Review of Data-Driven Robust Optimization

Diantiny Mariam Pribadi
Master Program of Mathematics, Faculty of Mathematics and Natural Sciences,
Universitas Padjadjaran, Indonesia
Email: diantiny152@gmail.com

Sukono, Riaman
Department of Mathematics, Faculty of Mathematics and Natural Sciences,
Universitas Padjadjaran, Indonesia
Email: sukono@unpad.ac.id; riaman@unpad.ac.id

Kalfin
Doctor Program of Mathematics, Faculty of Mathematics and Natural Sciences, Universitas
Padjadjaran, Indonesia
kalfin17001@mail.unpad.ac.id

Abdul Talib Bon
Department of Production and Operations, University Tun Hussein Onn Malaysia, Malaysia
talibon@gmail.com

Abstract

The optimization model generally assumes complete data is known. But in reality, a lot of data is not known for certain. This uncertainty problem can be solved by several approaches, one of which is robust optimization. Uncertainty parameters in optimizing robust are solved by using the set of uncertainties. However, the set of uncertainties yields less conservative results to be applied to the data as a whole. With the abundance of data in recent years, the determination of the set of uncertainties can be done based on data, this method is called Data-Driven Robust Optimization (DDRO). Robust optimization based on data with a machine learning approach presents new challenges. This paper reviews several papers on DDRO and their applications on inventory, scheduling, portfolio selection, industries, and transportation issues.

Keywords:
Data-driven optimization, robust optimization, uncertainty set.

1. Introduction

In general, researchers focus on models with complete known information where opportunities are known, limited, and discrete (Bertsimas et al, 2018). But in reality, only historical data are available, while the distribution is not known with certainty (Klabjan et al, 2013). This causes many optimization problems that meet uncertain conditions. The solution to this problem is finding the constraint randomization with the set of uncertainties (Campi & Garatti, 2008). The solution to the model with uncertainty is to estimate the distribution of historical data parameters. In the last decade, many industries began to make data-based decisions. The development of data science has been a development in recent years. This causes data availability to be abundant. Data has become an increasingly valuable asset in various fields. (Ning & You, 2016). Many researchers use available historical data as a practical approach to determining the characteristics of information distribution (Qiu et al. 2019).

This abundance of data makes it possible to utilize machine learning in estimating data density. Likewise in building a set of robust optimization uncertainty with machine learning becomes a new challenge (Qiu et al. 2019). Uncertainty and availability of data influence researchers to determine the set of uncertainties based on data. Campi & Garatti (2008) set a new feasible solution for constraint randomization with a data-driven method for robust optimization. Wang et al. (2016) propose the problem of Likelihood robust optimization in an environment of uncertainty based on data with unknown input distribution. Bertsimas et al (2018) proposed a data set uncertainty set for robust optimization using statistical hypothesis testing. Data-driven robust optimization is significantly superior to traditional robust optimization.

Besides, some researchers are applying hypothesis testing in data-driven optimization. Goldfarb & Iyengar (2003) examined the set of uncertainties for the mean and covariance of distributions using linear regression and $t-$test on portfolio selection problems to resolve optimal solution sensitivity and estimating market parameters. Chi et al. (2019) use data-driven problems with Peer-to-Peer (P2P) lending investment.

Klabjan et al. (2013) proposed a minimax robust model with distributionally robust dynamic programs based on the Pearson χ^2-test for inventory problems based on historical demand data. They prove that the optimal policy inventory problem from the robust model is the same as the stochastic dynamic programming counterpart. If the demand distribution is known, the robust model converges to the stochastic model. Qiu et al. (2019) also apply data-driven robust optimization to the problem of multi-product inventory. They use Support Vector Clustering (SVC) to build a set of uncertainties based on data. Robust counterpart is developed by using absolute robustness criteria into a linear programming model. Zhao & You (2018) discusses the supply chain under uncertain production capacity using robust optimization.

Data-driven robust optimization has also been applied to industries, including steam systems, biodiesel and oil production. Zhao et al. (2019) apply data-driven robust optimization to the steam system. The uncertain parameters of the turbine model are derived from the semiempirical model and historical data. The set of uncertainties is determined by the estimated kernel density. In another study, Zhao et al. (2019) also implemented a data-driven robust optimization of the steam system in ethylene plants. Mohseni et al. (2019) proposed a linear integer programming model for biodiesel production using liquid waste sludge. They overcome the parameter uncertainty by supporting Support Vector Clustering (SVC). Dai et al. (2020) resolve the uncertainties in mixing crude oil using data-driven optimization. In his research, the blending effect model is used to extract the uncertainty of oil components from production data using the recursive least square method. The set of uncertainties was built by combining Principle Component Analysis and Robust Kernel Density Estimation based on historical data. Zhang et al. (2018) apply data-driven to the uncertainty of environmental and operational conditions in the process industry so that the robust optimization model adaptation is obtained. This study reviews several previous papers regarding the determination of the set of uncertainties based on data in Robust Optimization (DDRO) and their applications in the fields of inventory, scheduling, transportation, investment portfolio, and industry.

2. Robust Optimization

Robust optimization is one of the optimization methods used to deal with uncertainty. When the parameter is only known to have a certain interval with a certain level of confidence and the value covers a certain range of variations, then the robust optimization approach can be used. The purpose of robust optimization is to find a robust solution for all possible and optimal data realizations.

There are several approaches to resolve uncertainty directly including stochastic programming, chance-constrained programming, fuzzy programming, and robust optimization. Stochastic programming can be used to deal with uncertainty, but the solution is very sensitive to a predetermined probability distribution so that it can deviate significantly (Zhao et al, 2019). In robust optimization, the model parameters are not necessarily solved by determining the set of uncertainties, then optimizing them against the worst-case set. The selection of the right set of uncertainties will result in the best solution to the optimal problem (Bertsimas et al, 2018). The Robust optimization approach gives very conservative results.

The robust optimization approach can be converted from linear programming problems with uncertainty to deterministic robust counterpart problems. Robust Counterpart must be computationally tractable. Robust Counterpart is a semi-infinite optimization problem. This problem cannot be solved numerically efficiently. In robust optimization, it is important to turn the robust counterpart for generic convex problems into explicit convex optimization problems. The set of uncertainties can be obtained by reformulating the problem into a tractable problem. There are several sets of uncertainties, including intervals, ellipsoidal, and polyhedral (Ben-Tal et al, 2000).

The set of uncertainties in robust optimization by following mild assumptions about uncertainty is explained by several researchers, including Ben-Tal & Nemirovski (2000), Bertsimas & Sim (2004), Ben-Tal et al. (2009), Bandi & Bertsimas (2012), Chan et al. (2010). The approach to determining the set of uncertainties fulfills two key properties, namely robust computationally tractable constraints and the set of uncertainties implying probability guarantees to ensure that feasible solutions to robust constraints will always be feasible.

3. Data-Driven Robust Optimization

Based on the development of time, the set of L1, L2, infinite norm, ellipsoidal, and polyhedral uncertainty set is considered as the set of classical indeterminacy (Zhang et al, 2018). The Data-Driven Robust Optimization (DDRO) method integrates big data with robust optimization. Several studies have been carried out for robust methods based on data optimization. Furthermore, the set of uncertainties based on data is determined based on the probability distribution and quantile values. This set of uncertainties has been proven to reduce the conservatism of robust solutions. Data-driven robust optimization is done either by hypothesis testing or not. Campi & Garatti (2008) proposed a data-driven method for robust optimization not based on hypothesis testing. While the application of hypothesis testing in robust optimization is carried out by Klabjan et al. (2013), Goldfarb & Iyengar (2003), and Bertsimas et al. (2018).

Goldfarb & Iyengar (2003) propose an alternative deterministic model that is robust to parameter uncertainty and error estimation. They use multivariate linear regression to justify the structure of uncertainty. Bertsimas et al. (2018) made several contributions to data-driven robust optimization, including schemes to build a set of uncertainties from data using hypothesis testing, so that the robust optimization problem that is generally tractable to the set is obtained. Data-driven methods using hypothesis testing motivate the use of statistical numerical techniques and propose new approaches to model several uncertain constraints simultaneously to obtain optimal solutions. Besides, they also implemented data-driven robust optimization in the queue and portfolio allocation.

In general, the geometric characteristics of probability guarantees occur at ε level only if $\mathrm{P}^*(\overline{\mathbf{u}} \in \mathrm{U}) \geq 1-\varepsilon$. However, when P is unknown, the region of P trust with hypothesis testing will contain P with a probability of more than $1-\alpha$. Bertsimas et al (2018) designed several sets of uncertainties, including the set of uncertainties from discrete distributions, Kolmogorov-Smirnov test, Forward and Backward deviation, marginal samples and non-independent potential components.

The set of uncertainties from the discrete distribution P^* is assumed to be known and limited. Bertsimas et al (2018) consider two hypothesis tests, namely the Pearson X^2 test and the G. test both of these tests use the $H_0 : \mathrm{P}^* = \mathrm{P}_0$ hypothesis where P_0 is some specified measure. The next set of uncertainties assumes continuous P^*, but the marginal distribution is known and mutually independent, among the set of uncertainties built from the Kolmogorov-Smirnov Test, motivated from forward and backward deviation, marginal samples, and potential non-independent components.

The set of uncertainties from the Kolmogorov-Smirnov test assumes a $s\mathrm{upp}\left(\mathrm{P}^*\right)$ known and limited, $\left[\hat{\mathbf{u}}^{(0)}, \hat{\mathbf{u}}^{(N+1)}\right] \equiv \left\{\mathbf{u} \in \mathrm{R}^d \middle| \hat{u}_i^{(0)} \leq u_i \leq \hat{u}_i^{(N+1)}, i = 1,\ldots,d\right\}$ Both are given the univariate $\mathrm{P}_{0,t}$ size set, then apply the Kolmogorov-Smirnov goodness-of-fit test to the marginal I, so the null-hypothesis $H_0 : \mathrm{P}_t^* = \mathrm{P}_{0,t}$. previously this method was used by Chen et al. (2007) which focus on non-data-driven. Bertsimas et al. (2018) use data-driven where the average distribution and its support are unknown, they use a multivariate hypothesis test approach based on a combined univariate test. Forward and backward deviations from the univariate P_t distribution are known as follows:

$$\sigma_{fi}(\mathrm{P}_i) = \sup_{x>0} \sqrt{-\frac{2\mu_i}{x} + \frac{2}{x^2}\log\left(\mathrm{E}^{\mathrm{P}_i}\left[e^{x\overline{u}_i}\right]\right)}, \text{ and } \sigma_{bt}(\mathrm{P}_i) = \sup_{x>0} \sqrt{-\frac{2\mu_i}{x} + \frac{2}{x^2}\log\left(\mathrm{E}^{\mathrm{P}_i}\left[e^{x\overline{u}_i}\right]\right)}, \tag{1}$$

where $\mathrm{E}^{\mathrm{P}_i}\left[\overline{u}_i\right] = \mu_i$. Next use three null hypotheses, which are:

$$H_0^1 : \mathrm{E}^{P_i^*}\left[\overline{u}\right] = \mu_{0,i},\ H_0^2 : \sigma_{fi}\left(\mathrm{P}_i^*\right) \leq \sigma_{0,fi},\ H_0^3 : \sigma_{bi}\left(\mathrm{P}_i^*\right) \leq \sigma_{0,bi} \tag{2}$$

Next is the set of uncertainties from marginal samples, Bertsimas et al (2018) analyzed samples from marginal P^* distributions separately. The multivariate hypothesis is given as follows:

$$H_0 : \mathrm{VaR}_{\varepsilon/d}^{\mathrm{P}^*}(e_i) \geq \overline{q}_{i,o} \text{ dan } \mathrm{VaR}_{\varepsilon/d}^{\mathrm{P}^*}(-e_i) \geq \underline{q}_{i,0} \text{ for all } i = 1,\ldots d. \tag{3}$$

The set of uncertainties for non-independent potential components assumes samples from an infinite P^* combined distribution. Bertsimas et al (2018) consider a goodness-of-fit test on a linear-convex model with a null hypothesis $H_0 : \mathrm{P}^* = \mathrm{P}_0$.

A set of uncertainties can also be built with machine learning, including Support Vector Clustering (SVC), Dirichlet Process Mixture Model (Ning & You, 2017), and Principle Component Analysis (PCA). Shang et al. (2017), Qiu et al. (2019), and Mohseni et al. (2019) develop a set of uncertainties using Support Vector Clustering (SVC). SVC is an unsupervised learning approach. SVC methods are used to model complex high dimensional data with uncertainties and solve nonparametric grouping problems. In SVC algorithms, data points are mapped from the input space to the feature space using the kernel function. Assume a set is a set of samples. Non-linear mapping of $\phi(\mathbf{u}) : \mathsf{R}^n \rightarrow \mathsf{R}^K$, SVC searches for the smallest sphere that includes all the data by formulating optimization problems $\min_{P,R}\left\{R^2 \middle| \left\|\phi\left(\mathbf{u}^{(i)}\right) - \mathbf{P}\right\|^2 \leq R^2, i = 1,\ldots,N\right\}$. Shang et al. (2017) obtained the following set of data-driven uncertainties:

$$\mathsf{U}(\mathsf{D}) = \left\{\mathbf{u} \middle| K(\mathbf{u},\mathbf{u}) - 2\sum_{i=1}^{N}\alpha_i K\left(\mathbf{u},\mathbf{u}^{(i)}\right) + \sum_{i=1}^{N}\sum_{j=1}^{N}\alpha_i\alpha_j K\left(\mathbf{u}^{(i)},\mathbf{u}^{(j)}\right) \leq R^2\right\} \tag{4}$$

Shang et al. Propose covariate information on the Generalized Intersection Kernel so that the Weighted Generalized Intersection Kernel (WGIK) is obtained. Therefore, the set of data-driven uncertainties is as follows:

$$\mathsf{U}_v(\mathsf{D}) = \left\{\mathbf{u} \middle| \sum_{i \in CV}\alpha_i \left\|Q\left(\mathbf{u} - \mathbf{u}^{(i)}\right)\right\|_1 \leq \sum_{i \in CV}\alpha_i \left\|Q\left(\mathbf{u}^{(i_i)} - \mathbf{u}^{(i)}\right)\right\|_1, i' \in BSV\right\} \tag{5}$$

Ning & You (2018) and Dai et al. (2020) used Principle Component Analysis and Robust Kernel Density Estimation based on historical data to construct a set of uncertainties. Zhang et al. (2018) employ a set of density-based uncertainties by minimizing the tolerable probability of uncertainty scenarios and the empirical density function of uncertainties.

4. Data-Driven Robust Optimization Application

Data-driven methods in the robust optimization model have been applied in several fields, including inventory, portfolio, scheduling, industry, and transportation problems.

4.1. Data-Driven Robust Optimization Application on Inventory Issues

On inventory issues, most researchers focus on models with known demand distribution. But in reality, the distribution of requests is largely unknown and only historical data is available. Klabjan et al. (2013) apply a data-driven method to solve robust models of inventory problems by reordering. They propose a robust stochastic model for multi-period lot-sizing problems in models with unknown demand distributions. The convergence of the results for the model based on the chi-square test shows that the robust stochastic approach converges with the stochastic programming solution with a fairly large sample size. Whereas in the relatively small sample size, the solution obtained from the robust model is not affected by the uncertainty in the demand distribution.

Zhao & You (2018) discusses the application of data-driven robust optimization to supply chain problems with uncertain production capacity. The proposed model is a fractional two-stage model with resilience and economic objectives. The objective of resilience is to maximize supply chain resilience in the worst conditions based on the ratio between accumulated and uninterrupted supply chain performance, while the economic goal is to minimize nominal costs without interruption, including facility location costs, additional capacity, and operational costs. In achieving these two objectives, Zhao & You (2018) divided the model into two stages, the first stage for location decisions, the production capacity of each facility, and transportation was completed using a combination of parametric algorithms and affine decision rules. The model in the first stage is transformed into a robust and static linear integer assist programming problem. Parametric functions have important properties so that when reformulated they still have an optimal solution that is identical to the original problem with the fractional objective function. The second stage for work capacity decisions for each facility and recovery schedule is completed by the heuristic method which results in less computing time with more real quality solutions.

4.2. Data-Driven Robust Optimization Application on Portfolio Selection Issues

Portfolios are allocating capital to several available assets to get maximum investment returns and minimal risk. The selection of the portfolio was first formulated by Markowitz (1952). In the Markowitz portfolio selection model, portfolio returns are measured as the expected value of random portfolio returns, while the risk is the amount of variance of portfolio returns. There are several robust mean-variance portfolio selection models, including the robust variance minimization model, the robust return maximization model, the Sharpe robust ratio maximization problem, the robust Value-at-Risk (VaR) portfolio model, and the robust portfolio allocation with an uncertainty covariance matrix.

Goldfarb & Iyengar (2003) propose an alternative deterministic model that is robust to the uncertainty of parameters and estimated errors, where the market parameter disturbances in the model are unknown and the optimization problem is solved by assuming the worst-case disturbance. They developed a robust factor model for return on assets, viz:

$$\mathbf{r} = \boldsymbol{\mu} + \mathbf{V}^{\mathbf{T}}\mathbf{f} + \boldsymbol{\varepsilon} \tag{6}$$

where $\mathbf{r} \in \mathbf{R}^n$ random asset return vector, $\boldsymbol{\mu} \in \mathbf{R}^n$ average return vector, $\mathbf{f} \in \mathbf{R}^m$ random market return vector factor, $\mathbf{V} \in \mathbf{R}^{m \times n}$ matrix loading factor, and $\boldsymbol{\varepsilon}$ residual return vector. Besides, they show that the set of uncertainties for market parameters is defined as a statistical procedure for estimating parameters of market return data, and the set of uncertainty robust optimization problems can be re-formulated as Second-Order Cone Programs (SOCPs). Goldfarb & Iyengar (2003) shows the stages to solve robust portfolio selection problems, i.e. collect data returns from assets and returns from factors, use one asset at a time, evaluate least-square estimates of averages and factor loading matrices, then select boundaries trust and specify a bootstrap confidence interval around the variance or use the worst error variance estimate. Next, determine the projections S_m for the length of the $\boldsymbol{\mu}$ and S_v projections along the $\mathbf{V}$ vector, and solve the robust problem.

Generally, researchers estimate the covariance matrix factors from known and stable market models. But the complexity of the market model makes it possible for uncertainties in the covariance matrix. Goldfarb & Iyengar (2003) developed the structure of covariance matrix inequality, namely the inverse structure of inverse covariance and covariance. Structural uncertainty for inverse covariance considering the following covariance matrix factors:

$$S_{f^{-1}} = \left\{ \mathbf{F} \middle| \mathbf{F}^{-1} = \mathbf{F}_0^{-1} + \boldsymbol{\Delta} \succeq 0, \boldsymbol{\Delta} = \boldsymbol{\Delta}^T, \left\| \mathbf{F}_0^{1/2} \boldsymbol{\Delta} \mathbf{F}_0^{1/2} \right\| \leq \eta \right\}, \tag{7}$$

where $\mathbf{F} \succ \mathbf{0}$, $\left\| \mathbf{F}_0^{1/2} \boldsymbol{\Delta} \mathbf{F}_0^{1/2} \right\| = \max_i \left| \lambda_i \left(\mathbf{F}_0^{1/2} \boldsymbol{\Delta} \mathbf{F}_0^{1/2} \right) \right|$ with $\left\{ \lambda_i \left(\mathbf{F}_0^{1/2} \boldsymbol{\Delta} \mathbf{F}_0^{1/2} \right) \right\}$ is the eigenvalue of $\mathbf{F}_0^{1/2} \boldsymbol{\Delta} \mathbf{F}_0^{1/2}$.

Chi et al. (2019) use data-driven problems with Peer-to-Peer (P2P) lending investment. P2P lenders can invest a portion of each loan. Thus, P2P loan investment decisions can be turned into a matter of optimizing the loan portfolio. In the P2P problem, there are two challenges, namely the unavailability of information on historical loan data and the uncertainty of loan distribution. These two things make assessing new loan risk very challenging. Chi et al. (2019) use an instance-based assessment framework to estimate return expectations and kernel regression of returns and risks to investigate nonlinear relationships between random variables.

Kang et al. (2018) present data-driven optimization in the selection of a mean-CVaR robust optimization portfolio under the uncertainty of distribution. They use a nonparametric bootstrap approach to deal with the non-stop and show that the selection of robust portfolios varies with the value of the input. Kang et al. (2018) combine the uncertainty of mean, covariance, and distribution using one set of distributions to replace a single distribution in the worst-case.

4.3. Data-Driven Robust Optimization Application in Scheduling Problems

Ning & You (2016) propose a new scheduling approach based on a model that is combined with Mixed Integer Linear Programming (MILP). The set of uncertainties is obtained from uncertain historical parameter data. Robust counterpart is reduced to a tractable conic quadratic Mixed-Integer Programming. The robust data-driven scheduling model aims to maximize profits and meet the epigraphic reformulation constraints of the objective function, assignment constraints, time constraints, batch size constraints, mass balance and storage constraints, demand constraints, and other constraints.

Qiu et al. (2019) use data-driven to determine the set of uncertainties in multi-product inventory problem requests. The set of uncertainties is built using SVC. The robust counterpart was developed using the absolute robustness criterion into a linear programming model. The results of his research show that the robust model of data-based optimization with the SVC method is superior to the set of box and ellipsoidal uncertainty.

4.4. Data-Driven Robust Optimization applications in the industry

Optimization methods are widely applied in the steam system industry because they can improve efficiency and economic benefits and save energy. Zhao et al. (2019) apply data-driven robust optimization to the uncertainty of the semiempirical turbine model parameters using historical data. They use the kernel density estimation method to determine the set of uncertainties. At first, the semi-empirical model of the steam turbine was developed based on the process mechanism and operational data. Robust counterpart from data-driven robust optimization is derived as a Mixel Integer Linear Programming (MILP) model. In another study, Zhao et al. (2019) also implemented a data-driven robust optimization of the steam system at an ethylene plant.

In the process industry, Zhang et al. (2018) apply data-driven to robust optimization to deal with uncertain environmental and operational conditions. The set of uncertainties is defined by probability density contours. Their research also estimates the nob-convex set of uncertainties. From the results of his research, it was found that about 2% of fluctuations in gas fuel consumption can be controlled.

4.5. Data-Driven Robust Optimization Application on Transportation Problems

Xie et al. (2018) analyzed the application of data-driven robust optimization to determine the location and size of electric vehicle charging stations. They use a two-stage model to solve this problem. The first stage is a linear programming model for determining location using Monte Carlo simulations for spatial and temporal distribution of requests. This first stage relies on Value-at-Risk (VaR). In the second stage, they develop a data-driven robust optimization model to optimize renewable capacity based on Conditional Value-at-Risk (CVaR). The results of Xie et al. (2018) shows that optimal solutions with data-driven robust optimization can overcome various types of distribution uncertainty.

Chassein et al. (2019) analyze the shortest path problem with the set of uncertainty based on data. They found several results on this problem, including the convex-hull solution showing good sample performance, but it was not stable against scenarios outside the sample; easy and fast interval solutions are obtained on a small scale, but the performance is generally not good; Budgeted uncertainty sets are less effective in this problem; Ellipsoidal sets as a whole have good performance; whereas symmetrical permutohull solutions tend to be less strong, but provide excellent performance. Based on this, Ellipsoidal uncertainty set provides high-quality solutions with computational efforts, so Chassein et al. (2019) discuss in more detail the set of ellipsoidal uncertainty. Next, to get the optimal extremely efficient solution, they use the Naïve algorithm.

5. Conclusion

In the current era of big data, where data availability becomes abundant, the uncertainty of parameters in robust optimization can be determined based on data. Data-driven robust optimization has been applied in several fields, including inventory issues, portfolio selection, scheduling, transportation, and industry. The set of uncertainties based on data can be done with a hypothesis test or without a hypothesis test. Besides, the use of machine learning in determining the set of robust optimization uncertainty becomes a challenge. Based on the discussion, the data-driven uncertainty set is superior to the traditional set. The robust optimization model that is built with a set of data-driven uncertainties provides good performance.

References

Bandi, C., & Bertsimas, D. (2012). Tractable Stochastic Analysis in High Dimensions via Robust Optimization. *Mathematical Programming*, 1-35.

Ben-Tal, A., & Nemirovski, A. (2000). Robust Solution of Linear Programming Problems Contaminated with Uncertain Data. *Mathematical Programming*, 88(3): 411-424.

Ben-Tal, A., El Ghaoui, L., & Nemirovski, A. (2009). *Robust Optimization*, Princeton University Press.

Bertsimas, D., & Sim, M. (2004). The Price of Robustness. *Operations Research*, 52(1): 35-53.

Bertsimas, D., Gupta, V., & Kallus, N. (2018). Data-Driven Robust Optimization. *Mathematical Programming*, 167: 235-292.

Campi, M.C., & Garatti, S. (2008). The Exact Feasibility of Randomized Solutions of Uncertain Convex Programs. *SIAM Journal on Optimization*, 19(3): 1211-1230.

Chassein, A., Dokka, T., & Goerigk, M. (2019). Algorithms and Uncertainty Sets for Data-Driven Robust Shortest Path Problem. *European Journal of Operational Research*, 274(2): 671-686.

Chi, G., Ding, S., & Peng, X. (2019). Data-Driven Robust Credit Portfolio Optimization for Investment Decisions in P2P Lending. *Mathematical Problems in Engineering*, 2019.

Dai, X., Wang, X., He, R., Du, W., Zhong, W., Zhao, L., & Qian, F. (2020). Data-Driven Robust Optimization for Crude Oil Blending under Uncertainty. *Computers & Chemical Engineering*, 136.

Goldfarb, D., & Iyengar, G. (2003). Robust Portfolio Selection Problems. Mathematics of Operations Research, 28(1): 1-38.

Kang, Z., Li, X., Li Z., & Zhu, S. (2018). Data-Driven Robust Mean-CVaR Portfolio Selection under Distribution Ambiguity. *Quantitative Finance*.

Klabjan, D., Simchi-Levi, D., & Song, M. (2003). Robust Stochastic Lot-sizing By Means of Histograms. Production and Operations Management, 691-710.

Markowitz, H. M. (1952). Portfolio Selection. *Journal of Finance*, 7: 77-91.

Mohseni, S., & Pishvaee, M. S. (2019). Data-driven Robust Optimization for Wastewater Sludge-to-Biodiesel Supply Chain Design. *Computers & Industrial Engineering*, 139.

Ning, C., & You, F. (2016). Data-driven Robust MILP Model for Scheduling of Multipurpose Batch Processes under Uncertainty. *In 2016 IEEE 5th Conference on Decision and Control*, Las Vegas, USA.

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.

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.

Qiu, R., Sun, Y., Fan, Z., & Sun, M. (2019). Robust Multi-Product Inventory Optimization under Support Vector Clustering-based Data-Driven Demand Uncertainty Set. *Soft Computing*, 24: 6259-6275.

Shang, C., Huang, X., & You, F. (2017). Data-Driven Robust Optimization based on Kernel Learning. *Computer & Chemical Engineering*, 106: 464-479.

Wang, Z., Glynn, P. W., & Ye, Y. (2016). Likelihood Robust Optimization for Data-Driven Problems. *Computational Management Science*, 13: 241-261.

Xie, R., Wei, W., Khodayar, M. E., Wang, J., & Mei, S. (2018). Planning Fully Renewable Powered Charging Stations on Highways: A Data-Driven Robust Optimization Approach. *IEEE Transactions on Transportation Electrification*, 4(3): 817-830.

Zhang, Y., Feng, Y., & Rong, G. (2018). Data-Driven Rolling-Horizon Robust Optimization for Petrochemical Scheduling using Probability Density Contours. *Computers and Chemical Engineering*, 15: 342-360.

Zhao, L., Ning, C., & You, F. (2019). A Data-Driven Robust Optimization Approach to Operational Optimization of Industrial Steam Systems under Uncertainty. *In Proceedings of the 29th European Symposium on Computer Aided Process Engineering*, Netherlands.

Zhao, L., Zhong, W., & Du, W. (2019). Data-Driven Robust Optimization for Steam Systems in Ethylene Plants under Uncertainty. *Processes*, 7(10), 744.

Zhao, S., & You, F. (2018). Resilient Supply Chain Design and Operations with Decision-Dependent Uncertainty using a Data-Driven Robust Optimization Approach. *AIChE Journal*, 65(3).

Biographies

Diantiny Mariam Pribadi is a magister student in the Department of Mathematics, Faculty of Mathematics and Natural Sciences, Universitas Padjadjaran, the field of applied mathematics, with a field of actuarial science.

Sukono is a lecturer in the Department of Mathematics, Faculty of Mathematics and Natural Sciences, Universitas Padjadjaran. Currently serves as Head of Master's Program in Mathematics, the field of applied mathematics, with a field of concentration of financial mathematics and actuarial sciences.

Riaman is a teaching staff in the mathematics study program at the Mathematics Department of the University of Padjadjaran since 1997. The mathematics degree was won from the Mathematics Department of the University of Padjadjaran in 1995, while the Masters of Actuarial Science was obtained from the Bandung Institute of Technology in 2000. The field of research currently being undertaken is the field of applied mathematics. More specifically are: fields of financial mathematics, actuarial mathematics, survival models, mathematical statistics, and reliability. Since 2001 until now he has joined as a member of IndoMS, and since 2017 joined IAENG.

Kalfin is currently a Student at the Doctoral Program at Padjadjaran University, Indonesia since 2019. He received his M.Mat in Mathematics from Padjadjaran University (UNPAD), Indonesia in 2019. His current research focuses on financial mathematics and actuarial sciences.

Abdul Talib Bon is a professor of Production and Operations Management in the Faculty of Technology Management and Business at the Universiti Tun Hussein Onn Malaysia since 1999. He has a PhD in Computer Science, which he obtained from the Universite de La Rochelle, France in the year 2008. His doctoral thesis was on topic Process Quality Improvement on Beltline Moulding Manufacturing. He studied Business Administration in the Universiti Kebangsaan Malaysia for which he was awarded the MBA in the year 1998. He's bachelor degree and diploma in Mechanical Engineering which his obtained from the Universiti Teknologi Malaysia. He received his postgraduate certificate in Mechatronics and Robotics from Carlisle, United Kingdom in 1997. He had published more 150 International Proceedings and International Journals and 8 books. He is a member of MSORSM, IIF, IEOM, IIE, INFORMS, TAM and MIM.