Journal of Cleaner Production 420 (2023) 138393

Contents lists available at ScienceDirect

Journal of Cleaner Production

journal homepage: www.elsevier.com/locate/jclepro

Mathematical modelling of waste flows and treatment based on reconstruction of historical data: Case of wastewater sludge in Czech Republic

Jaroslav Pluskal [a,*], Radovan Šomplák [a], Lucie Němcová [a], Jiří Valta [b], Martin Pavlas [a]

[a] *Institute of Process Engineering, Faculty of Mechanical Engineering, Brno University of Technology – VUT Brno, Technická 2896/2, 616 69, Brno, Czech Republic*
[b] *Czech Environmental Information Agency, Moskevská 1523/63, 101 00, Praha 10, Czech Republic*

ARTICLE INFO

Handling Editor: Cecilia Maria Villas Bôas de Almeida

Keywords:
National waste management database
Data verification
Data reconciliation
Network flow problem
Sewage sludge
Dry matter treatment

ABSTRACT

The contribution of supporting tools developed for logistic optimization and processing infrastructure planning is highly dependent on the quality of input data. Databases containing reports on waste production and treatment are often used for waste management case studies and in the aggregated form at the national level provide sufficient information with negligible errors. On the other hand, analysis in greater detail at the smallest administrative unit usually reveals significant inconsistencies, especially in the case of mass balance control. This often results in the impossibility to track waste flows in the system and evaluate final treatment. This paper aims to present a tool for database verification in order to satisfy mass balances at all levels of the investigated system. As part of the mathematical modelling, the database is first subjected to a statistical analysis where selected inconsistencies between records can be removed. Subsequently, data reconciliation is applied with appropriate weights assigned to each entity in the system. The weights represent credibility and reflect various errors, which are identified in the previous stage of the approach. The correction of mass balances enables an evaluation of quality or composition in greater detail of individual entities and also determines the feasible interval of monitored properties. The tool is applied to the treatment evaluation of dry matter in wastewater sludge in the Czech Republic. Thanks to the developed tool, disposal and material or energy recovery can be quantified at micro-regional.

1. Introduction

Waste management (WM) is a crucial aspect of ensuring the sustainability and environmental well-being of our society. The pressure to increase recycling and reduce waste is not only coming from the European Union (STOA, 2017) but also from the general public, who are becoming more aware of the negative effects of waste on the environment and public health (Abubakar et al., 2022). Therefore, new legislation has been implemented to address this issue, such as an increase in fees for landfill disposal, and a ban on landfilling waste with a calorific value greater than a specified amount in order to utilize it for energy recovery (Kumar and Samadder, 2022). These measures are aimed at reducing the environmental impact of waste and promoting recycling and other forms of WM in the established hierarchy.

WM is a regulated and closely monitored business in many states, where entities are forced to report data on the production and treatment of its waste and waste handed over. Data is reported regularly and further processed by the authorities and used for decisions (Sileryte et al., 2022). With the advancement of technology and digitalization, there will be an increasing amount of data available on waste, including information on the types and quantities of waste generated, as well as the methods used for its treatment (Sepasgozar et al., 2021). However, if this data is not utilized properly, all investments in data collection will be in vain. Therefore, it is essential to systematically analyze data using statistical tools to identify the potential for increased recycling and to plan the processing chain. It can help to improve the efficiency and effectiveness of WM and reduce its environmental impact (Tsai et al., 2020).

The WM supply chain is a complex system process that involves multiple entities and stages (Eghbali et al., 2022). The study (Sileryte et al., 2022) closely discusses the importance of studying flows and relations, determining the mass quantity and material content with a high

* Corresponding author.
E-mail address: Jaroslav.Pluskal@vutbr.cz (J. Pluskal).

https://doi.org/10.1016/j.jclepro.2023.138393
Received 28 February 2023; Received in revised form 11 July 2023; Accepted 7 August 2023
Available online 8 August 2023

level of geographical information. One of the biggest challenges is to identify the flow of waste from the producers to the place of its final treatment, which could, in general, include many handovers of waste between entities. In addition, there are discrepancies in the balance of waste production and treatment not only at individual entities but also in regions or states. Therefore, the whole chain of waste handling is not usually known and it is difficult to have an overview of the whole system and to identify where the WM system is working well and the areas where improvements are needed.

Without any development in waste flow tracking, the above-mentioned issues can lead to incorrect analyses and conclusions, because aggregated or averaged values may be used to determine the indicators of WM. The necessity of material flow analysis in greater detail is highlighted by Islam and Huda (2019) in order to open area for investigation from a holistic regional material-cycle perspective. Specifically, in the Czech Republic, the production and treatment of wastewater sludge are only available in the form of a slurry (i.e. with high content of water) and the dry matter content is only known at the point of generation. Wastewater sludge can be characterized as byproduct of sewage treatment with organic matter and nutrients like nitrogen and phosphorus, and also some potentially harmful substances (Bennamoun et al., 2013). With proper management, sludge can be reused in agriculture, energy generation, and construction, but the environmental aspects resulting from material recovery are crucial (Kacprzak et al., 2017). However, the dry matter content with useful substances in slurry varies in general and depends on the technology used (Halecki et al., 2016). Due to errors in the database and information loss caused by flows through multiple entities, it is impossible to identify all chains in the system and therefore the nationwide percentage average of dry matter content is currently used for dry matter treatment evaluation instead of regional estimates. As a result, the indicators of WM may deviate significantly from reality and a seemingly functioning system may be inefficient.

Overall, it highlights the importance of proper data governance in WM and the need for an accurate and comprehensive tool to maintain the reliability of data. This can help to minimize the environmental impact of waste, optimize resource recovery, and improve the overall efficiency of WM planning. The use of sophisticated models for the assessment of waste streams and their treatment can be considered more advantageous than increasing the duties of entities in the reporting system. The goal of this paper is to effectively manage collected data, fix errors, reconstruct missing data, provide material quantity including its quality, and overall gain better insight into the monitored system. The key characteristic is the analysis of a large amount of data within the entire state, as individual entities are often interrelated and cannot be decomposed into several parts. In addition, in the case of WM, there is usually only a single set of data without the possibility to make additional measurements or verification. Unlike the management of errors in production processes, it defines unique conditions for data handling and its further evaluation.

2. State-of-the art and literature review

Data errors can occur in a variety of ways, such as human error, system failures, or inaccuracies in data collection (Câmara et al., 2017). These errors can have a significant impact on the accuracy and reliability of the data, typically considered independent (Brown et al., 2018). Therefore, data is analyzed using statistical or optimization techniques in a variety of industries. Statistical methods are commonly used for data standardization and quality control. These methods can be used to cleanse and validate data, as well as to identify and correct errors or outliers in the data. For example, Material or Energy Flow Analysis (MFA, EFA) through the Sankey diagram is often used to analyze processes, explore interactions, identify the largest losses and accurately target interventions to improve efficiency (Bowman et al., 2022). Other techniques such as Principal Component Analysis (Kim et al., 2022) or Cluster Analysis (Chao et al., 2019) can be used to identify patterns and relationships in the data and to identify areas where data is missing or inconsistent. These methods can also be used to estimate missing data and to correct errors in the data.

Optimization models are more likely used to optimize the process itself and find optimal settings in order to achieve higher efficiency (Chen et al., 2018). However, techniques based on mathematical modelling can be also used to identify and correct errors in the data. In this respect, data reconciliation (DR), which compares and adjusts different sets of data to ensure consistency and accuracy, can be considered the most utilized technique as it is commercially available (Câmara et al., 2017). It is often used in industrial and financial applications to ensure that data from different sources, such as sensors, databases, and manual inputs, agree with one another (Cochinwala et al., 2001). The goal of DR is to provide a single, reliable set of data that can be used for decision-making, analysis, and reporting. It can be performed using neural networks or genetic algorithms, which mimic the process of natural selection to find the best solution to a problem (Hu et al., 2022). However, they represent a black box technique, which can make it difficult to control the whole procedure and understand the underlying processes that led to the solution.

DR can be considered a well-known principle for several decades (Kuehn and Davidson, 1961), which in certain variations is widely used in industry to align the physical laws of a process with measured data. The main area of application is the production process or energy sector, which is carefully monitored today, as it is connected to almost every segment of society. The study (Sharma et al., 2022) utilizes DR in the dynamic control of renewable energy sources, which allows for rectifying inconsistencies in sensor measurements. Yu et al. (2022) describe a challenge in maintaining the accuracy of calculations for the isentropic efficiency of a steam turbine stage. There are widely existing gross errors in steam turbine measurements, which can invalidate modelling results and hinder model-based monitoring and optimization. They propose a DR model that includes entropy increase constraints and uses nonlinear inequality constraints to detect gross errors.

Another article (Badings and van Putten, 2020) focuses on the integration of DR into allocation tasks, specifically in the petrochemical industry. The authors address the identification and correction of errors in data related to pipe flow, while also addressing the mixing of multiple products with different properties. The methodology is described generally with mathematical notation, but the mixing of multiple flows is solved by linearization using a simulation tool specialized for this branch. As generally known, the outcome of DR is highly dependent on the covariance matrix (Gurevich et al., 2022), which there contains only the diagonal values reflecting the standard deviation of individual measurements. This methodology cannot be used in the case of WM, as the reported data can vary from year to year. Therefore, only one data point is available, and in some cases, it is missing and needs to be reconstructed.

In the case of WM, studies analyzing errors in the system are again focused on specific processes in waste handling. For example, Behnami et al. (2019) present an approach for WM, specifically for improving the accuracy of measurements in the process of a wastewater treatment plant. However, such approaches are limited to a specific process, while it is useful to generalize and expand this local case to fill the gap between typical MFA at a national or regional level and point-based analyses. The study Šomplák et al. (2019) addresses such extension by applying DR to fix errors between waste production and treatment. The result of the presented case study related to bulky waste shows that the presented model provides a solution at the regional level but is not capable to provide insights into the required detail for each entity in the system.

Big data at a national level are used in forecasting future development using the DR (di Fonzo and Girolimetto, 2022). In the field of WM, Pavlas et al. (2017) present an approach for waste generation forecasting with the aim of having the most detailed outlook of investigated territory. However, historical data contains many errors, outlier values, etc., so it is appropriate to use the territorial hierarchy and make forecasts even on aggregated areas that are more consistent. The forecast of different levels usually does not meet mass balances, and therefore individual time series are subsequently corrected using DR. A similar principle is also used in forecasting of waste treatment for individual countries in the EU (Smejkalová et al., 2022).

Wider studies focusing on monitoring WM or tracking waste streams across the whole country usually do not address the issue of data inconsistency and only implement procedures related to basic analysis or data mining to find specific patterns (Li et al., 2019). The paper (Zhan et al., 2022) highlights the necessity of data handling and provides an approach based on mining and fusion of the reported values and available web text data. The goal is to identify the challenges and provide policy implications for the optimization by supplementation of reported data by another data set. Nevertheless, the consistency of data is ensured by sorting and cleaning reported data without further explanation. It may therefore be beneficial to have a general approach to any waste database and to ensure that any subsequent studies are based on the most reliable input data.

It is obvious from the literature review that the development of a comprehensive tool to ensure the consistency of large-scale databases in WM is not sufficiently addressed. Moreover, generally used MFA does not provide information about quality of individual stream or does not track its specific content. An example of improved MFA is presented by Li et al. (2022), where authors developed a probabilistic approach based on monte-carlo simulations. The study aims to evaluation of phosphorus content in steelmaking process. However, Monte-Carlo simulations are very limited by the size of the problem, which poses difficulties in the case of large data. Therefore it is important to develop a new approach to maintaining data consistency and increasing data reliability in the field of WM with subsequent advanced MFA. It can be achived by combinations of several methods for data analysis, DR and network flow problems, which can be generally utilized for any waste to assess the environmental impact and further optimize the material recovery. Unlike other research studies, this article deals with a broad database at a national level rather than a single process.

The presented approach works with a large amount of data in a unique branch of waste management reporting. It is characterized by the need to work with the highest possible level of detail while having only one measurement of each record. This prevents the use of common weights for reconciling inconsistencies based on variance and requires the development of specific indicators. To overcome this problem, a tailor-made covariance matrix for DR is developed. After fixing inconsistencies, it is possible to track the composition and quality of individual waste streams within the reported data, allowing for a more accurate evaluation of waste handling and providing insights into the system at a regional level. The presented paper related to data management is unique in both its scope and combination of various mathematical apparatus with a specific focus on wastewater sludge and the calculation of its dry matter content. The novelty and key points of the study are as follows:

- Step-by-step approach for maintaining waste management big data in micro-regional detail.
- Transportation error fixing by searching specific inconsistency patterns.
- Data reconciliation with the tailor-made covariance matrix.
- Material flow analysis with quality or content tracking including point and interval estimates.
- Wastewater sludge case study in Czech Republic.

3. Methodology

The goal of the presented paper is to develop a general approach for any waste and similar monitored systems, by correcting inconsistencies and evaluating waste content in the greatest possible detail, providing adequate insight. In the case of waste tracking and monitoring its quality or specific content, the problem is often more complex, as waste can be mixed at some point. In such cases, it is necessary to implement a weighted average, which represents a nonlinear problem. As is generally known, nonlinear tasks are difficult to solve in the scope of whole countries in the maximum possible detail and thus it is necessary to develop a new approach with combinations of various techniques. For this reason, the solution was divided into several steps, where the first goal is to deal with errors in the database and ensure mass balance at all levels of the system. A special form of DR with defined weights reflecting expert estimates and verification of incorrect reports is used here. The individual flows between entities are afterwards known and therefore the evaluation of waste stream content can be performed in the form of point or interval estimates. The general framework of the methodology is shown in Fig. 1.

As can be observed, direct identification of possible improvement from the originally reported data is difficult, because many calculated metrics are biased by errors and usually not evaluated with the necessary detail of the investigated system. The use of precious methodology to evaluate waste content in individual entities needs to deal with inconsistency in waste transfer, where values averaging can lead to the loss of important links. The best solution is to perform a step-by-step approach in sequence ensure consistency using algorithmic corrections and DR, evaluate waste content and derive implications. The waste content tracking needs to deal with inconsistencies in data and satisfy mass balances, which futher results into optimization problem with zero degree of freedom in the case of point estimates. Interval estimates are independed on point estimates, therefore its evaluation can be performed parallely to speed up the calculation.

3.1. Pre-processing

Before the actual DR, a thorough analysis of the data is necessary to identify system errors and prepare the structure for further reconciliation. The reported data is based on WM codes, which must be reported annually along with the waste quantity. The principles of reporting, including the relevant codes and rules, are derived from European legislation, specifically for the Czech Republic, details can be found in the relevant decree (European Commission, 2022). The relevant sheet for entities called "Sheet 2 - Reporting of summary data from ongoing records of WEEE processors for the reporting year" can be also found there as well as many rules and reporting sequences. The basic principles can be summarized as follows:

- The waste handling codes are divided into plus (production and reception) and minus (treatment and transfer) categories.
- The first letter of the code represents the origin of waste (A = own waste, B = others, C = accumulated amount from previous years.
- Each record has an entity as an announcer and an entity as a partner. The partner can be also the same as the announcer (e.g. for production and treatment).
- The mass balances have to be equal to 0 for every entity of the system (input versus output).
- The amount of transferred waste has to be equal to its reception along a flow between two entities.

Mass balances are often disrupted, with waste either in excess or lacking in some nodes. Furthermore, discrepancies can arise during waste transportation between two entities, with conflicting reports of how much waste was transferred and received. Ideally, records should be consistent across pairs, with the other entity listed as a partner and the same quantity for each transaction. For some types of waste, small deviations in values may be acceptable, due to factors such as water evaporation, but in some cases, records differ by tens of percent, or one side does not even record the interaction. The designed DR (see Section 3.2) should handle these errors based on links in the system, but some

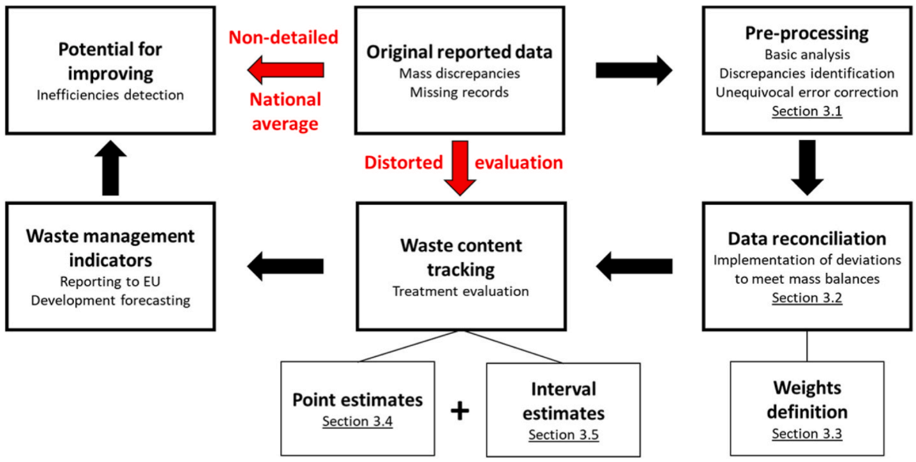


Fig. 1. General framework of the proposed step-by-step.

transport discrepancies can be pre-solved and the number of degrees of freedom can be reduced. This mainly refers to the case where two records correspond in value, but one record contains a bad partner that does not indicate such interaction. This can often happen due to the report at the company's headquarters, while the waste is received at a subsidiary. Using the algorithm shown in Fig. 2, these errors can be detected and corrected.

The principle of correction is based on a systematic analysis of all recorded flows, which are compared with each other. In order for a correction to be made, both flows must have the same amount of discrepancy in opposite meanings (error in reception vs. error in transmission) and must have the same subject on one side of the record. If there is more than one possible correction, the error should be left for DR. The principle of correction can be also used for a combination of records when the effort relies on to findings of several records B, in which the sum of errors is equal to the error of record A. The pre-processing corrections do not modify production or treatment and it is achieved with proper redirections of waste transfers and receptions.

3.2. Data reconciliation

Pre-processing data thoroughly can fix many inconsistent database entries, but errors can persist. These uncorrected mass balances must be tackled using a sophisticated method that considers all links between entities links and finds the most likely solution. A mathematical approach using DR principles has been developed to implement the minimum possible deviations into reported values and meet all conditions defined for the system.

The mathematical model for DR is based on nodes and oriented arcs. Due to a large number of possible arcs between all nodes, only those arcs between two entities that show one-sided or mutual interaction between them are generated which allows for avoiding computational inefficiencies and reducing memory demand. This mathematical approach incorporates predefined weights (see Section 3.3), reflecting logical rules and expert input, to resolve the waste discrepancies because the variance of measurement for the typical covariance matrix is not possible due to a single data set. It can also be meaningful to consider off-diagonal values, which could represent error chaining caused by

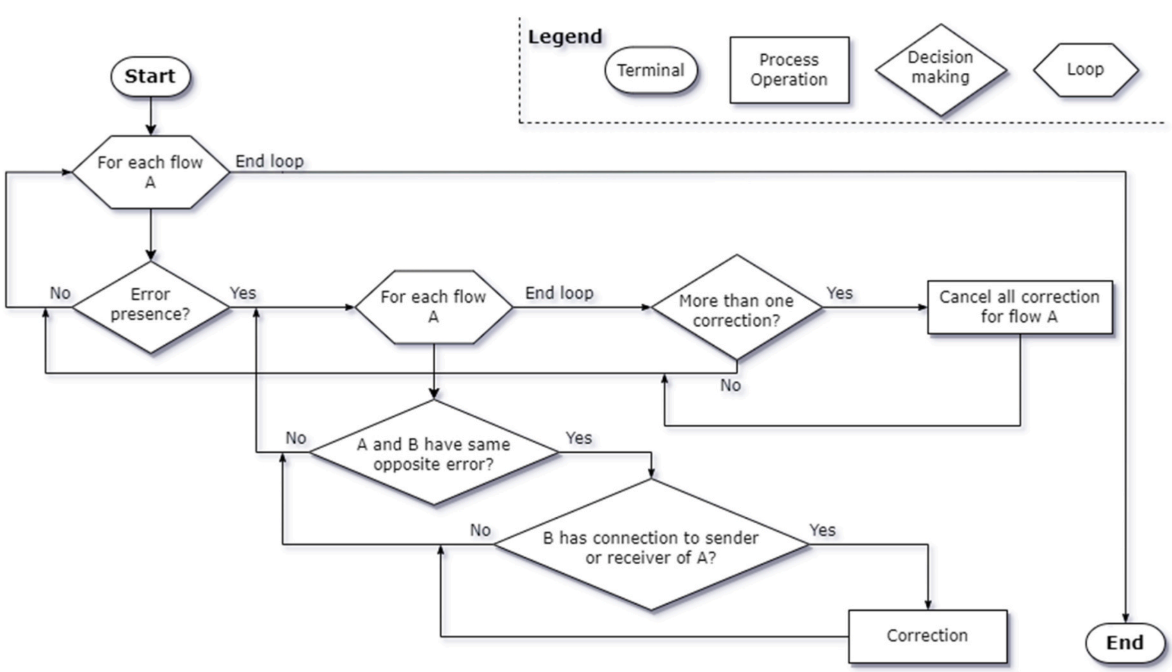


Fig. 2. Flowchart describing the transport error correction.

mutual influencing, for example, waste collection companies, which usually provide final values for reporting to individual entities.

In the following section, the mathematical model for the balance calculation will be presented, which is conceived as a linear model. Such a model has convenient properties for optimization in terms of solvability and finding the global optimum. Moreover, the linear dependence of the implemented deviation to the objective function leads to the selection of the median. It follows the idea that more reliable and consistent entities in the system are more trusted. This property is desirable since it is assumed that in most cases the record will be correct. Thus, the result of DR lies primarily in a selection of correct values rather than averaging them as in the case of quadratic dependence. First, a list of the symbols used is given,

Sets

$a \in A$ set of arcs representing flows between entities
$i \in I$ set of entities in the database (for example municipalities)

Parameters

F_a^+ reported amount of waste received along a, [kg]
F_a^- reported amount of waste transferred along a, [kg]
$M_{a,i}$ incidence matrix determining input and output of all arcs, [-]
P_i reported amount of produced waste in entity i, [kg]
T_i reported amount of treated waste in entity i, [kg]
w_i^P weight for produced waste in entity i, [-]
w_i^T weight for treated waste in entity i, [-]

Variables

$\gamma_a^{F_+}$ implemented deviation in received waste along a, [kg]
$\gamma_a^{F_+^+}, \gamma_a^{F_+^-}$ deviation division into positive and negative parts (received waste), [kg]
$\gamma_a^{F_-}$ implemented deviation in transferred waste along a, [kg]
$\gamma_a^{F_-^+}, \gamma_a^{F_-^-}$ deviation division into positive and negative parts (transferred waste), [kg]
γ_i^P implemented deviation in produced waste in entity i, [kg]
$\gamma_i^{P^+}, \gamma_i^{P^-}$ deviation division into positive and negative parts (waste production), [kg]
γ_i^T implemented deviation in treated waste in entity i, [kg]
$\gamma_i^{T^+}, \gamma_i^{T^-}$ deviation division into positive and negative parts (waste treatment), [kg]

The objective function of the optimization model is equal to the sum of all deviations implemented to the data in order to satisfy all mass balances and links in the system. These deviations must be split into positive and negative parts for correct implementation (see below for a description of equations (5) and (6)). Each element of the system has an associated confidence level, which is implemented in the model using weights representing values in the diagonal of the covariance matrix. The goal is to implement minimum deviations into the database, this sum of deviations is thus minimized. The corresponding mathematical notation is given by Eq. (1), where deviations for produced ($\gamma_i^{P^+}, \gamma_i^{P^-}$) and treated ($\gamma_i^{T^+}, \gamma_i^{T^-}$) waste are presented with matching weight (w).

$$\min\left(\sum_{i\in \mathrm{I}} w_i^P\left(\gamma_i^{P^+}+\gamma_i^{P^-}\right)+\sum_{i\in \mathrm{I}} w_i^T\left(\gamma_i^{T^+}+\gamma_i^{T^-}\right)\right). \tag{1}$$

The following formulas express the necessary constraints for the optimization model. Eq. (2) describes the equality between received and transferred waste along all modelled arcs. The reported amount of received waste F_a^+ including estimated deviation $\gamma_a^{F_+}$ must be equal to the reported amount of transferred waste F_a^- with estimated deviation $\gamma_a^{F_-}$.

$$F_a^+ + \gamma_a^{F_+} = F_a^- + \gamma_a^{F_-}, \forall a \in \mathrm{A}. \tag{2}$$

Next Eq. (3) provides the mass balance within a single entity. The sum of the production records P_i with the corresponding deviation γ_i^P minus the sum of the treated waste T_i with deviation γ_i^T plus the waste transport $M_{a,i}(F_a^+ + \gamma_a^{F_+})$ must be equal to zero. Here, the incidence matrix $M_{a,i}$ provides the necessary information about which arcs enter a given node and which ones leave. The sum of all arcs multiplied by the incidence matrix produces the overall transport mass balance of an entity.

$$P_i + \gamma_i^P - T_i - \gamma_i^T + \sum_{a\in \mathrm{A}} M_{a,i}\left(F_a^+ + \gamma_a^{F_+}\right) = 0, \forall i \in \mathrm{I}. \tag{3}$$

Eq. (4) provides the mass balance within the whole balanced system. All waste produced must be treated. Thus, the sum of the corrected productions must equal the sum of the corrected treatments.

$$\sum_{i\in \mathrm{I}}\left(P_i + \gamma_i^P\right) = \sum_{i\in \mathrm{I}}\left(T_i + \gamma_i^T\right). \tag{4}$$

Since the deviation can take any real value, negative deviations would reduce the value of the objective function under minimization. Therefore, for correct implementation of deviations, it is necessary to consider their absolute value, which, however, from the optimization point of view, distorts the favourable conditions (linear model) for finding the optimal result. This problem can be solved by dividing the deviations into positive and negative parts as described in Eq. (5) and Eq. (6). At the same time, it is necessary to ensure that these variables are non-negative. This condition is described by Eq. (7) and Eq. (8).

$$\gamma_a^{F_+} = \gamma_a^{F_+^+} - \gamma_a^{F_+^-}, \gamma_a^{F_-} = \gamma_a^{F_-^+} - \gamma_a^{F_-^-}, \forall a \in \mathrm{A}, \tag{5}$$

$$\gamma_i^P = \gamma_i^{P^+} - \gamma_i^{P^-}, \gamma_i^T = \gamma_i^{T^+} - \gamma_i^{T^-}, \forall i \in \mathrm{I}, \tag{6}$$

$$\gamma_a^{F_+^+}, \gamma_a^{F_+^-}, \gamma_a^{F_-^+}, \gamma_a^{F_-^-} \geq 0, \forall a \in \mathrm{A}, \tag{7}$$

$$\gamma_i^{P^+}, \gamma_i^{P^-}, \gamma_i^{T^+}, \gamma_i^{T^-} \geq 0, \forall i \in \mathrm{I}. \tag{8}$$

In order to maintain the meaningfulness of the problem to be solved, there must not be a situation where the deviation causes a negative flow, production or treatment of waste. For this reason, Eqs. (8)–(10) are defined.

$$F_a^+ + \gamma_a^{F_+} \geq 0, \forall a \in \mathrm{A}, \tag{9}$$

$$P_i + \gamma_i^P \geq 0, \forall i \in \mathrm{I}, \tag{10}$$

$$T_i + \gamma_i^T \geq 0, \forall i \in \mathrm{I}. \tag{11}$$

As an additional condition of the DR, it is appropriate to implement a constraint to ensure that waste transfers and receipts are not balanced across the boundaries of the recorded values. In fact, this constraint will ensure that in particular those cases where the records on both sides of the reporting flow correspond in value are fixed. These flows can be considered correct and it is not desirable to modify them. The constraint is implemented using Eq. (12), where the magnitude of an error on a given flow in absolute value must be equal to the sum of the implemented deviations to flow.

$$\left|F_a^+ - F_a^-\right| = \gamma_a^{F_+^+} + \gamma_a^{F_+^-} + \gamma_a^{F_-^+} + \gamma_a^{F_-^-}, \forall a \in \mathrm{A}. \tag{12}$$

The following Eqs. (13)–(16) are optional, but it is recommended to use at least one defined combination based on expert knowledge. However, at most one configuration can be chosen, otherwise, the problem is likely to be infeasible. The following may be chosen:

- Eq. (13) and Eq. (14): the optimization model is forced to compare total production and treatment in order to meet the mass balance of the system. However, the logical solution seems to be a procedure where the lower value is increased by the given difference. It can be assumed that information is missing in the system rather than extra records being given. Thus, simultaneous consideration of Eq. (13) and Eq. (14) will ensure that the lower value (production or treatment) will be at least equal to the second value reported. However, individual entities are not affected by this condition.

- Eq. (15): Using this equation, it is possible to set the maximum credibility on the reported production. Thus, it is strictly forbidden to modify the production in any entities and the model only adjusts the amount of treatment.
- Eq. (16). Using this equation, it is possible to set the maximum credibility on the reported treatment. Thus, it is strictly forbidden to modify the processing in any entities and the model only adjusts the amount of production.

$$\sum_{i\in \mathrm{I}}\left(P_i+\gamma_i^P\right) \geq \sum_{i\in \mathrm{I}} P_i, \tag{13}$$

$$\sum_{i\in \mathrm{I}}\left(T_i+\gamma_i^T\right) \geq \sum_{i\in \mathrm{I}} T_i, \tag{14}$$

$$\gamma_i^P = 0, \forall i \in \mathrm{I}, \tag{15}$$

$$\gamma_i^T = 0, \forall i \in \mathrm{I}. \tag{16}$$

3.3. Weights definition

The weights in the mathematical model are a crucial input that plays a key role in determining the final solution. These weights must be defined based on expert estimates, as in the case of WM, only one set of data is available and it is not possible to assume any form of records verification or additional measurements. The weights are defined based on the identified discrepancies, which further determine the rate of credibility for each entity in the system. The goal is to adjust more likely the faulty entity instead of trying to distribute the error among all in the system, which should be carried out as the last possible solution. This effort is also supported by the linear form of the model. The principle of weights definition can be summarized as the statement that an entity showing more error rate is modified more than the other.

The most of formulas are given for the weights related to production, and if the exception is not mentioned, they are applicable also to the weights related to treatment. The greater value of weight means higher credibility of an entity and less tendency to modification. The goal is to keep weights between zero and one, while the value zero is not desirable because it leads to the possibility of unlimited changes in an entity. The greater value of weight means higher credibility of an entity. The current knowledge obtained during the development and testing of the tool has led to the following rules, which define individual parts used for the final calculation of weights.

The first part of the modelled weight is related to the error rate in transport reporting. The aim is to implement greater deviations to entities showing higher relative error, as these entities are more likely to not consistently follow the principles of reporting. The relevant coefficient is expressed by Eq. (17), where the sum of errors on the arcs is divided by the average of reported values, including only those arcs connected to the entity being evaluated. The resulting value is subtracted from one, with a desirable minimum value. In the case of no reception or transfer of waste, the value is set to one.

$$k_i^{1,P} = \begin{cases} 1, \text{for no connected transport} \\ \max\left(0.1; 1 - \dfrac{\sum_{a\in \mathrm{A}}\left|M_{a,i}\left(F_a^+ - F_a^-\right)\right|}{\sum_{a\in \mathrm{A}}\left|M_{a,i}\dfrac{F_a^+ + F_a^-}{2}\right|}\right), \text{otherwise} \end{cases}, \forall i \in \mathrm{I}. \tag{17}$$

The following Eq. (18) adjusts the weight based on the error in the mass balance within an entity. The total mass imbalance is divided by the sum of production and treatment. An entity without production and treatment can be considered a transit node and it is not desirable for transport imbalances to cause a change in these records. Therefore, in such a case, the weight is strictly set to a value of 1. The weight is also influenced by Eq. (19) and Eq. (20) depending on whether it is related to production or treatment. In the case of an entity with more waste, the goal is to reduce the treatment weight, so that the model tends to increase the amount of treated waste rather than reduce production. The same principle is used in the case of a negative balance, where the aim is to increase the production of waste. However, it is necessary to consider all other links in the system first, and it should only be reflected in the result as the last step. Therefore, a coefficient of 0.9 is chosen, which does not significantly affect the resulting weight but ensures the desired output.

$$k_i^{2,P} = \begin{cases} \max\left(0.1; 1 - \dfrac{\left|P_i + \gamma_i^P - T_i - \gamma_i^T + \sum_{a\in \mathrm{A}} M_{a,i}\left(F_a^+ + \gamma_a^{F^+}\right)\right|}{P_i + T_i}\right), \text{for } P_i + T_i > 0 \\ 1, \text{otherwise} \end{cases}, \forall i \in \mathrm{I}. \tag{18}$$

$$k_i^{3,P} = \begin{cases} 0.9, \text{for mass balance} < 0 \\ 1, \text{otherwise} \end{cases}, \forall i \in \mathrm{I}. \tag{19}$$

$$k_i^{3,T} = \begin{cases} 0.9, \text{for mass balance} > 0 \\ 1, \text{otherwise} \end{cases}, \forall i \in \mathrm{I}. \tag{20}$$

It can be expected that some records regarding reality may be missing in the database, which can escalate to a missing entire entity. The tool is capable of detecting these subjects if others report some interaction with them. Otherwise, missing reports cannot be identified from the investigated database, for example, when an entity produces and treats only its waste. However, if a subject is listed in a record without any own records, it is assumed that the necessary reports related to waste handling are not provided, and it has the lowest credibility. This principle is ensured by the following Eq. (21), which leads that the model assigns a corresponding production or treatment regarding connected transport records for a given entity.

$$k_i^{4,P} = \begin{cases} 0.1, \text{without any own records} \\ 1, \text{otherwise} \end{cases}, \forall i \in \mathrm{I}. \tag{21}$$

The purpose of the original DR is to distribute deviations in the system regarding qualitative characteristics, in order to satisfy the mass balance. In the case of quadratic form DR, data normalization is often utilized to make the relative deviation share roughly equal for all entities in the system. For linear form DR, an additive form of such principle must be used instead of a multiplicative form, since otherwise, the qualitative weight could only have a negligible effect. This is due to the linear dependency, where if an entity has a smaller weight, it is modified instead of deviations being proportionally divided among all entities. Therefore, the goal is to propose a rule that, when a deviation needs to be implemented between two entities with similarly equal qualitative weight, the larger entity should be modified, because it represents a smaller relative change. Eq. (22) describes the calculation of the normalization part of the weight, where the square root ensures a slower decrease in order to keep differences in the case of lower values of production and treatment. It should be noted that the form of the normalization part of weight should be adjusted accordingly to the amount of produced waste and size variability of individual entities in the system. This particular proposal is designed for the case study of wastewater sludge presented in Section 4.

$$v_i^P = \begin{cases} \dfrac{1}{\sqrt{P_i + T_i}}, for\ P_i + T_i > 100 \\ 0.1, otherwise \end{cases}, \forall i \in \mathrm{I}. \tag{22}$$

All the above parts together form a weight whose value is implemented in the DR. The following Eq. (23) expresses the final calculation, where all parts of qualitative characteristic are multiplied together the normalization part is added.

$$w_i^P = k_i^{1,P} k_i^{2,P} k_i^{3,P} k_i^{4,P} + v_i^P, \forall i \in \mathrm{I}. \tag{23}$$

3.4. Waste content tracking

The evaluation of waste treatment regarding waste quality is presented in the case of wastewater sludge and monitoring of its dry matter content. The wastewater sludge is considered in the form of a slurry (high content of water), while its solid part with important nutrients is referred to as dry matter. For example, in the Czech Republic, the monitoring of the treatment of the dry matter in wastewater sludge is based on a nationwide average calculated at the point of production, where information about dry matter content is available. This approach is used due to inaccuracies in transportation and difficulties in tracking waste within the system's network. The main issue is that all waste becomes group B (following waste handling) after the first transfer and further tracking of the entire chain from the producer is obscured by subsequent takeovers at this node (representing an entity). However, such simplification can introduce significant errors in the evaluation of indicators and it is advisable to average the data at least regionally or directly of the entities themselves. The principle is illustrated in Fig. 3.

The problem of tracking flows from producers to treatment point is solved by assuming that all waste entering the node is ideally mixed and then handled. This results in dividing each entity into two parts, with the dry matter ratio between parts being equal. The dry matter ratio in the node is calculated as a weighted average of all incoming streams. A further assumption is that the secondary production of waste from pretreatment is the same as the average in the entity because there is no additional information related to the dry matter content of pretreatment. The task is based on the network flow problem, where the amount of slurry is always known thanks to DR and only the dry matter ratio in individual nodes is unknown. The following is a list of sets and a description of the mathematical model,

Sets

$a \in A$ set of arcs
$i, j \in I$ set of nodes

Parameters

D_i^P percentage content of dry matter in produced wastewater sludge in node i in a given year, [-]
D_i^V percentage content of dry matter in stored wastewater sludge in node i in a previous year, [-]
F_a amount of wastewater sludge transported along arc a, [kg]
$M_{a,i}^{IN}$ incidence matrix for input flows a in node i, [-]
$M_{a,i}^{OUT}$ incidence matrix for output flows a in node i, [-]
P_i amount of wastewater sludge produced primarily in node i in a given year, [kg]
O_i amount of wastewater sludge produced secondary in node i in a given year, [kg]

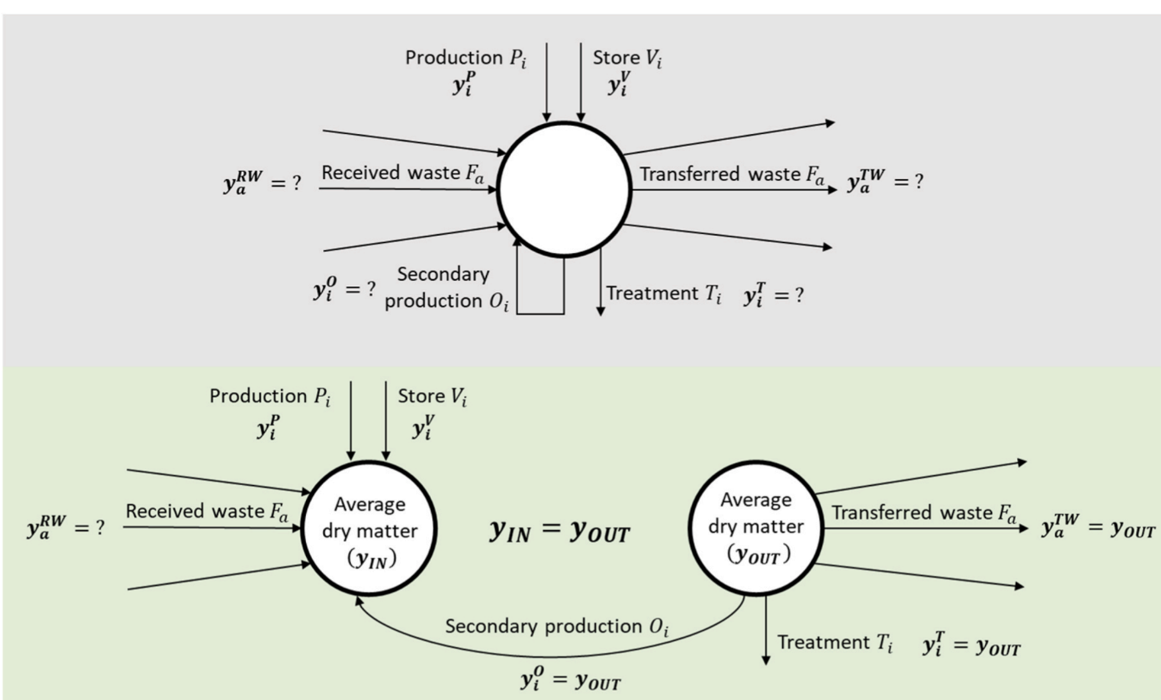


Fig. 3. Description of node division for evaluation of the average dry matter.

V_i amount of wastewater sludge stored in node i in a previous year, [kg]

Variables

y_i average percentage content of dry matter in wastewater sludge in node i, [-]

The evaluated system has zero degrees of freedom thanks to DR and thus objective function is not defined. Only one constraint is needed to calculate desired waste quality. The following Eq. (24) describe the weighted average of input flows to calculate the percentage dry matter content, which is equal to total dry matter divided by total wastewater sludge in every node. The left side of the equation gradually introduces the amount of dry matter in produced $P_iD_i^P$, accumulated from the previous year $V_iD_i^V$, secondary produced O_iy_i and received slurry, while the right side is composed of the total amount of slurry multiplied by the average percentage variable y_i.

$$P_iD_i^P + V_iD_i^V + O_iy_i + \sum_{a\in A}\left(M_{a,i}^{IN}F_a\sum_{j\in I}\left(M_{a,j}^{OUT}y_j\right)\right) = y_i\left(P_i + V_i + O_i + \sum_{a\in A}\left(M_{a,i}^{IN}F_a\right)\right), \forall i \in \mathrm{I}. \quad (24)$$

3.5. Interval estimates

The approach presented in the previous chapter provides only a point estimate, but it is useful to have information about variability and corresponding interval estimates when evaluating results. Since there is no information available on the standard deviation, the construction of interval estimates is realized through an optimization calculation that determines the limits of the possible proportion of dry matter in slurry. Therefore, the term interval estimate can be more likely considered as feasible boundaries for dry matter content.

The calculation is based on a network flow problem with multiple commodities, where each waste producer generates its own commodity. By minimizing or maximizing the proportion of dry matter at a selected node, the corresponding limits can be obtained. However, this construction is computationally intensive, especially in terms of computer memory. At the same time, it is necessary to perform calculations for a large number of scenarios, which is twice the number of entities with waste treatment in the system being analyzed. To maintain linearity, it was necessary to link the proportion of dry matter in secondary production to the national average, but from the perspective of the total amount of waste under consideration, the secondary production represents a negligible change. This also allowed for a well-arranged notation for the different waste production sources, which are indexed here compared to the previous chapter. The following is a mathematical model with a description of the symbols used,

Sets

$a \in A$ set of arcs
$i,j \in I$ set of nodes
$n \in N$ set of waste sources

Parameters

$D_{j,n}$ percentage content of dry matter in produced slurry in node j from source n, [-]
F_a amount of slurry transported along arc a, [kg]
$M_{a,i}$ incidence matrix determining input and output of all arcs, [-]
$P_{i,j,n}$ amount of slurry produced primarily in node i ($i = j$) from source n, [kg]
T_i amount of slurry treated in node i, [kg]

Variables

$t_{i,j,n}$ amount of slurry from producer j treated in node i, [kg]
$x_{a,j,n}$ amount of slurry from producer j and source n transported along arc a, [kg]
y_i percentage content of dry matter in slurry in node i, [-]

The objective function Eq. (25) is defined by the percentage dry matter content of treated wastewater sludge in a selected entity, which is minimized or maximized to obtain desired boundaries.

$$\min\ y_i \ \text{ or } \ \max\ y_i \quad (25)$$

The following constraint Eq. (26) represents maximum flow along an arc, which capacity F_a is defined by results from DR. The sum of each commodity and different sources transported along an arc must be equal to the given flow.

$$\sum_{j\in I}\sum_{n\in N} x_{a,j,n} = F_a, \forall a \in \mathrm{A}. \quad (26)$$

Eq. (27) describes the mass balance for each commodity in every entity in the system. Production $P_{i,j,n}$ with the sum of received and transferred waste $M_{a,i}x_{a,j,n}$ must be equal to treatment $t_{i,j,n}$. The sum of treated waste from various producers must be equal to the total treatment T_i in the node determined by DR, which is expressed by Eq. (28).

$$P_{i,j,n} + \sum_{a\in A} M_{a,i}x_{a,j,n} - t_{i,j,n} = 0, \forall i \in \mathrm{I}. \quad (27)$$

$$\sum_{j\in I}\sum_{n\in N} t_{i,j,n} = T_i, \forall i \in \mathrm{I}. \quad (28)$$

The percentage dry matter of treated slurry is calculated by Eq. (29). The sum of all dry treated dry matter $D_{j,n}t_{i,j,n}$ divided by the amount of slurry T_i is equal to the percentage value.

$$T_iy_i = \sum_{j\in I}\sum_{n\in N} D_{j,n}t_{i,j,n}, \forall i \in \mathrm{I}. \quad (29)$$

The last condition for the model is the nonnegativity of variables, which is expressed by Eq. (30) and Eq. (31). These two constraints ensure also the nonnegativity of y_i.

$$t_{i,j,n} \geq 0, \forall i \in \mathrm{I}, \forall j \in \mathrm{I}, \forall n \in \mathrm{N}. \quad (30)$$

$$x_{a,j,n} \geq 0, \forall a \in \mathrm{A}, \forall j \in \mathrm{I}, \forall n \in \mathrm{N}. \quad (31)$$

4. Results of the case study in the Czech Republic

The presented approach, consisting of a combination of several mathematical methods, is subsequently applied to a case study in the Czech Republic. The aim is to evaluate the treatment of the dry matter in waste sludges, according to the waste catalogue (SEPA, 2015), which corresponds to code 19 08 05, "sludges from treatment of urban wastewater." The ISOH database (ISOH, 2023) is used, which collects all information on waste production and handling in the Czech Republic.

Some aggregated information is publicly available, but the necessary structure of data for the approach verification is secret. Thanks to research collaboration with the CENIA agency (CENIA, 2023), the related database was provided in the form of raw data, exactly as reported by individual entities. The detail of data follows the lowest local administrative units, which are municipalities with their own authority. In the case of investigated wastewater sludges, there are approximately 1500 units, which can be further divided into entities based on juridical classification (municipality, company, citizen).

4.1. Database reconstruction

The preprocessing phase aims to identify errors in the system and provide basic statistics. Overall, 954 kt of slurry are reported as production, while the reported treatment is 971 kt. The total difference is 17 kt, which does not have a clearly defined origin at the point of treatment. At first glance, it may seem like an acceptable difference (approximately 2%), but when looking at the task in greater detail at the level of individual entities, larger relative errors become apparent. Often, there is only a record of production, and the subsequent information on handling is missing, or some entities in the system are completely missing, even though there are records, which confirm their participation in WM of wastewater sludge.

In the case of waste transfer and reception analysis, four categories can be considered. The first is related to the case when the transfer record shows the same amount as the receipt record and both partners of the transactions are correctly identified, making it possible to match them. This type represents the only correct case of reporting waste handling. Another type is transportation where records can be paired, but the reported amounts on both sides do not match. The most problematic type of reporting is missing evidence, where one side reports a transfer, but the other side does not. This case can be divided into two types, either where only the receipt record exists or only the transfer record exists. The subsequent analysis of the transported amounts of sludge in 2020 is graphically displayed in Fig. 4. The analysis is stated in the form of the raw database and after pre-processing, utilizes the correction principle explained in Fig. 2.

The graph shows that a more detailed analysis at the level of individual entities revealed over 30% discrepancies in terms of quantity. Specifically, over 10% of the transported waste does not have a corresponding record. Moreover, it cannot be assumed that all of these errors can be paired with each other, which also suggests differences in the quantity between transferred and received waste. Another group, where there are paired records, but the quantities do not match, represents about 20% of the transport records. It could be assumed that the waste changed its properties during transport (e.g., water evaporation), but in many cases, this represents a change in the order of (units vs thousands) tons, which could subsequently lead to unrealistic numbers greater than 100% in terms of dry matter monitoring.

Data pre-processing results in significant correction of incorrect transport inconsistencies, particularly by greatly assisting in the matching of transactions that had no pair record in the original database. At the same time, this approach has also reconciled the waste transfers where different values are reported. It should be mentioned that all of the pre-processing correction do not modify production or treatment and it is achieved with only proper redirections of transport records. All remaining system errors have been solved using pre-processing corrections and DR, and the mass balances are now met at all levels of the system. Manual verification of corrections is not possible for all error cases. There are inconsistencies where it is not possible for an expert to determine the truth, as it is impossible to assess multiple chains simultaneously that can influence each other. The discrepancies, which can be corrected in the pre-processing, are mostly corrected in the same way by the DR, except for cases where the same entity shows waste transfer to itself. In such a case, from the perspective of the node mass balance, it is a plus and minus record of the same amount, meaning it does not affect the result, and a randomly selected value between recorded values is selected. Therefore, it is always beneficial to perform pre-processing, even though it represents a negligible part of the evidence.

4.2. Dry matter evaluation

As outlined in the introduction of this paper, in the Czech Republic, the evaluation of treatment indicators currently relies on the national average without considering possible sources and varying quality of treated waste. In order for this approach to be considered acceptable (i.e. deviations from reality will not be significant), the histogram must follow a normal distribution with the least amount of variability possible. To verify this approach, individual productions are analyzed and the amount of wastewater sludges categorized by the proportion of dry matter are shown in Fig. 5.

From the histogram, it is apparent that the variability of the dry matter ratio is considerable, with the data roughly following a bimodal distribution with the mean between the peaks. In such a case, calculating treatment indicators based on the national average dry matter ratio can

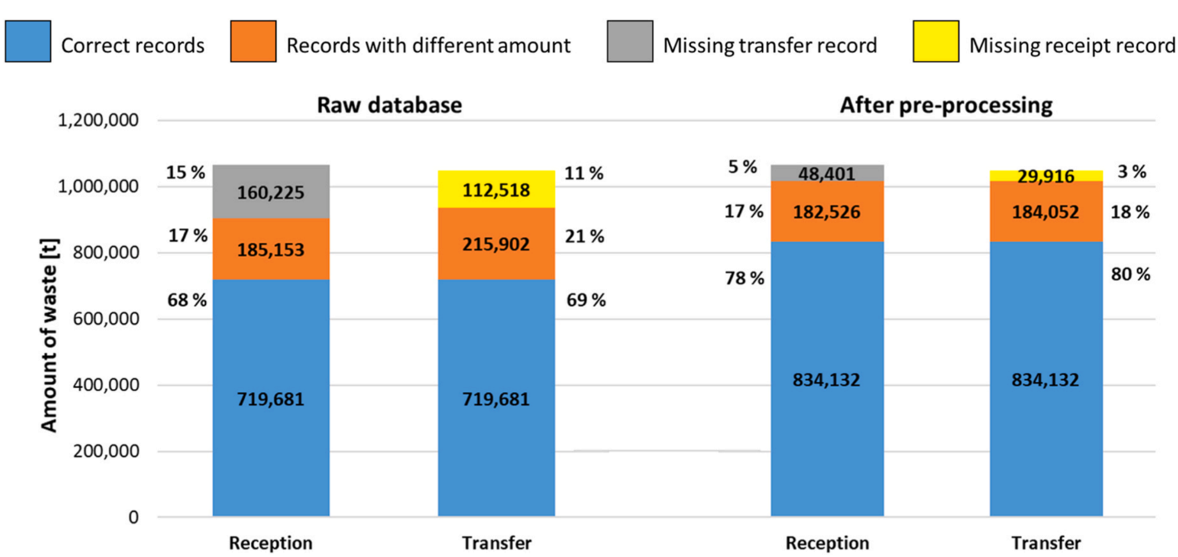


Fig. 4. Analysis of transferred – received waste in the system.

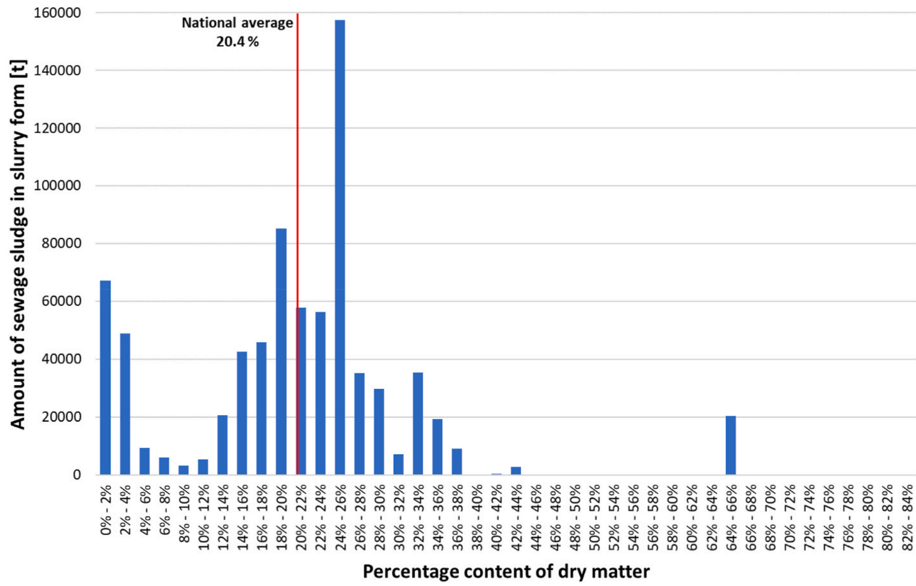


Fig. 5. Amount of wastewater sludge in slurry form sorted by percentage content of dry matter.

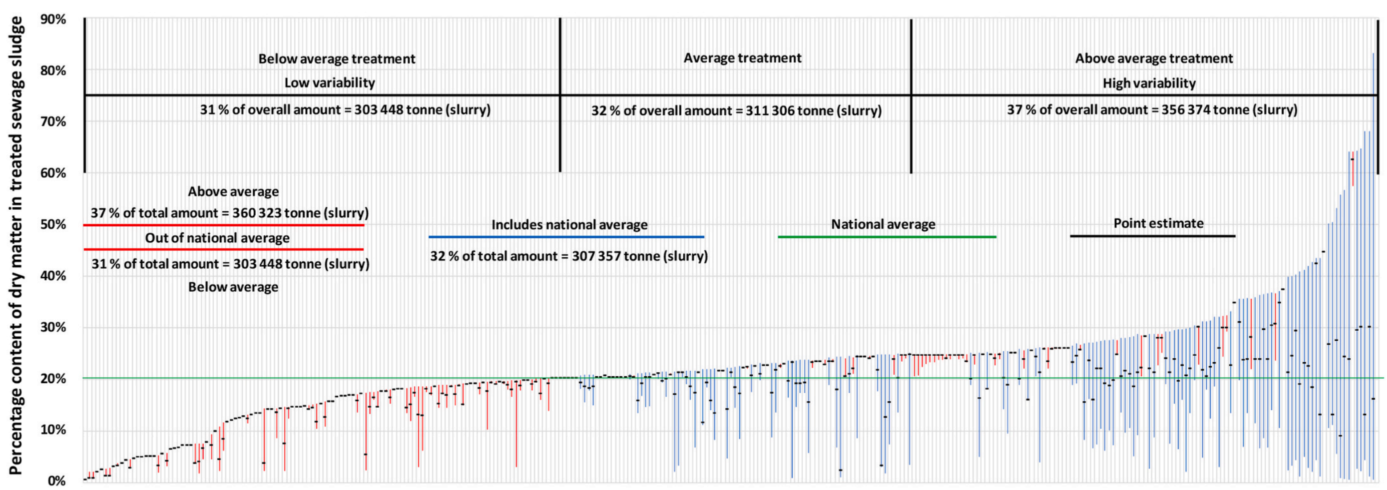
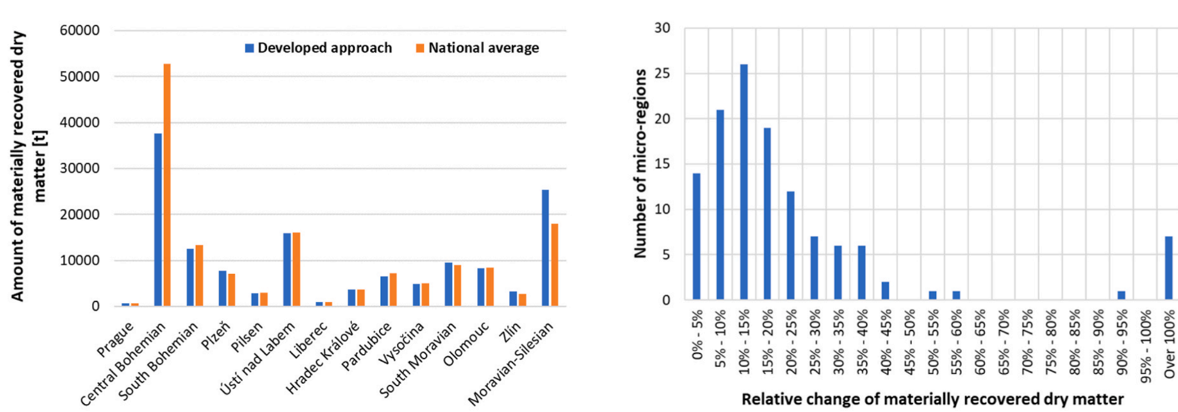


Fig. 6. Point and interval estimates of dry matter ratio of entities reporting final treatment.



a: Absolute amount of recovered dry matter in regions.

b: Histogram of the relative change of recovered dry matter in micro-regions.

Fig. 7. Comparison of the developed approach with the national average for material recovery.

be entirely misleading, as the final treatment at the microregional level may be dominated by wastewater sludge with a different dry matter ratio than the calculated average. In this regard, it is worth mentioning the extreme values with almost zero dry matter ratio and with a ratio of approximately 65%. The newly developed approach thus has its justification here, and the comparison with the national average is shown in Fig. 6, where individual entities that report the final treatment are displayed after DR.

It is evident that only one-third of the sludge processing includes the national average of dry matter ratio, which can further greatly influence the evaluation of final handling. The new approach of averaging waste in individual entities always provides an acceptable solution, whereas the national average is only feasible in cases of high variability. This can be caused particularly by cases where an entity represents a transit node where a large amount of waste flows through, but only a minimal part is treated, which can contain an extreme ratio of dry matter. Furthermore, it can be mentioned that the below-average and above-average dry matter ratios make up approximately the same proportion. Whether this will affect the final evaluation of indicators depends on which part is materially recovered. The following graphs in the Fig. 7 show the individual regions and a comparison of the material use of dry solids.

The overall material recovery of wastewater sludge decreased by 6% in the Czech Republic, which can be considered a significant change in the case of achieving ambitious goals of the circular economy. These differences are even more evident in individual regions. Specifically, the Central Bohemian region utilizes the dry matter from wastewater sludge materially by 30% less, while the Moravian-Silesian region shows an increase of 40% compared to the national average. At the level of micro-regions, even more significant relative deviations can be observed, even exceeding 100%. This information can be a key to identifying inefficient areas and support in developing necessary infrastructure.

4.3. Results discussion

One of the significant advantages of this comprehensive step-by-step approach is the standardized methodology employed for data correction and the ability to shift evaluations towards a more detailed level. From the reconstruction point of view, it should not be expected that results are exactly identical to what happened in reality. However, this cannot be expected even in the case of the raw data without inconsistencies. For large companies in the WM industry, the database shows that waste transport is reported across the entire state where the company is headquartered, while in reality, the waste does not leave the origin region, where the subsidiary of the company is located. A similar case can be observed for waste collection companies, which often collect waste from multiple municipalities at once and then reallocate the waste proportionally or with a methodology in order to reduce landfill charges. Correcting the database, however, should improve the informative value of the database and also allow for the application of more detailed methodologies in analyses.

The following evaluation of specific content in the investigated waste stream is particularly relevant when calculating waste management indicators. In the case of wastewater sludge, the evaluation should be based on the treatment of dry matter, which contains nutrients suitable for material recovery in agriculture. However, the performance in achieving recycling goals and assessment of the treatment is nowadays calculated with water content. It may lead to wrong conclusions and therefore, this indicator does not provide accurate information for decision-making, especially at the microregional level. The developed approach enables a more accurate assessment of the situation and identifies specific areas where a considerable amount of dry matter remains unused. This information becomes instrumental in improving infrastructure and implementing targeted interventions to optimize resource utilization in those locations. Differences between results of indicator evaluation can be seen in Fig. 7, which are negligible in national aggregated form, but in the detail of micro-regions represent a significant change. However, it is crucial to acknowledge that even with the current methodology, there is still a level of simplification involved. Wastewater sludge flows are averaged at nodes, which does not fully capture the complexity of the system and behaviour in real-world conditions. Recognizing this limitation, ongoing research efforts are directed towards avoiding such simplification and developing innovative approaches to uncover the entire flow of sludge from the producer to the processor using available data sources.

5. Conclusion

The presented paper deals with the administration of data on WM monitoring at the national level, which is unique in terms of the volume of data handled, its detail and character. The aim is to provide a general framework that would help to eliminate inconsistencies and errors in data and thus enable more detailed analyses, which is often nowadays evaluated in an aggregated form. The paper's contribution lies in the integration of various mathematical techniques based on statistics, DR, and network flow problems, which enables effective step-by-step evaluation at the smallest administrative unit level both in terms of waste quantity and quality. Fixing mass errors among all entities in the system with DR is improved by tailor-made covariance matrix and the results of material flow analysis for waste composition is supported by point and interval estimates. By employing this novel approach, inefficient regions or processes can be identified, allowing for appropriate planning of the required infrastructure.

The presented approach is further applied to a case study evaluating wastewater sludge in the Czech Republic. Based on detailed data, significant inconsistencies in waste handover and transfer are identified, resulting in a difference of 17 kt between production and final treatment. These inconsistencies are corrected using DR and expert-designed weights. Subsequently, the handling of dry matter is evaluated in detail for individual entities, and the results are compared with the currently used nationwide average. The results showed that the nationwide average for two-thirds of entities is not within a feasible set. Overall, the material utilization indicator is examined, which, according to the developed approach, identified a 6% worse handling of waste in the Czech Republic, while deviations in the range of tens to hundreds of percent are found when analyzing smaller territorial units.

The future development will be focused on addressing other waste streams, which can help define further conditions for the covariance matrix. As outlined in the methodology, defining values out of the diagonal may be considered since many entities in the database may be mutually influenced. It is necessary to find a suitable criterion based on available information, which helps to estimate the presence of a connection between entities. Another part of the research will aim to identify the entire flow from the producer to the processor or provide a relevant estimate of where waste may have ended up with a high probability. This will enable a more precise evaluation of indicators and analysis of waste quality.

CRediT authorship contribution statement

Jaroslav Pluskal: Formal analysis, Methodology, Writing – original draft, Writing – review & editing. **Radovan Šomplák:** Conceptualization, Validation, Writing – review & editing. **Lucie Němcová:** Formal analysis, Writing – original draft. **Jiří Valta:** Resources, Investigation. **Martin Pavlas:** Supervision, Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

The authors do not have permission to share data.

Acknowledgements

The authors 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". This work was also supported by Grant No. GA 22-11867S of the Czech Science Foundation. The support is gratefully acknowledged.

Abbreviations

DR	Data reconciliation
WM	Waste management

References

Abubakar, I.R., Maniruzzaman, K.M., Dano, U.L., AlShihri, F.S., AlShammari, M.S., Ahmed, S.M.S., Al-Gehlani, W.A.G., Alrawaf, T.I., 2022. Environmental sustainability impacts of solid waste management practices in the global south. Int. J. Environ. Res. Publ. Health 19 (19), 12717. https://doi.org/10.3390/ijerph191912717.

Badings, T.S., van Putten, D.S., 2020. Data validation and reconciliation for error correction and gross error detection in multiphase allocation systems. J. Petrol. Sci. Eng. 195, 107567 https://doi.org/10.1016/j.petrol.2020.107567.

Behnami, A., Zoroufchi Benis, K., Shakerkhatibi, M., Fatehifar, E., Derafshi, S., Chavoshbashi, M.M., 2019. Integrating data reconciliation into material flow cost accounting: the case of a petrochemical wastewater treatment plant. J. Clean. Prod. 218, 616–628. https://doi.org/10.1016/j.jclepro.2019.01.218.

Bennamoun, L., Arlabosse, P., Léonard, A., 2013. Review on fundamental aspect of application of drying process to wastewater sludge. Renew. Sustain. Energy Rev. 28, 29–43. https://doi.org/10.1016/j.rser.2013.07.043.

Bowman, G., Ayed, L., Burg, V., 2022. Material and energy flows of industrial biogas plants in Switzerland in the context of the circular economy. Bioresour. Technol. Rep. 20, 101273 https://doi.org/10.1016/j.biteb.2022.101273.

Brown, A.W., Kaiser, K.A., Allison, D.B., 2018. Issues with data and analyses: errors, underlying themes, and potential solutions. Proc. Natl. Acad. Sci. USA 115 (11), 2563–2570. https://doi.org/10.1073/pnas.1708279115.

Câmara, M., Soares, R., Feital, T., Anzai, T., Diehl, F., Thompson, P., Pinto, J., 2017. Numerical aspects of data reconciliation in industrial applications. Processes 5 (4), 56. https://doi.org/10.3390/pr5040056.

CENIA, 2023. In: The Czech Environment Information Agency. https://www.cenia.cz/czech-environmental-information-agency/organisation-profile/. (Accessed 25 January 2023).

Chao, G., Sun, J., Lu, J., Wang, A.-L., Langleben, D.D., Li, C.-S., Bi, J., 2019. Multi-view cluster analysis with incomplete data to understand treatment effects. Inf. Sci. 494, 278–293. https://doi.org/10.1016/j.ins.2019.04.039.

Chen, Y., Yuan, Z., Chen, B., 2018. Process optimization with consideration of uncertainties—an overview. Chin. J. Chem. Eng. 26 (8), 1700–1706. https://doi.org/10.1016/j.cjche.2017.09.010.

Cochinwala, M., Kurien, V., Lalk, G., Shasha, D., 2001. Efficient data reconciliation. Inf. Sci. 137 (1–4), 1–15. https://doi.org/10.1016/S0020-0255(00)00070-0.

di Fonzo, T., Girolimetto, D., 2022. Forecast combination-based forecast reconciliation: insights and extensions. Int. J. Forecast. https://doi.org/10.1016/J.IJFORECAST.2022.07.001.

Eghbali, H., Arkat, J., Tavakkoli-Moghaddam, R., 2022. Sustainable supply chain network design for municipal solid waste management: a case study. J. Clean. Prod. 381, 135211 https://doi.org/10.1016/j.jclepro.2022.135211.

European Commission, 2022. Draft decree on details of the management of certain end-of-life products. In: Internal Market, Industry, Entrepreneurship and SMEs. https://ec.europa.eu/growth/tools-databases/tris/en/search/?trisaction=search.detail&year=2021&num=343. (Accessed 24 January 2023).

Gurevich, Y.G., Gurevich, A., Gurevich, H., 2022. Geometrical interpretation of data reconciliation and uncertainty reduction in multi-sensor systems. SSRN Electron. J. https://doi.org/10.2139/ssrn.4098750.

Halecki, W., Gasiorek, M., Gambus, F., Abram, R., 2016. The potential of hydrated and dehydrated sewage sludge discharges from soil reclamation appliances. Fresenius Environ. Bull. 25, 1935–1941.

Hu, G., Xu, L., Zhang, Z., 2022. Correntropy based Elman neural network for dynamic data reconciliation with gross errors. J. Taiwan Inst. Chem. Eng. 140, 104568 https://doi.org/10.1016/J.JTICE.2022.104568.

Islam, M.T., Huda, N., 2019. Material flow analysis (MFA) as a strategic tool in E-waste management: applications, trends and future directions. J. Environ. Manag. 244, 344–361. https://doi.org/10.1016/j.jenvman.2019.05.062.

ISOH, 2023. In: Waste Management Information System. https://www.cenia.cz/odpadove-a-obehove-hospodarstvi/isoh/. (Accessed 25 January 2023).

Kacprzak, M., Neczaj, E., Fijałkowski, K., Grobelak, A., Grosser, A., Worwag, M., Rorat, A., Brattebo, H., Almås, Å., Singh, B.R., 2017. Sewage sludge disposal strategies for sustainable development. Environ. Res. 156, 39–46. https://doi.org/10.1016/j.envres.2017.03.010.

Kim, M., Chang, J.W., Park, K., Yang, D.R., 2022. Comprehensive assessment of the effects of operating conditions on membrane intrinsic parameters of forward osmosis (FO) based on principal component analysis (PCA). J. Membr. Sci. 641, 119909 https://doi.org/10.1016/j.memsci.2021.119909.

Kuehn, D.R., Davidson, H., 1961. Computer control II. Mathematics of control. Chem. Eng. Process 57, 44–47.

Kumar, A., Samadder, S.R., 2022. Assessment of energy recovery potential and analysis of environmental impacts of waste to energy options using life cycle assessment. J. Clean. Prod. 365, 132854 https://doi.org/10.1016/j.jclepro.2022.132854.

Li, J., Panchabikesan, K., Yu, Z., Haghighat, F., Mankibi, M. el, Corgier, D., 2019. Systematic data mining-based framework to discover potential energy waste patterns in residential buildings. Energy Build. 199, 562–578. https://doi.org/10.1016/J.ENBUILD.2019.07.032.

Li, P., Shih, H., Ma, H., 2022. Applying probabilistic material flow analysis for quality control and management of waste recycling in steelmaking. Waste Manag. 144, 67–75. https://doi.org/10.1016/j.wasman.2022.03.011.

Pavlas, M., Šomplák, R., Smejkalová, V., Nevrlý, V., Zavíralová, L., Kůdela, J., Popela, P., 2017. Spatially distributed production data for supply chain models - forecasting with hazardous waste. J. Clean. Prod. 161, 1317–1328. https://doi.org/10.1016/J.JCLEPRO.2017.06.107.

SEPA, 2015. In: Guidance on Using the European Waste Catalogue (EWC) to Code Waste. Scottish Environment Protection Agency. https://www.sepa.org.uk/media/163421/ewc_guidance.pdf. (Accessed 10 February 2023).

Sepasgozar, S.M.E., Frances Mair, D., Tahmasebinia, F., Shirowzhan, S., Li, H., Richter, A., Yang, L., Xu, S., 2021. Waste management and possible directions of utilising digital technologies in the construction context. J. Clean. Prod. 324, 129095 https://doi.org/10.1016/j.jclepro.2021.129095.

Sharma, R., Agrawal, D., Kodamana, H., 2022. Data reconciliation frameworks for dynamic operation of hybrid renewable energy systems. ISA (Instrum. Soc. Am.) Trans. 128, 424–436. https://doi.org/10.1016/j.isatra.2021.12.006.

Sileryte, R., Sabbe, A., Bouzas, V., Meister, K., Wandl, A., van Timmeren, A., 2022. European waste statistics data for a circular economy monitor: opportunities and limitations from the Amsterdam metropolitan region. J. Clean. Prod. 358, 131767 https://doi.org/10.1016/j.jclepro.2022.131767.

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.

Šomplák, R., Nevrlý, V., Smejkalová, V., Šmídová, Z., Pavlas, M., 2019. Bulky waste for energy recovery: analysis of spatial distribution. Energy 181, 827–839. https://doi.org/10.1016/j.energy.2019.05.175.

STOA, 2017. In: Towards a Circular Economy – Waste Management in the EU. European Parliamentary Research Service. Study: IP/G/STOA/FWC/2013-001/LOT 3/C3. https://www.europarl.europa.eu/RegData/etudes/STUD/2017/581913/EPRS_STU%282017%29581913_EN.pdf. (Accessed 10 February 2023).

Tsai, F.M., Bui, T.-D., Tseng, M.-L., Lim, M.K., Hu, J., 2020. Municipal solid waste management in a circular economy: a data-driven bibliometric analysis. J. Clean. Prod. 275, 124132 https://doi.org/10.1016/j.jclepro.2020.124132.

Yu, J., Han, W., Chen, K., Liu, P., Li, Z., 2022. Gross error detection in steam turbine measurements based on data reconciliation of inequality constraints. Energy 253, 124009. https://doi.org/10.1016/j.energy.2022.124009.

Zhan, L., Zhao, R., Wu, Y., Zeng, S., Yuan, Y., 2022. Construction of a spatial–temporal metabolic path for hazardous waste management based on the fusion of reported data and web text data. Environ. Technol. Innovat. 28, 102541 https://doi.org/10.1016/J.ETI.2022.102541.