatorial optimization problems is a

fascinating issue and often requires significant specialized knowledge and trial-and-error. NP-hard problems are solved with exact methods, heuristic algorithms, or a combination of them. Although exact methods provide optimal answers, they have the limitation of performing inefficiently in time complexity. Heuristics are used to improve computational time efficiency and provide decent or near-optimal solutions [70]. According to the definition of Burke et al. [71], a hyper-heuristic is a searching mechanism that aims to select or generate appropriate heuristics to solve an optimization problem. However, the effectiveness of general heuristic algorithms is dependent on the problem being considered, and high levels of performance often require extensive tailoring and domain-specific knowledge. ML strategies have become a promising route to addressing these challenges, which led to the development of meta-algorithms to various combinatorial problems.

Solution approaches meta-heuristics and hyper-heuristics have been developed to tackle the NP-hard combinatorial optimization problem [72]. Recently, hyper-heuristics arise in this context as efficient methodologies for selecting or generating (meta) heuristics to solve NP-hard optimization problems. Hyper-heuristics are categorized into heuristic selection (Methodologies to select) and heuristic generation (Methodologies to generate) [71]. Deep RL is a possible learning method that can automatically solve various optimization problems [73]. Encouragingly, characteristics of the deep RL method have been found in comparison with classical methods, e.g., strong generalization ability and fast solving speed. RL methods can be used at different levels to solve combinatorial optimization problems. They can be applied directly to the problem, as part of a meta-heuristic, or as part of hyper-heuristics [74]. Utilizing advanced computation power with meta-heuristics algorithms and massive-data processing techniques has successfully solved various NP-hard problems. However, meta-heuristic approaches find good solutions which, do not guarantee the determination of the global optimum. Meta-heuristics still face the limitations of exploitation and exploration, which consists of choosing between a greedy search and a wider exploration of the solution space.

A way to guide Meta-heuristic algorithms during the search for better solutions is to generate the initial population of a genetic algorithm by using a technique of Q-Learning algorithm.

The hyper-heuristic for heuristic selection can use RL algorithms, enabling the system to autonomously select the meta-heuristic to use in the optimization process and the respective parameters. For example, Falcão et al. [74] proposed a hyper-heuristic module for solving scheduling problems in manufacturing systems. The proposed hyper-heuristic module uses an RL algorithm, which enables the system to autonomously select the meta-heuristic to use in the optimization process and the respective parameters. Cano-Belmán et al. [75] proposed a heuristic generation scatter search algorithm to address a mixed-model assembly line sequencing problem. Khalil et al. (Dai et al., 2017) developed a neural combinatorial optimization framework that utilizes neural networks and RL to tackle combinatorial optimization problems. The developed meta-algorithm automatically learns good heuristics for a diverse range of optimization problems over graphs. Mosadegh et al. [72] proposed novel hyper-simulated annealing (HSA) to tackle the NP-hard problem. They developed new mathematical models to describe a mixed-model sequencing problem with stochastic processing times (MMSPSP). The HSA applies a Q-learning algorithm to select appropriate heuristics through its search process [72]. The main idea is to conduct simulated annealing (SA)-based algorithms to find a suitable heuristic among available ones creating a neighbor solution(s).

Case study 1: Data-driven robust optimization under correlated uncertainty.

The first case study focuses on the production schedule. The data-driven robust optimization applied for an ethylene plant is predicted to hedge against the fluctuations generated from continuous production processes. For capturing and enrich the valid information of uncertainties, copulas are introduced to estimate the joint probability distribution and simulate mutual scenarios for uncertainties [43]. For this purpose, cutting planes are generated to remove unnecessary uncertain scenarios in the uncertainty sets. Then robust formulations induced by the cut set are proposed to reduce conservatism and improve the robustness of scheduling solutions. They consider the robust counterpart induced by the classical uncertainty set, where the difference to the best possible solution over all scenarios is to be minimized. Instead of focuses on simple uncertainty sets that are either finite or hyperboles, they considered problems with more flexible and realistic ellipsoidal uncertainty sets. In this research, the cut sets of flexible uncertainty sets are proposed. They used the historical data to correct the uncertainties and drive the reformulation of constraints with uncertainties. The new robust formulations induced by cut sets are derived for linear programming (LP) and mixed-integer linear programming (MILP) problems. Through the real-world ethylene plant example, the correlations between uncertain consumption rates of furnaces are analyzed.

In this research, Decision-makers prefer to obtain robust solutions immune to most high-frequency uncertain scenarios. Since in production scheduling problems, many uncertainties are associated with the entire production network, a process, or equipment, which makes them correlated and difficult to be separated. So, in this optimization research, uncertainties are assumed to be dependent. In this research, the cut sets of flexible uncertainty sets are proposed.

Deterministic solutions are regarded as theoretically optimal at most times, and robust solutions provide references for decision-makers, which may not be optimal but feasible and applicable. It is always neglected that stricter descriptions of uncertainties could also create great profits. The full coverage of uncertain values usually leads to unpractical and conservative results. The improper simplification of uncertainty scenarios will cause infeasibility when the solutions are implemented in the volatile production process. Thus, historical data should be introduced to correct the uncertainties and drive the reformulation of constraints with uncertainties. For eliminating the worst-case formulation scenario for robust optimization and decrease conservatism, the cut set of flexible uncertainty sets is constructed by introducing cutting planes. Cutting planes are generated to construct cut sets for the outer approximation of most uncertain scenarios. Since the size of the uncertainty set directly influences the quality of robust solutions, in this research, the more uncertain values are considered.

They stated that utilizing the data-driven robust optimization approach causes the decision-makers to have the ability to decide how many uncertain scenarios are considered in the model and to provide effective, economical, and robust scheduling plans. Finally, it causes fluctuations in the production performance captured and controlled below a lower level of conservatism.

Case study 2: wastewater sludge-to-biodiesel supply chain design.

Designing and optimizing the wastewater sludge-to-biodiesel supply chain facilitates the development of its large-scale production [42]. Hence, this case study evaluates Data-driven robust optimization for supply chain designing and optimization. The entire wastewater sludge-to-biodiesel supply chain over multiple periods is systematically designed and optimized based on the uncertainty sets constructed from

the data of uncertain parameters. In this research, a data-driven robust optimization has been adopted, which constructs the uncertainty sets from the data of uncertain parameters utilizing support vector clustering. In contrast, the conventional uncertainty sets are driven without incorporating the data, which results in a high cost of robustness. The developed uncertainty set in this research encloses the fuzzy support neighborhood of data samples that makes it practical even when the available data is limited. The research results show that the proposed data-driven robust optimization approach can yield robust supply chain decisions with the same degree of robustness but at a lower cost than robust conventional optimization approaches.

Case study 3: Forecasting fluctuating variation in electricity demand and generation.

Our third case study relates to forecasting fluctuating electricity demand and generation variation, aiming to develop an energy forecasting model with renewable energy technologies [54, 55]. Wind and solar energy sources are erratic and difficult to implement in renewable energy systems; therefore, circumspection is needed to implement renewable energy systems and policies. This translates into the DL-based models for forecasting fluctuating electricity demand and generation in renewable energy systems.

This study compares and evaluates DL models and conventional statistical models. The DL models include DNN, long short-term memory, gated recurrent unit, and the disadvantages of conventional statistical models such as multiple linear regression and seasonal autoregressive integrated moving average. Thus, they thoroughly compare and evaluate the forecasting models and select the best forecasting model for future electricity demand and renewable energy generation. They then utilized the proposed model for renewable energy scenarios for Jeju Island's policy design to achieve their energy policy. The optimal scenario is assessed by considering its strengths, weaknesses, opportunities, and threats analysis while also considering techno-economic-environmental domestic and global energy circumstances.

5. Conclusion and managerial implications

Data-driven optimization refers to the art and science of integrating the data-driven system based on ML to convert (big) data into relevant and useful information and insights, and the model-based system based on mathematical programming to derive the optimal and more accurate decisions from the information. As a direct implication, the generic approach proposed in data-driven optimization can be utilized to create an automated, data-driven, and intelligent DSS, which would increase the quality of decisions both in terms of efficiency and effectiveness. Recent advances in DL as a predictive model have received great attention lately. One of the distinguishing features of DNN is its ability to “learn” better predictions from large-scale data than ML methods. Hence, one of the primary messages of this overview chapter is to review the applicability of DL in improving DSS across core areas of supply chain operations.

Much data is generated at ever-faster rates by companies and organizations [76]. Applying the advanced DL techniques for predictive analytics becomes a promising issue for further research to improve the decision-making process. Although the conventional data-driven optimization paradigm has made significant progress for hedging against uncertainty, it is foreseeable that data-driven mathematical programming frameworks would proliferate in the next few years due to the generation

of large volumes of data and the complexity of relationships among elements. Nowadays, the increase in data acquisition and availability and the emergence of DL makes it imperative to develop data-driven mathematical programming to approximate complex systems under uncertainty. More specifically, a deep data-driven model paradigm, in which the rigorous mathematical model is developed based on neural networks to modeling complex systems and optimizing their operations, could be a promising research direction.

Furthermore, there are some research challenges associated with conventional data-driven optimization frameworks. For example, updating the data-driven system and learning based on real-time data in the data-driven model frameworks can be a key research challenge. Future research could be directed toward designing the data-driven system, in which learning takes place sequentially to extract useful and relevant information from real-time uncertainty data. The data-driven systems should be updated in an online fashion.

Developing the mathematical programming problems for an online-learning-based data-driven optimization paradigm creates another challenge. The model-based system can be devised based on the deep data-driven model paradigm and be leveraged the power of DL. Additionally, deep RL can be applied to developing efficient algorithms to solve the resulting online-learning-based mathematical programming problems. Applying deep RL in the paradigm of learning-while-optimizing also could be another promising research direction. Besides, multi-agent RL techniques could be explored by taking advantage of DL to develop complex systems and optimize their performance based on real-time data.

Also, RL is another ML area that has recently been used to model complex systems and problems and to optimize their performance and behaviors. RL is also considered an optimal solution in addressing challenges where many factors must be taken into account. More specifically, deep RL emerges as a new method to solve the various optimization problems automatically. Thereby, applying RL in optimization problems deserves further attention in future research.

IntechOpen

IntechOpen

Author details

Saeid Sadeghi¹, Maghsoud Amiri² and Farzaneh Mansoori Mooseloo^{3*}

1 Faculty of Management and Accounting, Department of Industrial Management, University of Tehran, Tehran, Iran

2 Faculty of Management and Accounting, Department of Industrial Management, Allameh Tabataba'i University, Tehran, Iran

3 Faculty of Management, Department of Industrial Management, University of Hormozgan, Bandar-Abbas, Iran

*Address all correspondence to: farzanmansoori7@gmail.com

IntechOpen

© 2021 The Author(s). Licensee IntechOpen. This chapter is distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/3.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. 

References

- [1] Biegler, L. T., & Grossmann, I. E. (2004). Retrospective on optimization. *Computers & Chemical Engineering*, 28(8), 1169-1192.
- [2] Sakizlis, V., Perkins, J. D., & Pistikopoulos, E. N. (2004). Recent advances in optimization-based simultaneous process and control design. *Computers & Chemical Engineering*, 28(10), 2069-2086.
- [3] Sahinidis, N. V. (2004). Optimization under uncertainty: state-of-the-art and opportunities. *Computers & Chemical Engineering*, 28(6-7), 971-983.
- [4] Darvazeh, S. S., Vanani, I. R., & Musolu, F. M. (2020). Big data analytics and its applications in supply chain management. In *New Trends in the Use of Artificial Intelligence for the Industry 4.0* (p. 175). IntechOpen.
- [5] Bertsimas, D., Gupta, V., & Kallus, N. (2018a). Data-driven robust optimization. *Mathematical Programming*, 167(2), 235-292.
- [6] Bertsimas, D., Gupta, V., & Kallus, N. (2018b). Data-driven robust optimization. *Mathematical Programming*, 167(2), 235-292.
- [7] Grossmann, I. E., Apap, R. M., Calfa, B. A., García-Herreros, P., & Zhang, Q. (2016). Recent advances in mathematical programming techniques for the optimization of process systems under uncertainty. *Computers & Chemical Engineering*, 91, 3-14.
- [8] Birge, J. R., & Louveaux, F. (2011). *Introduction to stochastic programming*. Springer Science & Business Media.
- [9] Nikzad, E., Bashiri, M., & Oliveira, F. (2019). Two-stage stochastic programming approach for the medical drug inventory routing problem under uncertainty. *Computers & Industrial Engineering*, 128, 358-370.
- [10] Quddus, M. A., Chowdhury, S., Marufuzzaman, M., Yu, F., & Bian, L. (2018). A two-stage chance-constrained stochastic programming model for a bio-fuel supply chain network. *International Journal of Production Economics*, 195, 27-44.
- [11] Mavromatidis, G., Orehounig, K., & Carmeliet, J. (2018). Design of distributed energy systems under uncertainty: A two-stage stochastic programming approach. *Applied energy*, 222, 932-950.
- [12] Lima, C., Relvas, S., & Barbosa-Póvoa, A. (2018). Stochastic programming approach for the optimal tactical planning of the downstream oil supply chain. *Computers & Chemical Engineering*, 108, 314-336.
- [13] Alipour, M., Zare, K., & Seyed, H. (2018). A multi-follower bilevel stochastic programming approach for energy management of combined heat and power micro-grids. *Energy*, 149, 135-146.
- [14] Ben-Tal, A., El Ghaoui, L., & Nemirovski, A. (2009). *Robust optimization*. Princeton university press.
- [15] Kim, J., Do Chung, B., Kang, Y., & Jeong, B. (2018). Robust optimization model for closed-loop supply chain planning under reverse logistics flow and demand uncertainty. *Journal of cleaner production*, 196, 1314-1328.
- [16] Aalaei, A., & Davoudpour, H. (2017). A robust optimization model for cellular

manufacturing system into supply chain management. *International Journal of Production Economics*, 183, 667-679.

[17] Lim, Y. F., & Wang, C. (2017). Inventory management based on target-oriented robust optimization. *Management Science*, 63(12), 4409-4427.

[18] Vitus, M. P., Zhou, Z., & Tomlin, C. J. (2015). Stochastic control with uncertain parameters via chance constrained control. *IEEE Transactions on Automatic Control*, 61(10), 2892-2905.

[19] Farina, M., Giulioni, L., & Scattolini, R. (2016). Stochastic linear model predictive control with chance constraints—a review. *Journal of Process Control*, 44, 53-67.

[20] Guo, Y., Baker, K., Dall'Anese, E., Hu, Z., & Summers, T. H. (2018). Data-based distributionally robust stochastic optimal power flow—Part I: Methodologies. *IEEE Transactions on Power Systems*, 34(2), 1483-1492.

[21] Carvalho, A., Lefèvre, S., Schildbach, G., Kong, J., & Borrelli, F. (2015). Automated driving: The role of forecasts and uncertainty—A control perspective. *European Journal of Control*, 24, 14-32.

[22] Russell, S., & Norvig, P. (2002). Artificial intelligence: a modern approach.

[23] Ngiam, K. Y., & Khor, W. (2019). Big data and machine learning algorithms for healthcare delivery. *The Lancet Oncology*, 20(5), e262-e273.

[24] Helm, J. M., Swiergosz, A. M., Haeberle, H. S., Karnuta, J. M., Schaffer, J. L., Krebs, V. E., ... & Ramkumar, P. N. (2020). Machine learning and artificial intelligence: Definitions, applications, and future directions. *Current reviews in musculoskeletal medicine*, 13(1), 69-76.

[25] Jakhar, D., & Kaur, I. (2020). Artificial intelligence, machine learning and deep learning: definitions and differences. *Clinical and experimental dermatology*, 45(1), 131-132.

[26] Goodfellow, I., Bengio, Y., Courville, A., & Bengio, Y. (2016). *Deep learning* (Vol. 1, No. 2). Cambridge: MIT press.

[27] Wang, H., Wu, Y., Min, G., Xu, J., & Tang, P. (2019). Data-driven dynamic resource scheduling for network slicing: A deep reinforcement learning approach. *Information Sciences*, 498, 106-116.

[28] Ning, C., & You, F. (2019). Optimization under uncertainty in the era of big data and deep learning: When machine learning meets mathematical programming. *Computers & Chemical Engineering*, 125, 434-448.

[29] Wong, A. K. C., & Wang, Y. (2003). Pattern discovery: a data driven approach to decision support. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, 33(1), 114-124.

[30] Yang, H., Jin, Z., Wang, J., Zhao, Y., Wang, H., & Xiao, W. (2019). Data-Driven Stochastic Scheduling for Energy Integrated Systems. *Energies*, 12(12), 2317.

[31] Esfahani, P. M., & Kuhn, D. (2018). Data-driven distributionally robust optimization using the Wasserstein metric: Performance guarantees and tractable reformulations. *Mathematical Programming*, 171(1), 115-166.

[32] Smith, J. E., & Winkler, R. L. (2006). The optimizer's curse: Skepticism and postdecision surprise in decision analysis. *Management Science*, 52(3), 311-322.

- [33] Calafiore, G. C., & El Ghaoui, L. (2006). On distributionally robust chance-constrained linear programs. *Journal of Optimization Theory and Applications*, 130(1), 1-22.
- [34] Hu, Z., & Hong, L. J. (2013). Kullback-Leibler divergence constrained distributionally robust optimization. Available at *Optimization Online*. Hyperheuristics: A survey of the state of the art. *Journal of the Operational Research Society*, 64(12), 1695-1724.
- [35] Wang, C., & Chen, S. (2020). A distributionally robust optimization for blood supply network considering disasters. *Transportation Research Part E: Logistics and Transportation Review*, 134, 101840.
- [36] Chiou, S. W. (2020). Data-Driven Stochastic Optimization for Transportation Road Network Design Under Uncertainty. In *Handbook of Research on Big Data Clustering and Machine Learning* (pp. 231-278). IGI Global.
- [37] Gao, J., Ning, C., & You, F. (2019). Data-driven distributionally robust optimization of shale gas supply chains under uncertainty. *AIChE Journal*, 65(3), 947-963.
- [38] Shang, C., Huang, X., & You, F. (2017). Data-driven robust optimization based on kernel learning. *Computers & Chemical Engineering*, 106, 464-479.
- [39] Shen, W., Li, Z., Huang, B., & Jan, N. M. (2018). Chance-constrained model predictive control for SAGD process using robust optimization approximation. *Industrial & Engineering Chemistry Research*, 58(26), 11407-11418.
- [40] Ning, C., & You, F. (2017). Data-driven adaptive nested robust optimization: general modeling framework and efficient computational algorithm for decision making under uncertainty. *AIChE Journal*, 63(9), 3790-3817.
- [41] Ning, C., & You, F. (2018). Data-driven decision making under uncertainty integrating robust optimization with principal component analysis and kernel smoothing methods. *Computers & Chemical Engineering*, 112, 190-210.
- [42] Mohseni, S., & Pishvaei, M. S. (2020). Data-driven robust optimization for wastewater sludge-to-biodiesel supply chain design. *Computers & Industrial Engineering*, 139, 105944.
- [43] Zhang, Y., Jin, X., Feng, Y., & Rong, G. (2018). Data-driven robust optimization under correlated uncertainty: a case study of production scheduling in ethylene plant. *Computers & Chemical Engineering*, 109, 48-67.
- [44] Erdoğ an, E., & Iyengar, G. (2006). Ambiguous chance constrained problems and robust optimization. *Mathematical Programming*, 107(1), 37-61.
- [45] Jiang, R., & Guan, Y. (2016). Data-driven chance constrained stochastic program. *Mathematical Programming*, 158(1), 291-327.
- [46] Calfa, B. A., Grossmann, I. E., Agarwal, A., Bury, S. J., & Wassick, J. M. (2015). Data-driven individual and joint chance-constrained optimization via kernel smoothing. *Computers & Chemical Engineering*, 78, 51-69.
- [47] Ji, R., & Lejeune, M. A. (2021). Data-driven distributionally robust chance-constrained optimization with Wasserstein metric. *Journal of Global Optimization*, 79(4), 779-811.
- [48] Ghosal & Wiesemann, W. (2018). Data-driven chance constrained

programs over Wasserstein balls. *arXiv preprint arXiv:1809.00210*.

[49] Khan, P. W., Byun, Y. C., & Park, N. (2020). IoT-Blockchain Enabled Optimized Provenance System for Food Industry 4.0 Using Advanced Deep Learning. *Sensors*, 20(10), 2990.

[50] Kraus, M., Feuerriegel, S., & Oztekin, A. (2020). Deep learning in business analytics and operations research: Models, applications and managerial implications. *European Journal of Operational Research*, 281(3), 628-641.

[51] Kilimci, Z. H., Akyuz, A. O., Uysal, M., Akyokus, S., Uysal, M. O., Atak Bulbul, B., & Ekmiş, M. A. (2019). An improved demand forecasting model using deep learning approach and proposed decision integration strategy for supply chain. *Complexity*, 2019.

[52] Weng, Y., Wang, X., Hua, J., Wang, H., Kang, M., & Wang, F. Y. (2019). Forecasting horticultural products price using ARIMA model and neural network based on a large-scale data set collected by web crawler. *IEEE Transactions on Computational Social Systems*, 6(3), 547-553.

[53] Yu, L., Zhao, Y., Tang, L., & Yang, Z. (2019). Online big data-driven oil consumption forecasting with Google trends. *International Journal of Forecasting*, 35(1), 213-223.

[54] Nam, K., Hwangbo, S., & Yoo, C. (2020a). A deep learning-based forecasting model for renewable energy scenarios to guide sustainable energy policy: A case study of Korea. *Renewable and Sustainable Energy Reviews*, 122, 109725.

[55] Nam, K., Hwangbo, S., & Yoo, C. (2020b). A deep learning-based

forecasting model for renewable energy scenarios to guide sustainable energy policy: A case study of Korea. *Renewable and Sustainable Energy Reviews*, 122, 109725.

[56] Li, Q., Loy-Benitez, J., Nam, K., Hwangbo, S., Rashidi, J., & Yoo, C. (2019). Sustainable and reliable design of reverse osmosis desalination with hybrid renewable energy systems through supply chain forecasting using recurrent neural networks. *Energy*, 178, 277-292.

[57] Loy-Benitez, J., Vilela, P., Li, Q., & Yoo, C. (2019). Sequential prediction of quantitative health risk assessment for the fine particulate matter in an underground facility using deep recurrent neural networks. *Ecotoxicology and environmental safety*, 169, 316-324.

[58] Zhu, W., Chebeir, J., & Romagnoli, J. A. (2020). Operation optimization of a cryogenic NGL recovery unit using deep learning based surrogate modeling. *Computers & Chemical Engineering*, 137, 106815.

[59] Gupta, V., & Rusmevichientong, P. (2017). Small-data, large-scale linear optimization with uncertain objectives. *Management Science*, 67(1), 220-241.

[60] Katz, J., Pappas, I., Avraamidou, S., & Pistikopoulos, E. N. (2020). Integrating deep learning models and multiparametric programming. *Computers & Chemical Engineering*, 136, 106801.

[61] Pfrommer, J., Zimmerling, C., Liu, J., Kärger, L., Henning, F., & Beyerer, J. (2018). Optimisation of manufacturing process parameters using deep neural networks as surrogate models. *Procedia CiRP*, 72, 426-431.

[62] Marino, D. L., & Manic, M. (2019). Modeling and planning under uncertainty

using deep neural networks. *IEEE Transactions on Industrial Informatics*, 15(8), 4442-4454.

[63] del Rio-Chanona, E. A., Wagner, J. L., Ali, H., Fiorelli, F., Zhang, D., & Hellgardt, K. (2019). Deep learning-based surrogate modeling and optimization for microalgal biofuel production and photobioreactor design. *AIChE Journal*, 65(3), 915-923.

[64] Tang, Z., & Zhang, Z. (2019). The multi-objective optimization of combustion system operations based on deep data-driven models. *Energy*, 182, 37-47.

[65] Swaminathan, J. M., Smith, S. F., & Sadeh, N. M. (1998). Modeling supply chain dynamics: A multiagent approach. *Decision sciences*, 29(3), 607-632.

[66] Mnih, V., Kavukcuoglu, K., Silver, D., Rusu, A. A., Veness, J., Bellemare, M. G., ... & Hassabis, D. (2015). Human-level control through deep reinforcement learning. *nature*, 518(7540), 529-533.

[67] Fuji, T., Ito, K., Matsumoto, K., & Yano, K. (2018, January). Deep multi-agent reinforcement learning using dnn-weight evolution to optimize supply chain performance. In *Proceedings of the 51st Hawaii International Conference on System Sciences*.

[68] Zou, L., Xia, L., Ding, Z., Song, J., Liu, W., & Yin, D. (2019, July). Reinforcement learning to optimize long-term user engagement in recommender systems. In *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining* (pp. 2810-2818).

[69] Zhang, C., Lesser, V. R., & Shenoy, P. J. (2009, July). A Multi-Agent Learning Approach to Online Distributed Resource Allocation. In *Ijcai* (Vol. 9, pp. 361-366).

[70] Dumitrescu, I., & Stützle, T. (2003, April). Combinations of local search and exact algorithms. In *Workshops on Applications of Evolutionary Computation* (pp. 211-223). Springer, Berlin, Heidelberg.

[71] Burke, E. K., Gendreau, M., Hyde, M., Kendall, G., Ochoa, G., Özcan, E., & Qu, R. (2013).

[72] Mosadegh, H., Ghomi, S. F., & Süer, G. A. (2020). Stochastic mixed-model assembly line sequencing problem: Mathematical modeling and Q-learning based simulated annealing hyper-heuristics. *European Journal of Operational Research*, 282(2), 530-544.

[73] Li, K., Zhang, T., & Wang, R. (2020). Deep reinforcement learning for multi-objective optimization. *IEEE transactions on cybernetics*.

[74] Falcão, D., Madureira, A., & Pereira, I. (2015, June). Q-learning based hyper-heuristic for scheduling system self-parameterization. In *2015 10th Iberian Conference on Information Systems and Technologies (CISTI)* (pp. 1-7). IEEE.

[75] Cano-Belmán, J., Ríos-Mercado, R. Z., & Bautista, J. (2010). A scatter search based hyper-heuristic for sequencing a mixed-model assembly line. *Journal of Heuristics*, 16(6), 749-770.

[76] Corbett, C. J. (2018). How sustainable is big data?. *Production and Operations Management*, 27(9), 1685-1695.