Quantile-based Sequential Learning and Optimization for Contextual Stochastic Vehicle Routing

Nele Bertling*, Michael Römer and Kevin Tierney
Bielefeld University, Germany
{nele.bertling, michael.roemer, kevin.tierney}@uni-bielefeld.de
* Corresponding author

Extended abstract submitted for presentation at the 12th Triennial Symposium on Transportation Analysis conference (TRISTAN XII)
June 22-27, 2025, Okinawa, Japan

February 10, 2025

Keywords: contextual stochastic optimization problems, vehicle routing, sequential learning and optimization, integrated learning and optimization, quantile regression

1 Introduction

In real-world routing problems key input parameters, such as the travel time between locations, are affected by uncertainty and usually depend on *contextual information* such as the time of day or the weather. Following the terminology proposed in Sadana *et al.* (2025), they form contextual stochastic (combinatorial) optimization (CSO) problems that can be addressed by combining Machine Learning (ML) with combinatorial optimization methods to provide estimates for the uncertain parameters.

One way to combine ML with optimization methods is the sequential learning and optimization (SLO) framework, als known as the (two-stage) predict-then-optimize (PtO) framework. In this approach, uncertain parameters are predicted using standard ML models (and loss functions) and then put into a deterministic model to make a decision. A major critique of SLO is, that minimizing the prediction loss may not align with the goal of making an optimal decision, which can lead to bad-quality decisions (Sadana *et al.*, 2025, Mandi *et al.*, 2024).

This motivates the use of integrated learning and optimization (ILO) approaches, also referred to as end-to-end predict-then-optimize or predict-and-optimize paradigm, which seek to directly minimize the decision loss. A challenging aspect of ILO is the dependence of the loss function on the solution of the optimization problem, particularly when decision variables are discrete, as this introduces non-differentiability and non-smoothness into the decision loss (Mandi *et al.*, 2024, Sadana *et al.*, 2025). Since many ML models are unable to deal with such loss functions, a common approach is to replace the exact decision loss by smooth and differentiable *surrogate loss functions* such the smart predict-then-optimize+ (SPO+) loss introduced by Elmachtoub & Grigas (2022) and the perturbed Fenchel-Young (PFYL) loss proposed by Berthet *et al.* (2020).

Next to SPO+ and PFYL, several other SLO and ILO methods are implemented in the PyEPO software package (Tang & Khalil, 2024) along with a set of computational experiments comparing the performance of these methods on various combinatorial optimization problems including the shortest path problem (SPP) and the traveling salesperson problem (TSP). Tang & Khalil (2024) show that ILO performs better for complex functions but requires longer training time compared to SLO methods.

This highlights the need for a more robust SLO approach which motivates the following contributions of this work. First, we propose a quantile-based SLO approach that predicts a quantile instead of the expected values of the uncertain parameters in the first stage of the SLO method. In our proposed approach, the quantile to be predicted is determined at training time by choosing the quantile that yields the best decision loss on the training data. Second, using PyEPO, we implement and compare various SLO and ILO methods for the capacitated vehicle routing problem (CVRP), which to the best of our knowledge has not yet been done. Our experimental results show, that our quantile-based SLO approach generally outperforms the classical expected value-based SLO approaches and that, for the CVRP, our quantile-based SLO approach reduces the average normalized regret by 0.03 points compared to the best-performing ILO method.

2 Quantile-Based Sequential Learning and Optimization

We consider contextual stochastic optimization problems with uncertain cost parameters $\mathbf{c}$ in the objective function, following the notation of Mandi *et al.* (2024):

$$\mathbf{x}^\star(\mathbf{c}) = \arg\min_{\mathbf{x}} f(\mathbf{x}, \mathbf{c}) \quad \text{s.t. } g(\mathbf{x}) \leq 0,\, h(\mathbf{x}) = 0. \tag{1}$$

The goal is to find an optimal solution $\mathbf{x}^\star(\mathbf{c}) \in \mathbb{R}^n$ that minimizes the objective function f. To predict costs $\mathbf{c}$, we use historical data comprising contextual observed characteristics $\mathbf{z}$, represented in the form $D = \{(\mathbf{z}_i, \mathbf{c}_i)\}_{i=1}^N$ (Sadana *et al.*, 2025, Mandi *et al.*, 2024).

In ILO, the costs $\hat{\mathbf{c}}$ are estimated by minimizing the regret associated with the decision $\mathbf{x}^\star(\hat{\mathbf{c}})$ in comparison to the optimal decision $\mathbf{x}^\star(\mathbf{c})$, based on the true value of $\mathbf{c}$ (Mandi *et al.*, 2024):

$$\mathit{Regret}(\mathbf{x}^\star(\hat{\mathbf{c}}), \mathbf{c}) = f(\mathbf{x}^\star(\hat{\mathbf{c}}), \mathbf{c}) - f(\mathbf{x}^\star(\mathbf{c}), \mathbf{c}). \tag{2}$$

As previously noted, this regret function is often non-differentiable, prompting the use of differentiable surrogate loss functions such as SPO+ and PFYL instead.

In SLO, any classical ML model (and loss function) can be used to estimate $\hat{\mathbf{c}}$ using the contextual data D. Typically, these models use the (conditional) expected value for this purpose. The key idea of our approach is to predict a quantile instead of the expected value using quantile regression (QR). Intuitively, this approach has two advantages: First, quantile predictions tend to be more robust against irregularities in data and outliers. Second, it allows us to adapt the prediction approach to the CSO problem by adjusting the quantile during training. Specifically, we aim at finding the quantile for which the predictions $\hat{\mathbf{c}}$ yield the smallest regret (Eq. (2)) on the training data. To do so, we use a binary search (Manna & Waldinger, 1987) for quantile values within the interval $[0.2, 0.8]$.

Similar to our approach, Wang & Yan (2022) propose using QR to predict the conditional distribution of parameters in an SLO setting. However, no experimental results are given. Mahmutoğulları & Guns (2023) explore the use of quantile predictions to leverage contextual information for robust solutions of the capacitated vehicle routing problem with time windows (CVRPTW) with uncertain demands. Quantiles are used to predict constraint parameters of a deterministic model as well as for constructing uncertainty sets for a robust optimization model. They note that a quantile must be specified by the decision-maker before initiating the optimization, meaning that unlike in our approach, quantile selection is not part of the training process.

3 Results

We conduct experiments using the PyEPO framework to evaluate the effectiveness of the quantile-based SLO approach for three logistics-related problems: SPP instances organized in a 5x5 node grid, as well as TSP and CVRP instances with 20 nodes each. The SPP and TSP instances are solved with Gurobi, while we use PyVRP for the CVRP instances (Wouda *et al.*, 2024). The data for the SPP and TSP is generated using the methodology outlined by Tang & Khalil

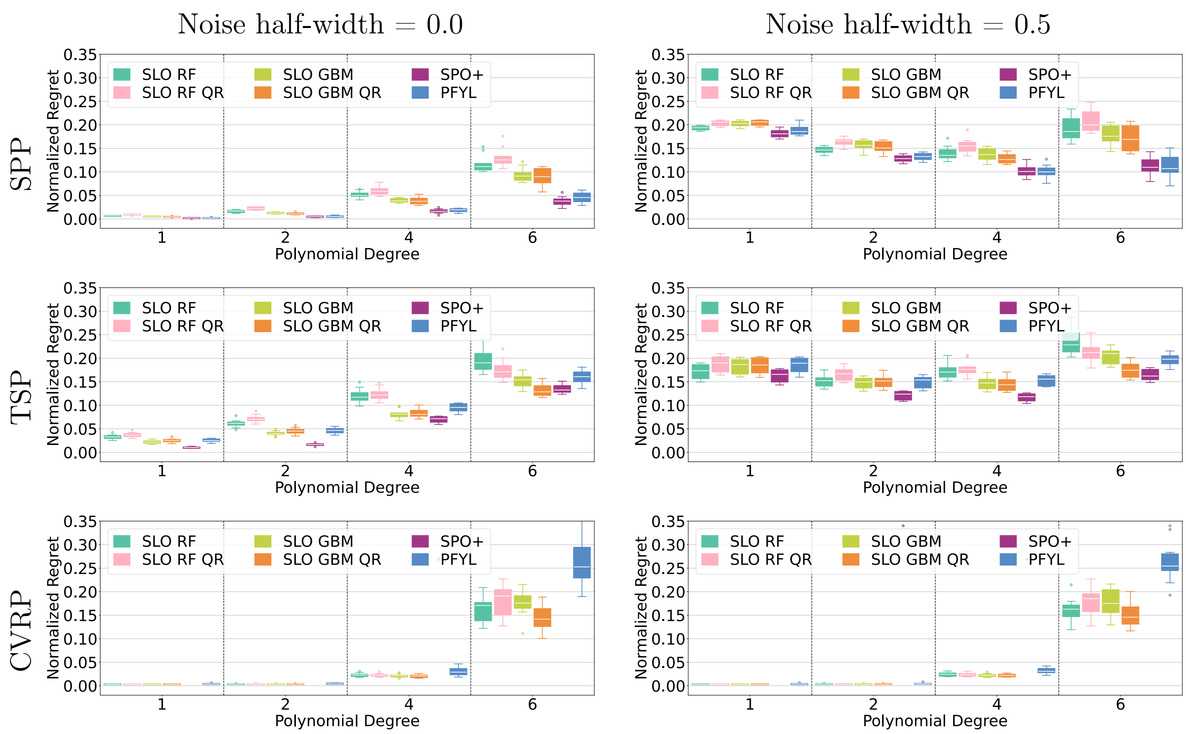
(2024), while the CVRP data follows the pipeline proposed by Uchoa *et al.* (2017). The costs and observed features are generated in accordance with Tang & Khalil (2024), where costs vectors $\mathbf{c}_i$ are defined by

$$\left[\frac{1}{3^{deg-1}}\left(\frac{1}{\sqrt{p}}(\mathbf{B}\mathbf{z}_i)_j + 3\right)^{deg}\right] \cdot \epsilon_{ij}. \tag{3}$$

The costs depend on a random matrix $\mathbf{B}$ following a Bernoulli distribution and the number of available features p. The polynomial degree deg and and the noise half-width, which affects the noise ϵ, can be changed to modify the cost function's complexity.

To compare SLO with ILO approaches, we evaluate random forest regression (RF), random forest quantile regression (RF QR), gradient boosting (GBM), and gradient boosting quantile regression (GBM QR) for SLO predictions, alongside the ILO methods SPO+ and PFYL. While PyEPO offers additional methods, our focus is on the most effective ones, as shown in Table 1.

Table 1 – *Experimental results for SPP, TSP and CVRP. For each problem, 10 experiments are conducted using a training dataset of 100 instances and evaluated with the normalized regret of the test datasets with 1000 instances. The uncertain costs (Eq. (3)) vary with different polynomial degrees and noise half-widths.*



For each type of problem, we compare the best-performing quantile-based SLO with the best-performing ILO approach. Across 80 experiments per problem, SPO+ yields a lower normalized regret than GBM QR by an average of 0.03 points in the SPP and 0.02 points in the TSP, outperforming GBM QR in 71 of the latter's trials. For the CVRP, GBM QR surpasses PFYL in 76 experiments, reducing normalized regret by an average of 0.03 points. Notably, in the case of the CVRP, SPO+ demonstrates markedly suboptimal performance relative to the other evaluated methods, with results that are exceedingly poor, placing them outside the plotted range.

All experiments run on an AMD Milan 7763 system with 4 CPUs and 32 GB RAM. For the SPP, SPO+ requires an average of 15.97 seconds for training, while GBM QR requires an average of 35.3 seconds. In the TSP, SPO+ needs an average of 140.37 seconds for training, compared to 187.46 seconds for GBM QR. For the CVRP, PFYL takes an average of 351.71 seconds for training, while GBM QR requires only 206.50 seconds. Although ILO methods are faster for the SPP and TSP, a similar trend as for the normalized regret is observed, namely that for the CVRP, GBM QR significantly outperforms PFYL in terms of training time.

4 Discussion

In this work, we show that quantile-based SLO methods can match or even beat the performance of ILO methods, while consistently surpassing classical SLO approaches based on the (conditional) expected value. For simpler CSO problems like the SPP, ILO methods consistently outperform SLO methods. However, for the TSP, SLO methods approach the effectiveness of ILO. Ultimately, SLO surpasses ILO in the more complex CVRP scenario. Interestingly, while SPO+ outperforms PFYL in the SPP and TSP, it lags behind PFYL in the CVRP.

Overall, our findings do not support the notion that ILO methods invariably outperform SLO methods. This observation challenges the assumption that surrogate functions are inherently better and underscores the necessity for approaches that do not depend on derivatives. In this regard, quantile regression offers a promising way forward, as it is a derivative-free method that prioritizes regret minimization while tuning the used quantile.

Future research will investigate the framework's scalability, as the current experiments are limited to relatively small problems. Despite the constrained training set sizes, we demonstrate that quantile-based SLO methods outperform ILO methods for the CVRP. However, Tang & Khalil (2024) indicate that the differences between ILO and SLO methods diminish with larger training set sizes, highlighting the need to understand the impact of such expansions. Furthermore, more realistic datasets should be considered to address problems that extend beyond normal or uniformly distributed data. In these scenarios, it is essential to evaluate the influence of heuristic solvers on the performance of these frameworks, as the reliance on solvers with optimality assurance may no longer be feasible.

References

Berthet, Quentin, Blondel, Mathieu, Teboul, Olivier, Cuturi, Marco, Vert, Jean-Philippe, & Bach, Francis. 2020 (June). *Learning with Differentiable Perturbed Optimizers.*

Elmachtoub, Adam N., & Grigas, Paul. 2022. Smart "Predict, Then Optimize". *Management Science*, **68**(1), 9–26.

Mahmutoğulları, Ali İrfan, & Guns, Tias. 2023 (Oct.). *Leveraging Contextual Information for Robustness in Vehicle Routing Problems.*

Mandi, Jayanta, Kotary, James, Berden, Senne, Mulamba, Maxime, Bucarey, Victor, Guns, Tias, & Fioretto, Ferdinando. 2024 (May). *Decision-Focused Learning: Foundations, State of the Art, Benchmark and Future Opportunities.*

Manna, Zohar, & Waldinger, Richard. 1987. The Origin of a Binary-Search Paradigm. *Science of Computer Programming*, **9**(1), 37–83.

Sadana, Utsav, Chenreddy, Abhilash, Delage, Erick, Forel, Alexandre, Frejinger, Emma, & Vidal, Thibaut. 2025. A Survey of Contextual Optimization Methods for Decision-Making under Uncertainty. *European Journal of Operational Research*, **320**(2), 271–289.

Tang, Bo, & Khalil, Elias B. 2024. PyEPO: A PyTorch-based End-to-End Predict-Then-Optimize Library for Linear and Integer Programming. *Mathematical Programming Computation*, **16**(3), 297–335.

Uchoa, Eduardo, Pecin, Diego, Pessoa, Artur, Poggi, Marcus, Vidal, Thibaut, & Subramanian, Anand. 2017. New Benchmark Instances for the Capacitated Vehicle Routing Problem. *European Journal of Operational Research*, **257**(3), 845–858.

Wang, Shuaian, & Yan, Ran. 2022. "Predict, Then Optimize" with Quantile Regression: A Global Method from Predictive to Prescriptive Analytics and Applications to Multimodal Transportation. *Multimodal Transportation*, **1**(4), 100035.

Wouda, Niels A., Lan, Leon, & Kool, Wouter. 2024. PyVRP: a high-performance VRP solver package. *INFORMS Journal on Computing*, **36**(4), 943–955.