



Network flow problem heuristic reduction using machine learning

M. Rosecký¹ · J. Pluskal² · R. Šomplák²

Received: 10 January 2023 / Revised: 23 June 2023 / Accepted: 21 August 2023
© The Author(s) 2023

Abstract

Most of the supporting tools developed for logistic optimization and processing in infrastructure planning are based on the network flow problem. The real-world application of these instruments can provide great insight and help to ensure long-term sustainability. The main limitation of these tools lies in great computing demand when there is the necessity of solving large-scale tasks in great detail. It means that the ability to find the optimal solution for real-world problems is limited. Thus, the detail of infrastructure is often reduced by data aggregation or heuristic approaches are used to find a suboptimal solution. This paper proposes a machine learning classification model to reduce the number of variables for an exact solution algorithm. First, the design of experiments is used to create a set of smaller problems that are possible to solve exactly. Artificial data are used at this stage, while domain knowledge is used to set appropriate distribution and parameters. Second, the classification model estimates the probability of the presence of each arc in the optimal solution. Features, which are related to costs and capacity, of each arc are utilized in the classification model. Models created on a subset of generated problems are then tested on the other problems. Finally, the proposed framework is applied to the waste management problem in the Czech Republic. The results of the verification show, that it is possible to remove 95% of arcs without impact on strategic decisions and without significant change of an objective function. The computing time of the reduced problem takes only 7% of the original task.

Keywords Model-size reduction techniques · Classification model · Arcs removal · Waste-to-Energy localization · Large-scale optimization

✉ M. Rosecký
Martin.Rosecky@vutbr.cz

¹ Faculty of Mechanical Engineering, Institute of Mathematics, Brno University of Technology, Brno, Czechia

² Faculty of Mechanical Engineering, Institute of Process Engineering, Brno University of Technology, Brno, Czechia

List of symbols

FOR	False omission rate
GAMS	General Algebraic Modelling System
ML	Machine learning
MMW	Mixed municipal waste
OOB	Out-of-back
RF	Random forest
SOS1	Special Ordered Set, type 1
WtE	Waste-to-Energy
WM	Waste management
A_Cap_Prod_ratio	Ratio of capacity and production of arc, [-]
A_TotC	Total cost of arc, [EUR]
A_TotC_Cap_grad	Gradient of the quotient total cost and capacity of arc, [-]
A_TotC_Fac_rank	Relative order of arc related to facility, [-]
A_TotC_grad	Gradient of total cost of arc, [-]
A_TotC_Mun_rank	Relative order of arc related to municipality, [-]
A_TotC_TotCMax_ratio	Ratio of total cost of arc and the most expensive arc, [-]
A_TrC	Transport cost of arc, [EUR]
A_TrC_ProcC_ratio	Ratio of transport and processing cost of arc, [-]
N_Cap_Mun_den	Capacity density related to municipalities, [t/km]
N_Prod_Fac_den	Production density related to facilities, [t/km]
N_Prod_Mun_den	Production density related to municipalities, [t/km]
N_TotC_kurt	Kurtosis of total cost of node, [-]
N_TotC_var	Variance of total cost of node, [EUR ²]
S_TotC_skew	Skewness of total cost of system, [-]
S_TotC_var	Variance of total cost of system, [EUR ²]

1 Introduction

Part of today's technological progress is the implementation of computer-controlled systems, including artificial intelligence, which allows a whole range of systems to be effectively managed without human intervention (Sharma 2021). These systems are currently very popular and are the subject of research and improvement in numerous research centers. However, there are still situations that require human decision-making, as these involve important conceptual decisions that impact various aspects of society. Alternatively, these may involve steps or procedures that are non-repetitive and lack sufficient data. For these purposes, supporting tools are created to provide insight into the issues addressed and to design or simulate solutions for various scenarios and uncertainties in the system. This enables managers to thoroughly analyze the problem, address all key questions, and ultimately implement a robust solution.

Many of these support tools for logistics and processing infrastructure planning are based on the principle of flow in the network, where the individual nodes constitute

municipalities and edges or arcs represent the transport infrastructure between them. However, with increasing digitization and data collection, the size and complexity of systems reach their limits. To provide an appropriate estimate, it is necessary to solve these tasks in the maximum possible detail and thus model the lowest territory division and infrastructure. In the case of the inclusion of real dependencies, which mostly represent nonlinear relationships, these tasks become practically unsolvable despite today's performance of computer technology. Moreover, the solvability of the problem can be limited by available PC memory. This is due to a combinatorial increase in parameters and variables for simulation or optimization software. Considering the significance of these tasks, it is essential to provide at least some solution, even if it deviates from reality to some extent. For this reason, individual studies are often approached by sacrificing necessary details or simplifying certain connections, potentially leading to the loss of crucial information and rendering the calculated recommendations invalid. It puts pressure on pre-processing before an optimization or simulation of a process as well as developing specific algorithms to solve specific problems.

In solving large-scale problems that cannot be computationally optimized by classical methods and approaches, tailored heuristic algorithms are developed (Zhou et al. 2021). These algorithms combine various approaches to search a set of feasible solutions, which can include expert estimation and experience with the problem. As a key part of efficient optimization, the proper formulation of a solved problem is necessary and can significantly speed up the calculation (Burre et al. 2022). According to a recent paper (Karimi-Mamaghan et al. 2022), heuristic algorithms can also be supported by machine learning (ML), which can provide key information based on proper analysis and extraction of available data.

Another approach, how to deal with large-scale problems, could be based on the number of variables reduction (Li and Balakrishnan 2016). This can be implemented in many ways. In the context of network flow tasks, reductions of nodes, edges or arcs are the most frequently mentioned. In principle, this is a pre-processing task, where based on selected parameters, specified structures in the network are removed or appropriately aggregated. In this area, individual approaches differ according to their focus and industry of application. At present, some optimization solvers already have algorithms for reducing variables before the calculation itself (GAMS 2022). However, it is mainly about removing redundant variables or those solutions, which are beyond the limits of the modeled constraints.

The best results can be expected with different combinations of many approaches for solving large-scale tasks. The aim of the presented paper is to create a new approach and tool for reducing network-based optimization tasks. ML can be effectively employed for this purpose, leveraging a sufficient amount of data and suitable parameters to identify optimal and unused arcs with relative accuracy. The proposed framework utilizes insights derived from a large number of small-scale tasks using a selected ML model, which are subsequently applied to large-scale tasks. ML was chosen as the preferred approach due to its ability to handle complex problems like network flow problems more effectively compared to traditional statistical techniques. However, it is important to note that ML models do not provide direct interpretability of individual feature effects. While ML models efficiently solve the problem, they do not offer

deeper insights to researchers. Results show, that with the help of an appropriate ML model, it is possible to significantly reduce the number of variables in the model and thus reduce computational time with minimal intervention in the final solution. The highlights and the main goals of the presented paper are as follows:

- Model reduction based on arc removal to cut down memory requirements and computing time.
- The main idea lies in learning on a small-scale task followed by application on a large-scale task.
- Development of ML model for network flow model in waste management (WM).
- Form a set of important parameters of the problem to characterize all arcs.
- Verification of the developed approach on the case study inspired by WM in the Czech Republic.

After the introduction, a literature review and detailed analysis of approaches addressing time-consuming optimization tasks are conducted, focusing on the novelty and paper contribution. Subsequently, the entire approach is described in more detail, starting from the design of a small test problem, the selection of investigated parameters, and the results obtained through arc removal based on ML. Finally, the approach is applied to solve a complex task within the context of the Czech Republic, including the evaluation and discussion of the acquired results.

2 Literature review

Among the mentioned possibilities of solving large-scale problems, this study is focused on the reduction of variables in the model. The main idea is to replace or limit the original model while preserving the original optimal solution within the feasible set. However, most approaches cannot guarantee the presence of the optimal solution in the modified task. Therefore, reduction techniques are usually employed at the expense of introducing an acceptable error, which should be quantifiable to maintain calculation accuracy. In reality, it is difficult to quantify the error exactly, and there should be at least an error estimate that is verified by many simulations. In order to reduce the size of the model, techniques are developed that allow to effectively modify the solved graphic structure of the network (nodes, edges, arcs). Individual approaches can be divided into several categories according to the technique used between aggregation of structures, compression (removal of redundant structures), simplification, and approximation (Martin et al. 2019).

Articles published during the last decade were studied within the literature review of the used approaches. The articles were searched using the keywords “*decomposition*”, “*clustering*”, “*large-scale model*”, “*model reduction techniques*”, “*network flow model*”, “*machine learning*” and their combinations. Table 1 lists the papers that met the study area, i.e. approaches that reduce the optimization or simulation task in the context of the number of variables to reduce computational time or avoid memory problems.

The overview of the found articles clearly shows the authors’ efforts to reduce the model in many industries. The simulation case studies, dynamic systems, logistic

Table 1 List of reviewed papers concerning model reduction techniques or approaches for large-scale problems in various applications

Article	Application	Approach
Lam et al. (2011)	Biomass supply chain	Graph connectivity reduction based on distance, removing unnecessary variables
Anderson et al. (2011)	Biology networks	Model decomposition, Reduction to underlying system
Kim et al. (2014)	Heat exchangers	Decomposition of finite elements method, zones with different detail
Ishizaki et al. (2014)	General graph structures	Clustering algorithm based on Krylov subspaces
Heo and Daoutidis (2015)	Chemical process network	Decomposition, aggregation of specific multiple cycles
Hajikolaie et al. (2015)	General large-scale optimization	Decomposition supported by quantification of variable correlation between variables
Li and Balakrishnan (2016)	Vehicle routing problems	Minimizing the sum of the shortest path between nodes
Jansen and Durlofsky (2017)	well control optimization	Orthogonal decomposition and trajectory piecewise linearization
Xie et al. (2018)	Dynamical systems, viscoelastic materials	Relative error indicator, adaptive second order Arnoldi method
Martin et al. (2019)	Urban traffic network	Scale-free graph, Reduction based on optimization special criteria, merging nodes
Bagattini et al. (2019)	Large scale geometrical optimization	Clustering methods to identify regions of local optimum as starting points for optimization
Basán et al. (2019)	Scheduling in shipbuilding industry	Decomposition
Luan et al. (2020)	Railway networks	Decomposition, distributed optimization approaches
Zhang et al. (2020)	Optical networks	Machine learning overview, various technique of learning
Gao et al. (2020)	Dynamical systems, fluid mechanics	Neural network, approximation
de Bona et al. (2021)	Public transportation systems	Skeleton extraction, detecting similarities
Orkun Baycik (2021)	Maximum flow network interdiction problem	Machine learning, decision tree and random forest regression
Ban and Yamazaki (2021)	Aerodynamic shape optimization	Clustering/classification methods and exploration strategy, machine learning

Table 1 (continued)

Article	Application	Approach
Zheng et al. (2021)	Transportation and resource allocation problems	Dual decomposition, clustering methods, subgradient method, parallel optimization
Meselhi et al. (2022)	Benchmark functions	Decomposition based on minimizing of common variables using interaction matrix
Morabit et al. (2022)	Routing problems	Machine learning to reduce size and speed up heuristic for pricing problem
Zeng et al. (2022)	General approach. Testing four case studies	Edge shedding, vertex degree distribution

tasks, and theoretical graphic structures are included. The aim of these approaches is to group points or edges and replace them with one element in the system, mainly based on the similarities of different parameters. There is also a decomposition of the system, which leads to the division of the structure into several subtasks, which are solved separately. Subsequently, they can be connected using various techniques with iterative calculations, or solutions from sub-regions can be used as a starting point for optimization, which can significantly reduce computing time.

Due to the topic of this paper being related to the network flow problem, it is necessary to focus primarily on articles dealing with the reduction of variable models linked to logistics or supply chains. There are often algorithms that try to identify and remove unnecessary variables based on certain parameters. The article (Lam et al. 2009) uses the P-graph framework, which generates the maximum structure or set of all feasible solutions, in order to minimize the energy cost and carbon footprint. Thus, before the calculation itself, some solutions are omitted due to the nature of the problem in the context of the given limitations or disadvantages, and, in addition, the trade-off between results accuracy and computing time is favorable. However, the main goal of the new approaches is a further reduction consisting of the selection of only advantageous structures with respect to the solved problem. This was done in a study (Lam et al. 2011) by simply limiting the edges to a certain distance and eliminating unnecessary variables that are meaningless in terms of the role of flow in the network (zero production or capacity).

The article (Martin et al. 2019) focuses on graph reduction while maintaining appropriate properties such as connectivity and eigenvector centrality. The authors conceived the reduction as an optimization problem, within which some nodes are aggregated based on the similarity of parameters. The result of their approach is a call as *scale-free* graph, which can be characterized as a graph that has few nodes with many connections and many nodes with only a few edges. This interpretation corresponds exactly to most allocation tasks and links between producers and processing facilities. There are

usually only a few processing facilities, as there is an effort to centralize the process and maximize efficiency.

The recent study (de Bona et al. 2021) related to reducing methodology for transportation-based networks presented algorithms, which extracted the skeleton of the original graph suitable for the next analysis. It was focused on 1- and 2- degree nodes, which can represent blind roads or single lines and can be aggregated and substituted by one node. The presented technique can be used for weighted and unweighted graphs and as a part of a complex study, the impact of the reduction technique on a graph can be measured with defined coefficients. A similar principle of graph size reduction technique was presented in the study (Zeng et al. 2022). It was developed to effectively storing of large-scale data to speed up solving algorithms. Two presented approaches for edge identifying were presented, which were retaining the expected node degree distribution and preserving network topology. They were also suitable for various types of graphs, even the bipartite. On the other, some graph modifications are not suitable for all network flow problems, especially the removal of a 2-degree node, because, for example, the flow of a commodity is often unknown and the production (demand) cannot be properly assigned to other nodes.

ML and neural networks also find their place in the field of solving complex large-scale problems. According to a study (Musumeci et al. 2019), these approaches have five typical applications, classification, regression, anomaly detection, clustering, and dimensionality reduction. The representation of these techniques is growing nowadays, as their potential in the case of large data processing is great. The study (Orkun Baycık 2021) dealt with ML to solve maximum flow network interdiction problems through a defined network, which are nondeterministic polynomial-time competed and in the case of large-scale tasks usually unsolvable in a reasonable time. The methodology used decision tree and random forest (RF) regression to learn the interdictor's decision patterns for subsequent prediction. Same as other studies dealing with large-scale problems, ML is used directly to solve the original problem. However, ML as the supporting tool for optimization was rarely used. The application of ML in the field of optimization is described in a review (Gambella et al. 2021). The study points to the issue of large-scale tasks, where heuristic methods are widely used. At the same time, the problem of data classification is highlighted, which can be an important part of pre-processing optimization tasks. There is also mentioned a challenge in the form of linking ML and optimization in practical applications. The paper (Morabit et al. 2022) aims to improve the performance of an existing heuristic (column generation) for routing and scheduling applications by using ML. In these applications, the problems are mostly defined on a network, and decomposition is used to reduce computational cost. Various classification models were tested by (Morabit et al. 2022) to reduce the size of a network of a subproblem of column generation called the pricing problem. RF classifier was found to be the best choice. The solution was able to save up to 40% of computational time.

It seems that ML can bring substantial improvement in the field of flow/network problems. Especially subfield of ML called supervised learning can be helpful. There are two kinds of problems in supervised learning, regression, and classification. In most cases, the algorithms are basically the same for both tasks and the main difference is the loss function used. The most traditional methods are linear regression for

regression tasks and logistic regression for classification tasks. These traditional statistical methods are usually a good choice for simple problems. However, nowadays the problems under scrutiny are usually quite difficult to deal with as they try to describe very complex phenomena. Sometimes, other simple models like nearest neighbors, support vector machines, or decision trees can perform well, but various kinds of ensemble methods or neural networks are understood to be state-of-the-art methods. Generally, ensemble methods (usually based on decision trees) are considered to be easier to tune than neural networks and thus are very popular and top performing in various real-life applications (Kalmar et al. 2016; Hastie et al. 2017).

3 Novelty of the contribution

From the researched studies and comprehensive revision papers, it was clear that the current effort of the authors is to approximate graphic structures using aggregations, which are based on cluster analysis and similarities in the system. The decomposition of the problem and the solution of partial subproblems are also widely utilized, but its application is highly dependent on the nature of the investigated problem. Regarding the detection of advantageous structures in the graph, the studies focus mainly on the analysis of the basic parameters of its elements. In the area of a network flow problem and a supply chain system, there is a potential research gap in the context of the identification or classification of variables. It is therefore desirable to focus on its further development. In particular, it offers the possibility of a combination of optimization and ML, which, as can be seen from the literature review, can provide the necessary apparatus for solving large-scale applications.

The contribution of the presented paper lies in a new approach to solving and optimizing network flow problems, which by their size and complexity exceed today's limits of computing power. The goal is to remove the edges or arcs between the nodes on the basis of several key parameters of the solved allocation or location problem and thus reduce the number of variables in the model. The principle of the approach for oriented arcs is illustrated in Fig. 1. When solving a large-scale problem, the optimization with maximum detail of input data is usually unfeasible due to high computation demands. However, the same problem can be efficiently solved with aggregated data. Such results can serve as a training set for ML because the essence of the task should be preserved. On the basis of key features, each arc of the original problem can be classified if it belongs to an optimal solution or not. This information can be used for model size reduction and to remove unnecessary arcs.

Unlike other approaches, there is no effort to directly select optimal arcs, but find those arcs, which are not in optimal solution with high probability. ML is used to classify the arcs, where the learning process is ensured with the help of smaller model tasks by a large range of generation of individual parameters. The advantage of the approach is that there is no aggregation or decomposition into subproblems. This makes it possible to avoid misinterpretations of the solutions found and their combination, which can result in a big mistake. Similar approach was identified in the study (Morabit et al. 2022) in order to speed up column generation algorithm. The main difference lies in reduction of generated subproblems, while newly presented approach deals

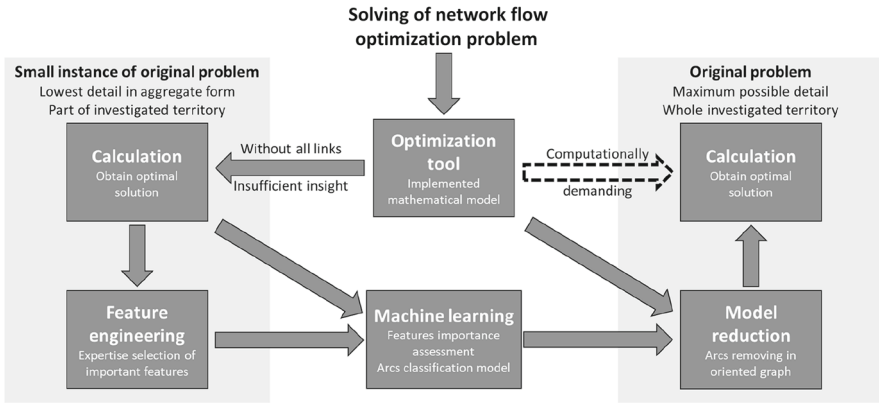


Fig. 1 The principle of the presented approach to solving large-scale network flow problems

with reduction of the original task using exact optimization algorithm. As another advantage can be mentioned the data collection for ML, where instead of solving many instances to optimum before further application, only one optimization process is needed. However, this fact can be influenced especially by different complexity of routing problems and location task.

The goal of the paper lies in identifying key parameters of arcs in problems inspired by WM, proving their importance, and verifying the approach based on learning in a small-scale task with further application on large-scale task. In the following sections, the process of ML is introduced with the description of obtaining suitable data. The importance of parameters (features) is individually evaluated. For the verification of the statistical model, the case study of WM in the Czech Republic is presented.

4 Data set for machine learning

A sufficient amount of data is needed for the ML model to be meaningful. After the learning process, it will be possible to identify unnecessary variables in the real study, which creates an unnecessarily large task. The approach is mainly aimed at reducing the number of arcs in the bipartite network. All arcs are therefore directed and go from the producer to the processing facility. For the needs of ML, a small model task was created, which can be efficiently solved and thus obtain the necessary data set. It is an allocation task, where the goal is to process all produced commodities in selected facilities. The task is based on the principle of network flow and the area of application is focused on WM. However, the approach can be applied to any field, but it is necessary to summarize important parameters from the area, see monitored parameters below. The ML allows the identification of key features of the arcs included in the optimal solution. On the contrary, this knowledge can also be used to identify arcs that are unlikely to belong to the optimal solution. The goal is to create an ML model, which enables to classify arcs and recommends their removal or not.

4.1 Generation of small model task

The artificial optimization problem setup is inspired by the WM case. The aim is to transport all waste from producers (municipalities) to processing facilities (Waste-to-Energy plants or landfills). The goal is to minimize the total costs, which include the transport costs and the cost of processing in the plant. The mathematical formulas in the optimization model are only linearly dependent on the amount of waste, which is transported or processed in the plant. The objective function is defined by Eq. 1 as follows

$$\min \sum_{i \in I} \sum_{e \in E} F_i M_{e,i} z_e + \sum_{e \in E} G_e D_e z_e, \quad (1)$$

where F_i stands for processing cost per tonne of waste in facility i , G_e is transportation cost per tonne of waste and kilometer, D_e is arc length, $M_{e,i}$ is incidence matrix and variable z_e represents the amount of waste transported along arc e . The first formula in the objective function is related to the total processing cost of waste, meanwhile, the second formula represents the total cost of transporting the waste from all producers to the processing facility. There is only one constraint defined by Eq. 2, which corresponds to the capacity limitation of each facility.

$$P_i + \sum_{e \in E} M_{e,i} z_e \leq C_i, \quad \forall i \in I, \quad (2)$$

where parameter C_i stands for the maximum capacity of processed waste in node i and parameter P_i for amount of waste produced in node i .

The size of the task is set at 206 nodes, which correspond to the infrastructure of the micro-region in the Czech Republic. The task is thus based on a real traffic network, which ensures a real link to transport costs. In order to ensure sufficient diversity of the obtained data, a generation of different scenarios is approached, which are different in average values of some parameters. Overall, 4 key parameters with two different levels are selected for the generation. They are defined as followed:

- Number of processing facilities: 10 or 100
- The ratio of total capacity and production: 1.01 or 1.1
- Average processing cost: 40 EUR/t or 120 EUR/t
- Average transportation cost: 0.5 EUR/t/km or 1 EUR/t/km

Thus, a total of 16 scenarios are generated covering all combinations of both levels of the above parameters. The individual parameters for each element in the system were generated from the Beta distribution (see Eqs. 3 and 4), which with its flexibility can create a symmetrical distribution, as well as negative and positive skewness. It is defined as

$$f(x; \alpha, \beta) = \frac{x^{\alpha-1} - (1-x)^{\beta-1}}{B(\alpha, \beta)}, \quad (3)$$

Table 2 The settings of beta distribution for scenario parameters generation

Scenario parameter	Alpha	Beta
Production	20	100
Capacity (10 facilities)	100	20
Capacity (100 facilities)	20	100
Processing cost	100	100
Transportation cost	100	100

where

$$B(\alpha, \beta) = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha + \beta)}, \quad (4)$$

where Γ is the gamma function. A symmetrical distribution is assumed for transport and processing costs. Positive skewness is used to generate production (it was the same for all tasks, only the ratio of total capacity and total production changed). In the case of a larger number (100) of small processing facilities, a positive skewness is also considered. On the contrary, in the case of a small number (10) of larger facilities, a negative skewness is considered. The change in skewness should better describe reality instead of its preservation. In addition, the Beta distribution settings made it possible to ensure that, for example, capacity and production are non-negative. The Beta distribution parameters for the generated task parameters are summarized in Table 2. Parameters from Table 2 are not empirically estimated, they only have to capture the probable type of skew of a given distribution.

4.2 Selected features

Within the monitored parameters, or features used to classify individual arcs, various characteristics are selected that apply to both individual network components (nodes, arcs) and values describing the entire task. The main indicators can include capacity C_i , production P_i , and price per tonne of waste (processing cost F_i and transportation cost $G_e D_e$), which are subject to various statistics and ratios. Abbreviations in brackets are stated only in that case when the feature is further mentioned in the ML process or discussion of results.

4.2.1 Characteristics of the system

- Total cost [EUR]–median, deviation (S_TotC_var), skewness (S_TotC_skew)
 - The data set consists of all arcs. Each arc is assigned the sum of processing and transportation cost per tonne of waste.
- Processing cost [EUR]–deviation

- o The data set consists of all processing facilities, where only processing costs are included.
- Average total cost related to processing facilities [EUR]–median
 - o The data set consists of all processing facilities
 - o Each processing facility is assigned the average of the total cost per tonne of waste (sum of processing and transportation cost) over all possible producers.
- Production [t]–average, median, deviation
 - o The data set consists of all municipalities where the amount of waste for processing is included.
- Capacity [t]–average, median, deviation
 - o The data set consists of all processing facilities, where the capacities are included.
- Ratio of overall capacity and production [–]
 - o The sum of all available capacity is divided by the total production in the system.
- Ratio of number of facilities and municipalities [–]
 - o The number of processing facilities is divided by the number of municipalities.

4.2.2 Characteristics of the nodes in the system

- Total cost of node [EUR]–deviance (N_TotC_var), skewness, kurtosis (N_TotC_kurt)
 - o The data set consists of all arcs related to selected node. Each arc is assigned the sum of processing and transportation cost per tonne of waste.
- Production density related to municipalities ($N_Prod_Mun_den$) [t/km]
 - o It expresses the concentration of waste in the vicinity of a selected municipality.
 - o The sum of all municipalities, which production is divided by the distance between a selected municipality and other municipalities.
- Production density related to facilities ($N_Prod_Fac_den$) [t/km]
 - o It expresses the concentration of waste in the vicinity of a selected processing facility.
 - o The sum of all municipalities, which production is divided by the distance between selected processing facility and other municipalities.
- Capacity density related to municipalities ($N_Cap_Mun_den$) [t/km]
 - o It expresses the concentration of processing capacity in the vicinity of a selected municipality.
 - o The sum of all processing facilities, which capacity is divided by the distance between a selected municipality and a processing facility.

4.2.3 Characteristics of the arcs in the system

- Total cost of arc (A_TotC) [EUR]
 - The sum of processing and transportation cost per tonne of waste along a selected arc.
- Ratio of total cost of arc and the most expensive arc ($A_TotC_TotCMax_ratio$) [-]
 - The sum of processing and transportation cost per tonne of waste along a selected arc divided by the total cost of the most expensive arc in the system.
- Transportation cost (A_TrC) [EUR]
 - The transportation cost per tonne of waste along a selected arc.
- Ratio of transportation and processing cost ($A_TrC_ProcC_ratio$) [-]
 - The ratio of transportation cost and processing cost at a processing facility of a selected arc.
- Ratio of capacity and production ($A_Cap_Prod_ratio$) [-]
 - The ratio of processing facility capacity and municipality production of selected arc connecting these two nodes.
- Relative order of arc [-]-municipality ($A_TotC_Mun_rank$), processing facility ($A_TotC_Fac_rank$)
 - Each arc is assigned the order of total cost in a specific set. The order is normalized by the cardinality of the considered set.
 - Municipality or processing facility considers the order within arcs connected to them.
- Gradient of total cost related to municipality (A_TotC_grad) [-]
 - The sum of processing and transportation cost per tonne of waste along a selected arc. Evaluated set contains arcs connected with only one specific municipality.
 - Each arc is assigned a value between 0 and 1. The value 0 represents the cheapest arc and the value 1 the most expensive one.
 - Other arcs are assigned the value according to a position of total cost between mentioned extreme values with linear dependency.
- Gradient of the quotient total cost and capacity ($A_TotC_Cap_grad$) [-]
 - The sum of processing and transportation cost per tonne of waste along a selected arc is divided by the capacity of the processing facility. Evaluated set contains arcs connected with only one specific facility.
 - Each arc is assigned a value between 0 and 1. The value 0 represents the arc with the lowest value and the value 1 the highest one.
 - Other arcs are assigned the value according to a position of the quotient between mentioned extreme values with linear dependency.

4.3 Machine learning

RF is selected as a suitable classification model for this task. since it is well know for its robustness, and good performance even with little to no need for parameter tuning (Kalmar 2016; Hastie et al. 2017). RF were also identified as most suitable by the study (Morabit et al. 2022) in similar application. Let's consider a classification problem with binary response Y and inputs X_1, \dots, X_p . The basic idea of RF is that, first, B Bootstrap samples are created from a training set, second, each bootstrap sample is used to grow a single tree (weak learner), while in each split only a random subset of features is used. Each of these individual trees partitions features space into subspaces (R_m) and then assigns a constant c_m to R_m (see Eq. 5) using indicator function I . In every node split (into regions R_1 and R_2) is done using the criterion from Eq. 6, where j is a variable and s is a split point. When making a prediction, data are passed to every single tree, in classification task B predictions (C_b) are obtained and RF prediction (\hat{C}) emerges by majority vote (see Eq. 7). For more details see (Hastie et al. 2017). The modeling part of this article is done in R language (R Core Team 2021) using *randomForest* library.

$$\hat{f}(X) = \sum_{m=1}^M c_m I\{(X_1, \dots, X_p) \in R_m\} \quad (5)$$

$$\min_{j,s} \left[\min_{c_1} \sum_{x_i \in R_1(j,s)} (y_i - c_1)^2 + \min_{c_2} \sum_{x_i \in R_2(j,s)} (y_i - c_2)^2 \right] \quad (6)$$

$$R_1(j, s) = \{X|X_j \leq s\}, R_2(j, s) = \{X|X_j > s\}$$

$$\hat{C} = \text{mode}\{C_b(x)\}_{b=1}^B \quad (7)$$

The classification task here is to assign a probability (class) to every arc based on its properties and the properties of the whole scenario. There are only 2 classes: 0—an arc is not present in the optimal solution, and 1—an arc is present in the optimal solution. By the nature of the network flow problem, most of the arcs are not used. Actually, from generated scenarios, only about 2% of arcs belong to class “1” (i.e. they are present in optimal solution).

Note that setting of the parameter *sampsiz*e is of great importance in the case of highly unbalanced classes (as in our case). This parameter allows accounting for the imbalance of the dataset by specifying how many cases of each class (here 0/1) should be used. E.g. when the default setting is used, out-of-bag (OOB) general classification error (i.e. error on unseen data) was as low as 1%, but this error reached almost 60% for class “1” (i.e. used arcs). Contrary, when RF is forced to give the same importance to both groups, OOB general classification error grows to 4%, but for class “1” is about 0.2%.

First, the most important features are identified by running RF on the training sample (70% of all cases used) from all created scenarios. Feature importance is

determined via Gini Index (see Fig. 2). It can be seen that features related to the price of the arc (actual cost or cost rank) are among the most important features. On the other hand, features related to waste generation, capacity of facilities, and general information about costs are of low importance. Only top six features are kept to reduce the complexity of an ML model. This helps to prevent the overfitting of a model. The exact number of features or threshold for Gini mean decrease is arbitrary, however since e.g. N_Cap_Mun_den, A_TrC_ProcC_ratio and A_TotC are quite similar, it is usually not a good idea to keep them all in a final model.

Second, a set of scenarios with 10 big facilities was used for the creation of a classification model. The setting of the classification model was the same as in the previous case. The classification model was then utilized to reduce the number of

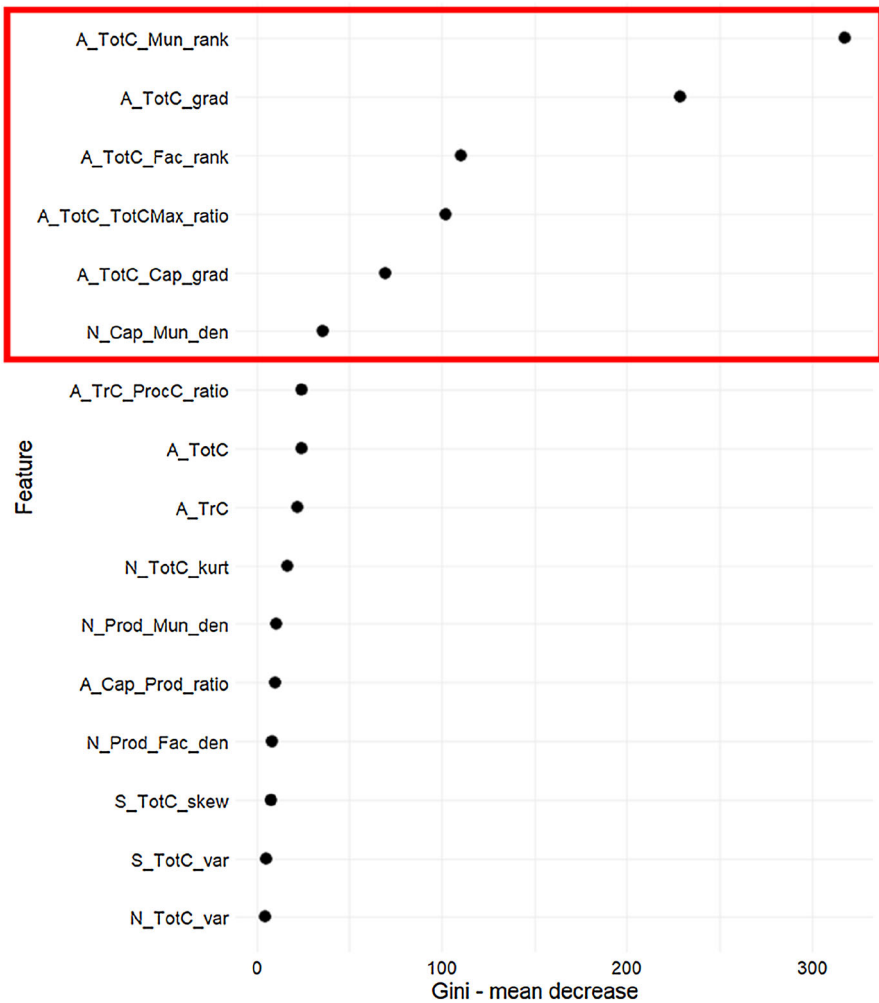


Fig. 2 The importance of all features via Gini Index

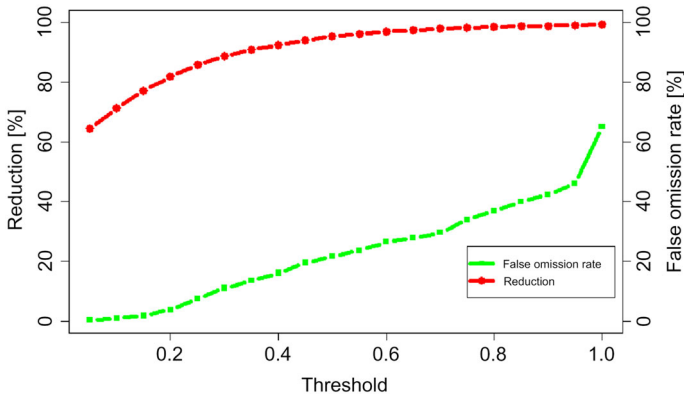


Fig. 3 False omission rate and arcs reduction dependency on the threshold

inputs for scenarios with 100 processing facilities (these have 10 times more arcs). The results of this approach are provided in Fig. 3.

It shows the false omission rate (FOR), i.e. % of arcs from optimal solution excluded by the classification algorithm, and reduction (% of all arcs excluded) for a given threshold. They are defined by Eqs. 8 and 9. It can be seen that for example when the threshold is set to 0.2, more than 80% of inputs are excluded at cost of a very small (1.4%) loss of arcs from the optimal solution.

$$FOR = 100 \frac{\text{number of arcs from optimal solution excluded by RF}}{\text{total number of arcs in optimal solution}} \quad (8)$$

$$Reduction = 100 \frac{\text{number of arcs excluded by RF}}{\text{total number of arcs}} \quad (9)$$

Figure 4 shows kernel density estimates of predicted probabilities for both classes of arcs. The concentration of the majority of probability mass near 0 for class and close to 1 for class shows that in most cases, “promising” and “non-promising” arcs are easy to distinguish.

The original task can be significantly reduced without removing optimal arcs and preserving insights into the solved problem with reduced computational demands. However, the previously presented results reflect the effort to create a general ML model for the arc classification of any network flow WM problem with different parameters. The performance of the approach can be further enhanced, if it is applied only to one problem using calculations with different detail, but with the same or aggregated input parameters.

5 Application on real case study

To verify the presented approach, it is applied to a real study of WM in the Czech Republic. In the Czech Republic, municipalities are responsible for WM and set

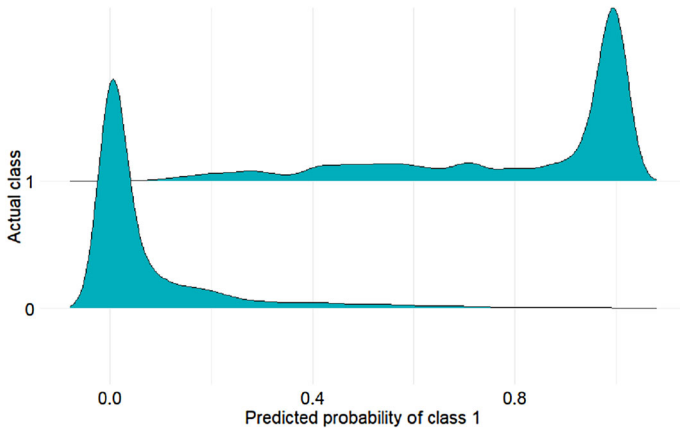


Fig. 4 Kernel density estimates of predicted probabilities for both classes of arcs

charges for waste to citizens in their cadastral area. Due to the new laws regulating the legislation in connection with the achievement of the set goals of the EU within the circular economy, the possibilities of landfilling and its price are gradually regulated. At the same time, waste treatment alternatives, especially recycling, are supported. However, not all waste can be materially recovered, so energy recovery should be also part of WM and the necessary infrastructure should be appropriately planned.

The presented case study is devoted to this problem, which is already outlined in previous study (Pluskal et al. 2022). The corresponding mathematical model with notations is available in Appendix. The goal is to process waste from all municipalities, while landfills and facilities for energy recovery, called Waste-to-energy (WtE) plants, are considered in the system. The types of waste taken into account include mixed municipal waste (MMW), bulky waste, and residuals after sorting the plastic and paper waste, which are not suitable for recycling, but their energy potential is desirable to use. The task deals only with waste for energy recovery and therefore the system does not include facilities providing material use in any form. The objective function is to minimize the total cost of allocating and processing all produced waste including investments in new WtE plants. Decisions such as the construction of WtE plants, which cannot be easily modified on a monthly or yearly basis, are classified as strategic decisions and represent the first stage of the model. Operational decisions, on the other hand, pertain to the planning of waste collection and the operation of individual facilities, which can be adjusted or modified based on current conditions. Strategic decisions are generally deemed more critical in WM planning.

There are 206 micro-regions with a total of 6258 municipalities in the Czech Republic, which produce the mentioned 3 commodities in a defined system. The forecast of waste production for the year 2030 is taken, in which it is possible to assume the completion of expensive facilities for energy use, which the optimization model will propose. Due to the uncertainty of future development and the possibility of various interventions in the legislation, two different production scenarios are expected. One scenario is defined as base scenario and another represents projection with different

rates of meeting the set goals. The forecasted data set is obtained by methodology presented in the paper (Smejkalová et al. 2022) using freely available data about Czech WM (ISOH 2022). It should be noted that more scenarios should be included to obtain robust results, but their number is limited to ensure solvability in a reasonable time due to many testing calculations.

Within the processing facilities, 101 WtE plants are considered, of which 4 are already existing. Other potential locations for new WtE plants are appropriately selected and targeted to places with a high demand for heat, which together with electricity production represent a significant revenue component. The selected micro-regions have at least 30 thousand inhabitants, which should satisfy the minimum heat demand. Based on the techno-economic model (Ferdan et al. 2015), the gate fee is determined, which is a condition for self-sufficiency and return on investment. In addition, more capacity options are planned in each location, which is reflected in the final price for waste treatment. The resulting price is thus a non-linear function depending on capacity. This dependency is implemented via SOS variables type 1 (Williams 2009), when 5 different capacities of each WtE plant are considered.

For optimal operation of a WtE plant, it is necessary to further ensure a suitable calorific value of incinerated waste, i.e. within the network flow task, not only quantitative indicators are met, but also qualitative ones. Therefore, each waste stream has a defined approximate calorific value, which has been determined within the recycling rate estimates in each municipality. Each WtE plant has a defined strict range to the calorific value of an incinerated mix of waste.

Other facilities in the system are cement plants, which can use high-calorific waste suitable for recycling as fuel. There are 5 of these facilities existing in the Czech Republic, while their annual capacity is considered to be only half due to the possibility of industrial waste use. However, the processing price is set to zero. WtE plants and cement plants are complemented by 115 landfills, which copy the existing infrastructure. The price for landfilling takes into account the rising costs, including environmental impacts, as the deposited waste can produce up to 1 kg of CO_{2eq} per kg of waste (Ferdan et al. 2018). This fact was recalculated in the context of finance according to emission allowances.

The transportation costs are in non-linear dependency on the overall traveled distance along an arc and they are calculated according to the study presented in the study (Gregor et al. 2017). The nonlinearity can be implemented without any additional integer variables thanks to the bipartite graph and it can be included already in pre-processing part. Within the permitted flows, it is necessary to state that the current legislation in the Czech Republic is going to prohibit the landfilling of high calorific waste above a certain calorific value after 2030 (Laws for people 2020). This fact is difficult to verify, but some waste types can be identified for which landfilling is going to be prohibited. In the context of this task, it is thus possible to process bulky waste only in WtE plants, as its calorific value exceeds the permitted limit. The residuals from waste sorting can only be processed in cement plants or WtE plants. MMW can be used for energy recovery in a WtE plant or deposited in a landfill.

Due to the nonlinear gate fee, which is modeled using SOS1 variables, the optimization model is defined as mixed integer programming. Such a model makes it possible

Table 3 Comparison of results with different detail of the solved problem

Detail of solved problem	Number of equations/variables in the model	Computing time	Objective value [thousand EUR]	WtE capacity [kt]	WtE locations/optimal location
Micro-regions	5,719/177,307	31 s	501,993	2,981	16/15
Municipalities	42,031/5,297,299	11,019 s	509,181	3,261	18/18

to maintain the properties of linear optimization, especially concerning the information of results optimality. The mathematical model for the network flow problem is implemented in GAMS (General Algebraic Modelling System) and the solver CPLEX (GAMS 2022) is used. A common computer with an Intel i7-11700U processor with 32 GB of RAM was used for the calculation.

A key part of verification is to show that the detail of the problem affects the results. The task is first solved on the micro-region transport network, within which aggregation of input data is performed. Afterward, the original problem with maximum detail is optimized. The parameters of both models and results are stated in Table 3.

The model of the original task contains a total of over 2.6 million arcs, and since 2 scenarios are solved, the number of variables corresponds to two times as much. It is a 30 times larger optimization problem than the aggregated task, which has only 88 thousand arcs. The impact of selected detail on the results can be clearly observed from strategic decisions related to WtE plants. The aggregated form of the modeled problem suggests 16 WtE plants with a total annual capacity of 2,981 kt. However, only 15 of them correspond to the same locations suggested by the optimization of the original problem. In other words, three optimal locations are omitted and one WtE plant is planned in a bad place.

The following part is devoted to improving the computation performance of the original problem. For comparison purposes, the removal of the longest arcs is used first, as it can be assumed that, by the nature of the problem, shorter arcs are more likely to be used. The following Table 4 shows the individual calculations, where the original problem is stated first and then the arcs of the model are gradually reduced according to the selected percentage reduction.

The methodology employed for arc removal based on distance represents a straightforward approach that does not require additional complex techniques, as evidenced by the results obtained. By applying this approach, it is possible to achieve a reduction of 60% in the original problem without significantly affecting the solution, particularly in the case of strategic decisions. However, it is important to note that further reduction beyond this point can lead to undesired consequences such as the mislocalization of Waste-to-Energy (WtE) plants and an increase in computation time. Therefore, a reduction of up to one-third can be considered a reasonable threshold to maintain a balance between reduction benefits and potential drawbacks.

The previous simple methodology of removing arcs by distance does not seem to be enough effective. Therefore, the newly developed approach using ML is further used

Table 4 Impact of the removing arcs by maximum distance on the results

Maximum distance	Arcs in the model	Optimal arcs removed	Computing time	Change of objective value	WtE capacity [kt]	WtE locations/optimal location
612 km	100%	0/0%	11,019 s	0%	3,261	18/18
297 km	80%	0/0%	13,061 s	0%	3,261	18/18
222 km	60%	0/0%	10,115 s	0%	3,261	18/18
163 km	40%	29/0.14%	4,810 s	9E-4%	3,251	18/18
106 km	20%	1,491/7.3%	10,927 s	0.26%	3,291	20/18
72 km	10% *	5,546/27%	36,000 s	2.38%	2,951	32/14
50 km	5% **	10,931/53%	–	–	–	–

*The optimization of the reduced problem reaches the time limit of 36,000 s

**The reduced problem is infeasible

to verify the contribution. The results from optimization in the detail of micro-regions will provide the necessary data for the classification model. The features taken into account remain the same, i.e. the 6 most important. Values of features for individual arcs are calculated using parameters from the base scenario. The classification model is created for each type of waste separately. The individual waste streams and their respective optimal arcs may differ in their characteristics and one model could be considered less significant than three separate models. This approach, therefore, reduces parameter uncertainty in the classification process, which could lead to bad evaluation. Subsequently, the created model was applied to the infrastructure of municipalities. The classification process takes 30 s. The probability density is shown in Fig. 5.

The ML model classifies the most of arcs, which are situated in the first peak of the density to the threshold equal to 0.05. Another significant peak can be seen on the

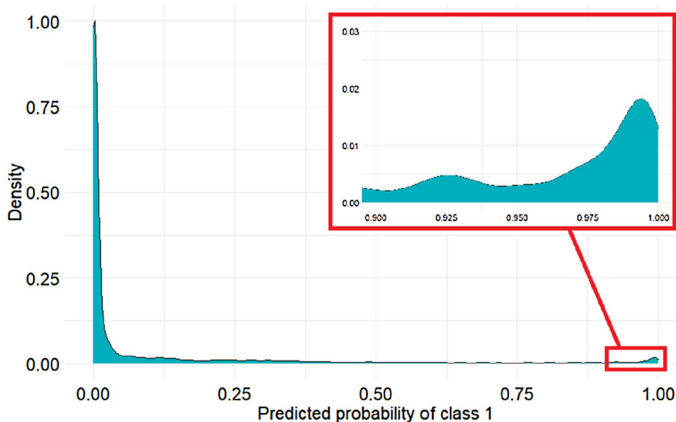


Fig. 5 The probability density of classified arcs in the case study

Table 5 Impact of the removing arcs by classification threshold on the results

Threshold	Arcs in the model	Optimal arcs removed	Computing time	Change of objective value	WtE capacity [kt]	WtE locations/optimal location
0	100%	0/0%	11,019 s	0%	3,261	18/18
0.015	20%	0/0%	6,062 s	0%	3,261	18/18
0.045	15%	3/0.015%	3,091 s	1E-5%	3,261	18/18
0.177	10%	15/0.074%	1,318 s	6E-5%	3,261	18/18
0.425	5%	144/0.71%	718 s	9E-4%	3,261	18/18
0.549	4%	268/1.3%	569 s	0.32%	3,281	20/18
0.749	3%*	638/3.1%	218 s	0.24%	3,301	19/17
0.93	2%*	2,042/10%	85 s	2.65%	3,421	21/17

*At least one arc to each type of permitted facility for each wast

other side of the graph around probability 1. The optimal arcs should take place in the detailed section. The proportion of these mentioned areas is diametrically different due to the ratio of total arcs in the system and the optimal ones. As the next step, it is important to verify, how many optimal arcs are removed and how the model and results are affected. Partial calculations with different threshold values are presented in Table 5.

The results show that the model created on micro-regions identifies the unnecessary arcs better than the previous removal based on distance. The mentioned peak in Fig. 5 has almost no optimal arc in the probability range up to 0.05. The size of the model can thus be reduced six times without changing the objective function and reducing the computational time almost to one-fourth. Further reduction is still possible, but a larger proportion of optimal arcs is being removed. However, more favorable arcs are left in the system and thus there is no significant change in the purpose function. The maximum possible reduction is achieved around a threshold of 0.5, where the model is approximately 50% larger than the micro-region model, taking into account a transport network 30 times larger. Further reductions no longer have a solution, mainly due to the removal of all arcs to some municipalities. Therefore, the condition for arc removal is added and consists of the rule, when at least one arc to each type of processing facility must be preserved. With this modification, the original problem can be further reduced, but the localization of WtE plants is affected.

It is important to acknowledge that the complexity of this task is not large and was used for testing purposes. It is necessary to consider more types of waste at once and to include additional waste treatment infrastructure, such as transfer stations. The importance of model reduction increases with the scope of tasks, especially with the number of processing facilities. Each municipality usually uses only one arc for one waste commodity, meanwhile, the original task contains an arc from a selected municipality to every permitted facility. Therefore, it can be expected that if the potential WtE plant locations were in every micro-region, the reduction could produce an even smaller optimization problem than the calculation with the detail of micro-regions.

Furthermore, it is necessary to consider numerous scenarios, which can potentially reach hundreds in number. All these aspects significantly increase the number of variables in the model, which subsequently cannot be solved through computational time or memory requirements. In such cases, this presented approach finds justification and utility.

6 Conclusion

The presented paper deals with the reduction of large-scale tasks. The subject of research is network flow tasks, especially in WM. Unlike existing studies, this research focuses on the identification and removal of unnecessary arcs. This is done by the classification model (namely random forest), which assigns to each arc a probability with which it occurs in the optimal solution. The cornerstone of this approach is an appropriate set of features, which sufficiently describes the nature of the system. The goal is to achieve maximum reduction with minimal error to make large-scale tasks solvable in a reasonable time. By reducing the memory and computational requirements, detailed studies can be completed in a fraction of the original time, allowing for faster feedback.

The key mechanic of this approach lies in the usage of ML on a small task followed by application on a large-scale detailed task. It is important to note that each network flow problem is unique, making it highly ambitious and practically impossible to create a general model for variable reduction. The results from testing instances related to WM show that ordered and ratio indicators play a crucial role, especially with a link to the price and available capacity. The approach is verified on the real case study from WM in the Czech Republic. The learning process is performed on results from the model based on micro-region infrastructure and afterward applied original network with the lowest administrative units—municipalities. The case study demonstrates that the model can achieve a reduction of nearly 95% without significantly impacting the objective function. Importantly, this reduction does not introduce any changes to strategic decision-making concerning the localization of WtE plants.

The presented approach is demonstrated on network flow problems, which can cover many real-world applications, especially in logistics (collection, distribution) and infrastructure planning (location and allocation problems). However, the methodology's underlying idea can be applied to any task where the number of nonzero variables is only a small fraction of the total variables involved. In the future, it is crucial to expand the range of monitored parameters, taking into consideration the specific challenges of real-world problems being addressed. Unusual constraints or nonlinearities, which must be taken into account in real models, can also be problematic. Additionally, it is important to conduct further analysis to determine which optimal arcs are removed first and the reasons behind it. These new insights can contribute to enhancing the efficiency of the developed approach.

Acknowledgements The authors also gratefully acknowledge the financial support provided by the Technology Agency of the Czech Republic, grant No. SS02030008 “Centre of Environmental Research: Waste management, circular economy and environmental security”.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

Appendix: Mathematical model for case study

The mathematical model for the verification of the presented approach was taken from the paper Pluskal et al. (2022):

Sets

E	Set of arcs between sources of waste and treatment facilities
F	set of potential locations for WtE plants
I	set of all nodes
$I^E \subset I$	subset of nodes including only WtE plants with corresponding capacities
$I_f^F \subset I$	subset of nodes including only WtE plants with capacities in location f
$I^P \subset I$	subset of nodes including only sources of waste
$I^T \subset I$	subset of nodes including only other types of treatment facilities
K	set of breakpoints for the piece-wise linear revenue function
S	set of scenarios

Parameters

$B_{i,j}$	Equal to 1 if there is a possible transfer of slag from the WtE plant in node j to a waste source node i from where the slag is transported to a landfill, and 0 otherwise, [-].
$C_{i,k}^{REV}$	revenue at plant i when heat power deviation is at breakpoint k , [EUR].
C_i^{TREAT}	treatment cost in the facility located at node i , [EUR/kt].
C_e^{TRANS}	transportation cost along arc e , [EUR/kt]
C_i^{WtE}	treatment cost at WtE plant at node i , [EUR]
G_i	treatment capacity of node i , [kt]
$H_{e,s}$	LHV of transported waste along arc e in scenario s , [-]
L_i^{MAX}	maximum LHV of incinerated waste in WtE plant i , [TJ/kt]
L_i^{MIN}	minimum LHV of incinerated waste in WtE plant i , [TJ/kt]
$M_{e,i}$	equal to 1 if arc e enters node i , equal to -1 if arc e leaves node i , and 0 otherwise, [-].
$O_{i,s}$	amount of waste at node $i \in I^P$ in scenario s , [kt].
P_s	probability of scenario s , [-].
Q_i^{MAX}	maximum heat power of WtE plant at node i , [TJ]
Q_i^{MIN}	minimum heat power of WtE plant at node i , [TJ]

- Q_i^{REF} preferred heat power of WtE plant at node i , [TJ]
 R_e coefficient of waste mass reduction by waste treatment assigned to arc e , [-]
 $U_{i,k}$ heat power deviation of WtE plant at node i and breakpoint k , [TJ].

Variables

- $x_{e,s}$ The amount of waste transported along arc e in scenario s , [kt]
 y_i binary decision variable, equal to 1 if WtE plant at node i is built, and 0 otherwise [-]

Special ordered sets, type 2

- $w_{i,k,s}$ Variable for the revenue function of the WtE plant at node i , when the heat power deviation is equal to breakpoint k in scenario s , [-]

$$\min \sum_{i \in I^E} C_i^{WtE} y_i + \sum_{s \in S} P_s \left(\sum_{e \in E} \sum_{i \in I^T} C_i^{TREAT} M_{e,i} x_{e,s} + \sum_{e \in E} C_e^{TRANS} x_{e,s} - \sum_{i \in I^E} \sum_{k \in K} C_{i,k}^{REV} w_{i,k,s} \right) \tag{10}$$

$$\sum_{i \in I_f^F} y_i \leq 1, \quad f \in F \tag{11}$$

$$0.5 y_i G_i \leq \sum_{e \in E} M_{e,i} x_{e,s} \leq y_i G_i, \quad i \in I^E, s \in S$$

$$\sum_{e \in E} M_{e,i} x_{e,s} \leq G_i, \quad i \in I^T, s \in S \tag{13}$$

$$O_{i,s} + \sum_{e \in E} \sum_{j \in I^E} R_e B_{i,j} M_{e,j} x_{e,s} + \sum_{e \in E} M_{e,i} x_{e,s} \leq 0, \quad i \in I^P, s \in S \tag{14}$$

$$\sum_{k \in K} U_{i,k} w_{i,k,s} = \sum_{e \in E} H_{e,s} M_{e,i} x_{e,s} - Q_i^{REF} G_i y_i, \quad i \in I^E, s \in S \tag{15}$$

$$\sum_{e \in E} H_{e,s} M_{e,i} x_{e,s} \leq Q_i^{MAX} G_i y_i, \quad i \in I^E, s \in S \tag{16}$$

$$\sum_{e \in E} H_{e,s} M_{e,i} x_{e,s} \geq Q_i^{MIN} G_i y_i, \quad i \in I^E, s \in S \tag{17}$$

$$\sum_{e \in E} H_{e,s} M_{e,i} x_{e,s} \leq \sum_{e \in E} L_i^{MAX} M_{e,i} x_{e,s}, \quad i \in I^E, s \in S \tag{18}$$

$$\sum_{e \in E} H_{e,s} M_{e,i} x_{e,s} \geq \sum_{e \in E} L_i^{MIN} M_{e,i} x_{e,s}, \quad i \in I^E, s \in S \tag{19}$$

$$\sum_{k \in K} w_{i,k,s} = 1, \quad i \in I^E, s \in S \tag{20}$$

$$w_{i,k,s} \geq 0, \quad i \in I^E, \quad k \in K, \quad s \in S \quad (21)$$

$$x_{e,s} \geq 0, \quad e \in E, \quad s \in S \quad (22)$$

$$y_i \in \{0; 1\}, \quad i \in I^E \quad (23)$$

References

- Anderson J, Chang YC, Papachristodoulou A (2011) Model decomposition and reduction tools for large-scale networks in systems biology. *Automatica* 47(6):1165–1174. <https://doi.org/10.1016/j.automatica.2011.03.010>
- Bagattini F, Schoen F, Tigli L (2019) Clustering methods for large scale geometrical global optimization. *Optim Methods Softw* 34(5):1099–1122. <https://doi.org/10.1080/10556788.2019.1582651>
- Ban N, Yamazaki W (2021) Efficient global optimization method via clustering/classification methods and exploration strategy. *Optim Eng* 22(1):521–553. <https://doi.org/10.1007/s11081-020-09529-4>
- Basán NP, Cóccola ME, García del Valle A, Méndez CA (2019) An efficient MILP-based decomposition strategy for solving large-scale scheduling problems in the shipbuilding industry. *Optim Eng* 20(4):1085–1115. <https://doi.org/10.1007/s11081-019-09457-y>
- de Bona AA, de Oliveira RM, Ono Fonseca KV, Lüders R (2021) A reduced model for complex network analysis of public transportation systems. *Physica a: Stat Mech Appl* 567:125715. <https://doi.org/10.1016/j.physa.2020.125715>
- Burre J, Bongartz D, Mitsos A (2022) Comparison of MINLP formulations for global superstructure optimization. *Optim Eng*. <https://doi.org/10.1007/s11081-021-09707-y>
- Ferdan T, Šomplák R, Závíralová L, Pavlas M, Frýba L (2015) A waste-to-energy project: a complex approach towards the assessment of investment risks. *Appl Therm Eng* 89:1127–1136. <https://doi.org/10.1016/j.applthermaleng.2015.04.005>
- Ferdan T, Pavlas M, Nevrlý V, Šomplák R, Stehlík P (2018) Greenhouse gas emissions from thermal treatment of non-recyclable municipal waste. *Front Chem Sci Eng* 12(4):815–831. <https://doi.org/10.1007/s11705-018-1761-4>
- Gambella C, Ghaddar B, Naoum-Sawaya J (2021) Optimization problems for machine learning: a survey. *Eur J Oper Res* 290(3):807–828. <https://doi.org/10.1016/j.ejor.2020.08.045>
- GAMS (2022) GAMS documentation. Solver manuals, CPLEX 12. https://www.gams.com/33/docs/S_CPLEX.html
- Gao H, Wang JX, Zahr MJ (2020) Non-intrusive model reduction of large-scale, nonlinear dynamical systems using deep learning. *Physica D* 412:132614. <https://doi.org/10.1016/j.physd.2020.132614>
- Gregor J, Šomplák R, Pavlas M (2017) Transportation cost as an integral part of supply chain optimisation in the field of waste management. *Chem Eng Trans* 56:1927–1932. <https://doi.org/10.3303/CET1756322>
- Hajikolaee KH, Pirmoradi Z, Cheng GH, Wang GG (2015) Decomposition for large-scale global optimization based on quantified variable correlations uncovered by metamodelling. *Eng Optim* 47(4):429–452. <https://doi.org/10.1080/0305215X.2014.895338>
- Hastie T, Tibshirani R, Friedman J (2017) The elements of statistical learning: data mining, inference, and prediction. Springer-Verlag, New York Inc
- Heo S, Daoutidis P (2015) Graph reduction for material integrated process networks with flow segregation. *IFAC-PapersOnLine* 48(8):586–591. <https://doi.org/10.1016/j.ifacol.2015.09.031>
- Ishizaki T, Kashima K, Imura J, Aihara K (2014) Model reduction and clusterization of large-scale bidirectional networks. *IEEE Trans Autom Control* 59(1):48–63. <https://doi.org/10.1109/TAC.2013.2275891>
- ISOH (2022) Waste management information system. CENIA. In Czech. <https://www.cenia.cz/odpadove-a-obehove-hospodarstvi/isoh/>
- Jansen JD, Durllofsky LJ (2017) Use of reduced-order models in well control optimization. *Optim Eng* 18(1):105–132. <https://doi.org/10.1007/s11081-016-9313-6>

- Kalmar M, Nilsson J (2016) The art of forecasting – an analysis of predictive precision of machine learning models (Dissertation). <http://urn.kb.se/resolve?urn=urn:nbn:se:uu:diva-280675>
- Karimi-Mamaghan M, Mohammadi M, Meyer P, Karimi-Mamaghan AM, Talbi EG (2022) Machine learning at the service of meta-heuristics for solving combinatorial optimization problems: a state-of-the-art. *Eur J Oper Res* 296(2):393–422. <https://doi.org/10.1016/j.ejor.2021.04.032>
- Kim EJ, Roux JJ, Kuznik F (2014) Decomposition and coupling of soil domain for modeling vertical ground heat exchangers using the state model size reduction technique. *Appl Therm Eng* 69(1–2):155–164. <https://doi.org/10.1016/j.applthermaleng.2014.04.043>
- Lam HL, Klemeš JJ, Kravanja Z (2011) Model-size reduction techniques for large-scale biomass production and supply networks. *Energy* 36(8):4599–4608. <https://doi.org/10.1016/j.energy.2011.03.036>
- Lam HL, Varbanov P, Klemeš JJ (2009) Optimisation of regional energy supply chains utilising renewables: P-graph approach. *Comput Aided Chem Eng* 26:1003–1008. [https://doi.org/10.1016/S1570-7946\(09\)70167-1](https://doi.org/10.1016/S1570-7946(09)70167-1)
- Laws for people (2020) Law 541/2020 Sb related to waste. In Czech. <https://www.zakonyprolidi.cz/cs/2020-541>
- Li G, Balakrishnan A (2016) Models and algorithms for network reduction. *Eur J Oper Res* 248(3):930–942. <https://doi.org/10.1016/j.ejor.2015.08.008>
- Luan X, de Schutter B, Meng L, Corman F (2020) Decomposition and distributed optimization of real-time traffic management for large-scale railway networks. *Transp Res Part b: Methodol* 141:72–97. <https://doi.org/10.1016/j.trb.2020.09.004>
- Martin N, Frasca P, Canudas-de-Wit C (2019) Large-scale network reduction towards scale-free structure. *IEEE Trans Netw Sci Eng* 6(4):711–723. <https://doi.org/10.1109/TNSE.2018.2871348>
- Meselhi M, Sarker R, Essam D, Elsayed S (2022) A decomposition approach for large-scale non-separable optimization problems. *Appl Soft Comput* 115:108168. <https://doi.org/10.1016/j.asoc.2021.108168>
- Morabit M, Desaulniers G, Lodi A (2022) Machine-learning-based arc selection for constrained shortest path problems in column generation. *INFORMS J Optim* 5:191–210. <https://doi.org/10.1287/ijoo.2022.0082>
- Musumeci F, Rottondi C, Nag A, Macaluso I, Zibar D, Ruffini M, Tornatore M (2019) An overview on application of machine learning techniques in optical networks. *IEEE Commun Surv Tutor* 21(2):1383–1408. <https://doi.org/10.1109/COMST.2018.2880039>
- Orkun Baycik N (2021) Machine learning based approaches to solve the maximum flow network interdiction problem. *Comput Ind Eng* 167:107873. <https://doi.org/10.1016/j.cie.2021.107873>
- Pluskal J, Šomplák R, Hrabec D, Nevrlý V, Hvattum LM (2022) Optimal location and operation of waste-to-energy plants when future waste composition is uncertain. *Oper Res Int J* 22(5):5765–5790. <https://doi.org/10.1007/s12351-022-00718-w>
- R Core Team (2021) R: a language and environment for statistical computing
- Sharma G (2021) A literature review on application of Artificial Intelligence in Human Resource Management and its practices in current organizational scenario. In: 2021 Fifth International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud) (I-SMAC), pp 594–600. IEEE. <https://doi.org/10.1109/I-SMAC52330.2021.9640954>
- Smejkalová V, Šomplák R, Pluskal J, Rybová K (2022) Hierarchical optimisation model for waste management forecasting in EU. *Optim Eng* 23(4):2143–2175. <https://doi.org/10.1007/s11081-022-09735-2>
- Williams HP (2009) Logic and integer programming. Springer, New York, N.Y.
- Xie X, Zheng H, Jonckheere S, van de Walle A, Pluymers B, Desmet W (2018) Adaptive model reduction technique for large-scale dynamical systems with frequency-dependent damping. *Comput Methods Appl Mech Eng* 332:363–381. <https://doi.org/10.1016/j.cma.2017.12.023>
- Zeng Y, Song C, Ge T, Zhang Y (2022) Reduction of large-scale graphs: Effective edge shedding at a controllable ratio under resource constraints. *Knowl-Based Syst* 240:108126. <https://doi.org/10.1016/j.knosys.2022.108126>
- Zhang Y, Xin J, Li X, Huang S (2020) Overview on routing and resource allocation based machine learning in optical networks. *Opt Fiber Technol* 60:102355. <https://doi.org/10.1016/j.yofte.2020.102355>
- Zheng Y, Xie Y, Lee I, Dehghanian A, Serban N (2021) Parallel subgradient algorithm with block dual decomposition for large-scale optimization. *Eur J Oper Res* 299(1):60–74. <https://doi.org/10.1016/j.ejor.2021.11.054>
- Zhou L, Zhen L, Baldacci R, Boschetti M, Dai Y, Lim A (2021) A heuristic algorithm for solving a large-scale real-world territory design problem. *Omega* 103:102442. <https://doi.org/10.1016/j.omega.2021.102442>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.