CHEMICAL ENGINEERING TRANSACTIONS

VOL. 56, 2017

A publication of AIDIC
The Italian Association of Chemical Engineering
Online at www.aidic.it/cet

Guest Editors: Jiří Jaromír Klemeš, Peng Yen Liew, Wai Shin Ho, Jeng Shiun Lim

ISBN 978-88-95608-47-1; **ISSN** 2283-9216

DOI: 10.3303/CET1756321

Input Data Validation for Complex Supply Chain Models Applied to Waste Management

Radovan Šomplák*[,a], Lenka Zavíralová[a], Martin Pavlas[a], Vlastimír Nevrlý[b], Pavel Popela[b]

[a]Institute of Process Engineering, Faculty of Mechanical Engineering, Brno University of Technology Technická 2896/2, Brno, 616 69, Czech Republic
[b]Institute of Mathematics, Faculty of Mechanical Engineering, Brno University of Technology, Technická 2896/2, 616 69 Brno, Czech Republic
somplak@upei.fme.vutbr.cz

This contribution focuses on comprehensive input data analysis prior to the application of complex network flow models. Network flow models are composed of vertices interconnected by edges, the number of which is dependent on particular problem studied. Any quantitative data in vertices may contain inaccurate information. The article is intent on computational system for simulation and forecasting in waste management incomplete data problems. The article introduces enhancements to the mathematical model of previously published tool and its elements. The focus is put on the suitable choice of weights for all input forecasting models and territorial divisions of relevant regions. Another model improvement is based on the support of adding new constraints that link multiple types of waste. The next part of article concentrates on forecasting of waste production by the application of the advanced model. The case study focused on the estimation of trends in data for municipal waste and composition of residual municipal solid waste in the Czech Republic is presented as well.

1. Introduction

Nowadays, there is an ongoing transformation of the implementation of waste management related law and principles in many countries of the EU. Possible division of EU countries into three groups according to the current level of the waste management maturity and according to the current challenges that these countries must face (diversion from landfilling, higher recovery rates, etc.) is described and discussed in Šomplák et al. (2014). There have been several modelling approaches developed, mostly based on mathematical programming, created in order to take into account interactions that occur in the tasks associated with the collection and processing of the waste Ghiani et al. (2014). The complex problem statement describing the competitive environment in the field of the waste management was presented in Šomplák et al. (2015). The effective approach for solving waste management challenges in non-EU countries is detailed in Santibañez-Aguilar et al. (2014). Nga et al. (2013) enriched the discussion on the economic aspects with the analysis of environmental impacts of waste-to-energy (WTE). The above mentioned articles represent different applications of so called supply chain modelling approach. Regarding the decision support of effective planning in the waste management area, among the recent major challenges we can find these related to incomplete input data and varying quality of the forecasts. Typically there is a waste production (of different types of waste) specified for all nodes that are included in the supply chain model. In case the supply chain is used for stable infrastructure planning (e.g., for allocation of waste processing units), there is a need to work with forecasts of the waste production, as well as, other input parameters. An interesting methodology for the modeling of future production of electrical waste has been introduced in Wang et al. (2013). This article introduces several possible approaches for creating the model of production. The idea that every approach may provide valuable information is considered. Thus the model which takes into account several kinds of available data on the production of electro waste was designed. However, the problem of dealing with multiple available models forming different input parameter values is general. Therefore, this article describes the original method how to deal with the problem of different input data generated by various models for the same location. An introduced computational tool allows taking into account all the available forecasting models and maximises the important outcome information to determine the final amount of waste that will be used in the calculation. Moreover, there is also considered: (1) hierarchical territorial division (region is divided into subregions and furthermore in their parts) and (2) the ability to add relations between the different types of waste (e.g. the more waste is separated, the less residual waste is available). From the mathematical point of view, the designed and implemented model is in the form of a mathematical program, which is based on finding a compromise between results obtained from various models forecasting the waste production. The computational tool implementing this mathematical program and related models was named *Justine*. The main idea and the initial demonstration of its application was presented in the article Zavíralová et al. (2015). The specific goal is to find a compromise between the available data sets with data often distorted by errors for certain nodes or completely missing for other nodes. In addition, mass balances within the territorial hierarchy must be met (the sum of waste production in areas of lower level of the hierarchy (subregion) must be equal to the waste production of higher level of hierarchy (region)). These additional equalities have to apply not only to historical data but also to forecasts, see model (1) in section 3. The need to deal with the differences between the measured values and the balance relations arising from mass and energy conservation law have already been used e.g. in article Farsang et al. (2014) focusing on the monitoring of energy consumption in chemical processes. Here Farsang only detects potential operational problems coming from the difference in measurement values, which are remote from the balance model. In this article, the method for update of inaccurate data for the waste management area based on the current balance of relations is presented. In the field of waste management a number of articles deal with the waste production forecasts for detailed time steps (e.g. Noori et al. 2009 in time steps of one week). The extensive data processing for the purpose of predicting the waste production in Finland is described in Korhonen et al. (2015). Obviously, socio-economic factors can significantly affect the amount of the generated waste. For example, the influence of tourism on the production of waste is discussed in Arbulú et al. (2015). Each of these approaches is interesting and has its strengths and weaknesses and it provides valuable information that may be relevant to the production estimate in the individual region. In our case of the use of multiple methods to estimate the generation of waste, it is also necessary to deal with various inconsistencies of individual forecasting models from the expert's point of view. Therefore, the presented optimisation tool *Justine* working with different models and related forecasts and estimates allows assignment of different weights by the experts who check their validity and credibility.

The tool *Justine*, presented in Zavíralová et al. (2015), described data tuning and generation related to the amount of the residual waste and its lower heating value (LHV) for the territory covering 230 nodes. It was assumed, that there is a relation between amount of this waste produced and its lower heating value (an effect of sorting described later on). The analysis focused only on one type of the waste, which was residual solid waste (RSW). For the computational purposes, the energy conservation was assumed, i.e. the amount of the waste multiplied by LHV is equal to energy potential. This equality must remain valid within the territorial hierarchy (the sum of energy contributions of the lower territorial units is equal to the energy for higher territorial unit, which includes the lower ones). Despite of limited amount of input data, the practical considerations show that the obtained results provide information about the likely realistic state of the RSW production and LHV in all considered territorial units. This information represents the valuable inputs for the related supply chain optimisation models, e.g., for designing new waste management facilities.

This article will at first describe the improvements of the mathematical model from Zavíralová et al. (2015) focusing on an appropriate setting of weights for individual territorial units. Another improvement is based on the support of possibility of adding new critical relations, which enable to link multiple types of waste (e.g. separated and residual). The next part of article will be devoted to forecasting of production of the waste applying the advanced model. In the last part of the article, there are results for the case study for residual municipal waste in the Czech Republic provided by the tool *Justine*.

2. Development of Mathematical Model

In this section there are elaborated areas of developed mathematical model, which are the major improvements of the tool *Justine*, introduced by Zavíralová et al. (2015).

The first part of this section describes the weighting approach for individual input data and the process of dealing with heteroscedasticity of the data. The following approach is valid in general for any type of territorial unit and any type of the waste.

With respect to the fact that territorial units have different area, population and thus different waste production, the estimated errors influence the objective function with various significance. For minimisation of the sum of square errors, this tool would give preferences to the reduced error in bigger territorial units, and so errors in smaller territorial units can increase. For this reason, it is necessary to design a system of weights for individual errors in order to be able to minimise the impacts of these errors almost uniformly. Since there are errors for various data types, these calculations of weights must be performed for all of them.

The goal in the weights construction process is to make all input data equally significant. Several approaches to solve this problem have been applied and related tests have been performed. The most suitable one was chosen by using the statistical regression where these weights were constructed in order to significantly reduce the heteroscedasticity of input data. In our case, this effect is achieved by using a square of weights in objective function for each territory and data type, where the weights are the inverse of average. By applying this procedure, the weights for all territorial units and particular data types look as follows:

$$w_{it} = \begin{cases} \frac{\sum_{d\in D} \delta_{td}^{T} \delta_{id}^{ID}}{\sum_{d\in D} \delta_{td}^{T} c_{id}} & \text{for nonzero } \sum_{d\in D} \delta_{td}^{T} c_{id} \\ 0, & \text{for } \sum_{d\in D} \delta_{td}^{T} c_{id} = 0 \end{cases} \quad (1)$$

where the following notation is utilised in (1) and (2):

sets and indices

$i \in I$	index of territorial unit
$d \in D$	index of particular data set
$t \in T$	index determining type of value
$k \in K \subset T$	index from subset of T
$l \in L \subset T$	index from subset of T
parameters	
δ_i^{I2}	$\{0,1\}$ indicator of territorial unit, equals 1 if territory i is in use
δ_{id}^{ID}	$\{0,1\}$ indicator of data availability for territory i and data set d
δ_{td}^{T}	$\{0,1\}$ indicator of data set type, equals 1 if data set d is of type t
c_{id}	two-dimensional parameter containing data for territory i and data set d
variables	
m_{it}	value for territory i and type t

The following part of this section describes how several types of waste are combined into one calculation. In the waste management area there are strong interactions between different types of waste. Municipal solid waste (MSW) fractions can be a good example. Assuming constant production, the more of the waste is sorted and separately collected, the less remains in residual solid waste (RSW). Moreover, the hierarchy of types of waste may be used in this example. E.g. MSW produced is available as separately collected (SEP) and RSW therefore the following relationship holds for the amounts of specified waste $MSW = SEP + RSW$. SEP can be further divided into individual fractions, e.g. paper (PAP), plastic (PLA), glass (GLA), and others (OTH). Therefore the following equation is valid for related amounts of waste types $SEP = PAP + PLA + GLA + OTH$. For all of these waste types and related supersets and subsets, it is possible to forecast the waste production (see Section 3). In the mathematical model, these equations are described by the general Eq(2).

$$\delta_i^{I2} \sum_{k\in K} m_{ik} = \delta_i^{I2} \sum_{l\in L} m_{il} \quad (2)$$

where sets *K* and *L* denote arbitrary (but meaningful) selected grouping of individual types of waste (e.g. paper, plastic, RSW, SEP components etc.).

3. Trend Analysis - Forecasting

There is a need to ensure appropriate quality of input data in order to obtain the output from the tool Justine which would represent a feasible input to the supply chain. This data may not be completely accurate (and dealing with this inaccuracy is a main benefit of tool Justine), but it is required that the collected data statistically well represent the uncertain real-world data for individual territorial units. This aspect can be usually relatively easily guaranteed for measured data (historical data). The situation becomes more complicated for the parameters that need to be forecasted. There are different approaches how to deal with this problem, having pros and cons related to the corresponding application. The main advantage of the tool Justine is its possibility to consider any number of the forecasting models for the computations, which is limited only by the time-consumption of the calculation process.

For smaller territorial units (county or municipalities), generally, there is a large variability in the historical data for the waste production. In combination with a short time series (data available for the Czech Republic is only for period 2009 - 2015, the reason is a change of processing methodology of historical data at 2009) there is a data fluctuation which restricts the possibility of application of a simple trend analysis by using common tools.

For this reason, it is advisable to correct the results of the detailed analysis carried out at lower territorial units (county) by aggregated data for higher territorial units (regions, the Czech Republic).
Proposed methodology includes the following steps:

- forecasting of individual types of waste (e.g. RSW) with respect to grouping of similar types of the waste (e.g. the SEP = paper + plastic) in all counties,
- forecasting of the waste of certain code number subject to the group of codes for all regions,
- forecasting of the waste of certain code subject to the group of codes for the Czech Republic,

when ensuring the consistency between the forecasts at county, regions and the Czech Republic (sum of forecasts for county in the region is equal to the forecast based on the data for region; analogically for the Czech Republic). In general, links between forecasts for various groups can be illustrated by Figure 1, see territorial units and/or catalog numbers summaries indicated by symbols "Σ", and performance of trend analysis (see symbol "→").

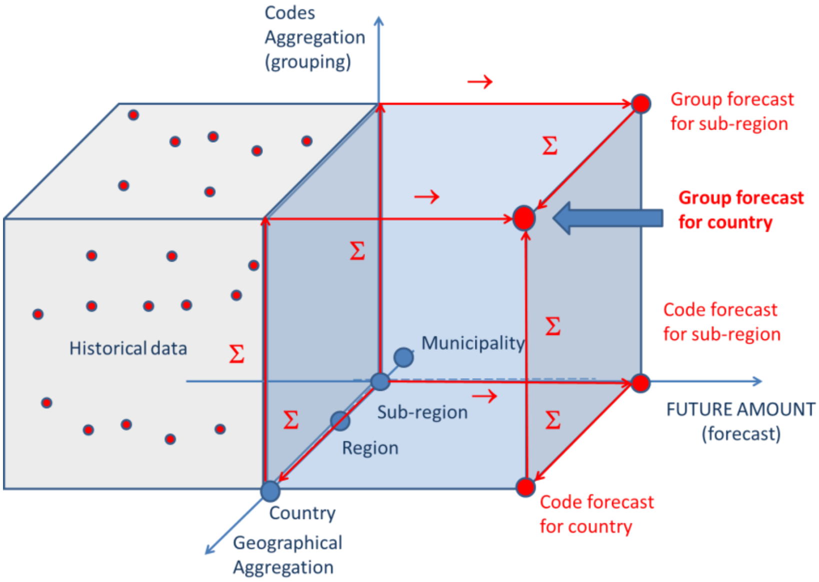


Figure 1: Diagram for considered forecasting approach

These links are included in the optimisation tool *Justine*. Results of trend analysis at all levels of territorial units (MWEP, regions, the Czech Republic) are corrected in such a way that the condition of consistency in the forecasting results is satisfied (does not matter which method will be applied, the results must always be the same, see Figure 1 for visualisation).
Already at the level of trend analysis, it is in general very difficult to define a unified extrapolation model which would accurately describe the situation in all counties, regions and the Czech Republic. Therefore given the nature of the data, there were two possible models considered for the production of all forecasted waste codes and their groups in 2024. Those considered models are as follows. For the trend (model 1 and 2 below) was considered a shifted power function, which is of the following form:

$$y = a + bx^c \quad (1)$$

where the coefficients *a*, *b*, *c* are based on the minimisation of the square deviations (difference between data and model - least squares method, LSM). The shape of the function is partly justified by exploratory data analysis showing that the linear growth over a long period is unrealistic. Models 1 and 2 differ in the order of execution of the forecast:

- Model 1 (denoted ForOfSum) – at first, the geographical aggregation of waste production for the entire region, respectively for the Czech Republic is performed and subsequently the forecast on these data is conducted (symbolically „Σ_G, →").
- Model 2 (denoted SumOfFor) – at first the forecasts are made individually for all counties, and then the sum is performed (symbolically „→, ΣG").

4. Case Study - Production of RSW in the Czech Republic

The aim of this study was to determine (forecast) the potential (the estimated production) of the waste, which would be suitable for waste-to-energy in the year 2024. During the estimation process it was necessary to handle the following steps:

1) Historical data processing - elimination of dubious data, creation of time series for subsequent forecasts.
2) Forecasting for waste type, codes and their groups considered (e.g., separately collected fractions and RSW).
3) Finishing of computation with tool Justine in order to maintain all of required balance equations keeping the change of the data on its minimum.

Resulting forecast of the production of the RSW for the Czech Republic is shown in Figure 2. From the graph it can be seen that the resulting prediction (obtained by using the tool *Justine*) is slightly above the input models. These models 1 and 2 (and computed trends) are based on the analysis of the time series - see Section 3. Shift of prediction to higher value has arisen due to the fact that the sum of the forecasts for input models for all territorial units (separate forecasts for each county) was higher than the forecast for the Czech Republic.

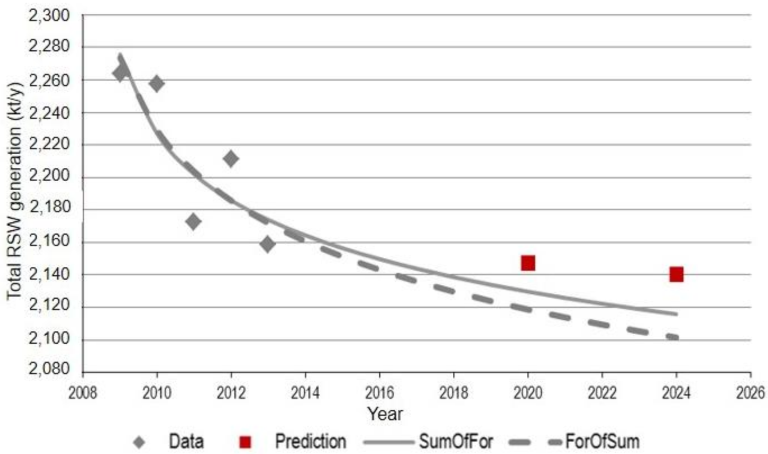
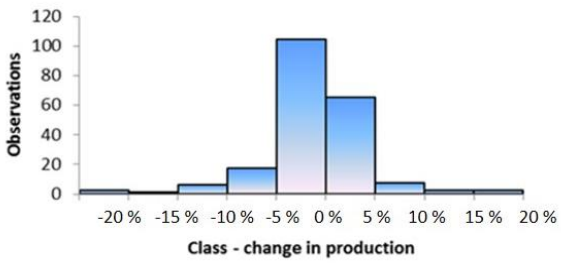


Figure 2: Forecast of RSW production in the Czech Republic

The resulting trend increases, respectively decreases, for each county, which represents the lowest level (the smallest territorial division), are shown in Figure 3. Results for most of county showed no significant change in the production. Despite the resulting overall decrease of production in the Czech Republic, there was an increase in production of RSW for 24 counties, when for 4 of them the increase was even over 10 % of the current production.

Trend	Description	Observations
x > 5 %	increase	11
2,5 % < x < 5 %	moderate increase	13
-2,5 % < x < 2,5 %	stagnation	128
-2,5 % < x < -5 %	moderate decrease	28
x < -5 %	decrease	26



Note: x denotes the percentage change of the production between 2013 (actual data) and 2024 (forecasted year)

Figure 3: Trend analysis of RSW production in the Czech Republic with the focus on county

5. Conclusion

This article shows the approach for forecasting of future trends with an application in the waste management area. The main problem of the modeling in this field is an inconsistency with respect to the waste types and

territorial hierarchy. Another common problem is the availability of different multiple models for one type of waste in one specific location. The article introduces an advanced mathematical model and its improved elements.
Tool *Justine*, which can integrate all available models for different types of the waste considering arbitrary hierarchy of territorial division, was enhanced. This tool suitably adjusts input models such that all the necessary balance equations are met while change of the input data was minimal. The focus is also put on the choice of weights for all forecasting models and territorial divisions of relevant regions. The support of new constraints that link multiple types of waste is included.
Another common problem in this field is the lack of detailed data on the lowest level of the territorial division. However, for some applications this information is essential. Tool *Justine* is able to distribute the data from the higher territorial units to smaller ones while maintaining the properties of smaller territorial units. These properties can be detected based on the regression analysis or by using of incomplete data at lower level of hierarchical territorial distribution of the intended area.

Acknowledgments

The authors gratefully acknowledge financial support provided by MEYS under the National Sustainability Programme I (Project LO1202) and financial support provided by Technology Agency of the Czech Republic within the research project No. TE02000236 "Waste-to-Energy (WtE) Competence Centre.

References

Arbulú I., Lozano J., Rey-Maquieira J., 2015, Tourism and solid waste generation in Europe: A panel data assessment of the Environmental Kuznets Curve, Waste Management 46, 628–636.

Farsang B., Nemeth S., Abonyi J., 2014, Synergy between data reconciliation and principal component analysis in energy monitoring, Chemical Engineering Transactions 39, 721-726.

Ghiani G., Laganà D., Manni E., Musmanno R., Vigo D., 2014, Operations research in solid waste management: a survey of strategic and tactical issues. Computers and Operations Research 44, 22–32.

Korhonen P., Kaila J., 2015, Waste container weighing data processing to create reliable information of household waste generation, Waste Management 39, 15-25.

Ng P.Q., Varbanov P.S., Klemeš J.J., Hegyháti M., Bertók B., Heckl I., Lam H.L., 2013, Waste to energy for small cities: economics versus carbon footprint, Chemical Engineering Transactions 35, 889-894.

Noori R., Abdoli M.A., Ameri Ghasrodashti A., Jalili Ghazizade M., 2009, Prediction of municipal solid waste generation with combination of support vector machine and principal component analysis: A case study of Mashhad, Environmental Progress and Sustainable Energy 28 (2), 249–258.

Santibañez-Aguilar J.E., Ponce-Ortega J.M., González-Campos J.B., Serna-Gonzaléz M., El-Halwagi M.M., 2014, Optimal planning of supply chains for multi-product generation from municipal solid waste, Chemical Engineering Transactions 42, 55-60.

Šomplák R., Pavlas M., Kropáč J., Putna O., Procházka V., 2014, Logistic model-based tool for policy-making towards sustainable waste management, Clean Technologies and Environmental Policy 16 (7), 1275-1286.

Šomplák R., Touš M., Pavlas M., Gregor J., Popela P., Rychtář A., 2015, Multi-commodity network flow model applied to waste processing cost analysis for producers, Chemical Engineering Transactions 45, 733-738.

Wang F., Huisman J., Stevels A., Baldé C. P., 2013, Enhancing e-waste estimates: Improving data quality by multivariate input–output analysis, Waste Management 33 (11), 2397–2407..

Zavíralová L., Šomplák R., Pavlas M., Kropáč J., Popela P., Putna O., Gregor J., 2015, Computational system for simulation and forecasting in waste management incomplete data problems, Chemical Engineering Transactions 45, 763-768.