



HAL
open science

Modelling and Forecasting Waste Generation – DECWASTE Information System

Jiří Hřebíček, Jiří Kalina, Jana Soukopová, Eva Horáková, Jan Prášek, Jiří
Valta

► **To cite this version:**

Jiří Hřebíček, Jiří Kalina, Jana Soukopová, Eva Horáková, Jan Prášek, et al.. Modelling and Forecasting Waste Generation – DECWASTE Information System. 12th International Symposium on Environmental Software Systems (ISESS), May 2017, Zadar, Croatia. pp.433-445, 10.1007/978-3-319-89935-0_36 . hal-01852620

HAL Id: hal-01852620

<https://inria.hal.science/hal-01852620>

Submitted on 2 Aug 2018

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

Modelling and forecasting waste generation – DECWASTE information system

Jiří Hřebíček¹, Jiří Kalina¹, Jana Soukopová², Eva Horáková³, Jan Prášek³ and
Jiří Valta³

¹Masaryk University, Institute of Biostatistics and Analyses,
Kamenice 126/3, 625 00 Brno, Czech Republic

²Masaryk University, Faculty of Economics and Administration,
Lipová 507/41a, 602 00 Brno, Czech Republic

³Czech Environmental Information Agency,
Vršovická 1442/65, 100 10 Prague 10, Czech Republic

hrebicek@iba.muni.cz, kalina@mail.muni.cz, soukopova@econ.muni.cz,
{eva.horakova,jan.prasek,jiri.valta}@cenia.cz

Abstract. The DECWASTE forecasting waste generation information system is presented. It is based on Annexes I, II and III of Regulation (EC) No 849/2010 amending Regulation (EC) No 2150/2002 of the European Parliament and of the European Council on waste (WSR). DECWASTE forecasts the quantity of waste generated for each waste category listed in Section 2(1) of Annex I of the WSR at the Czech national level. Its multi-linear regression forecasting model is based on environmental as well as economic and social predictors. These models use historical data of the waste information system (ISOH) of the Czech Environmental Information Agency and sets of indicators (predictors) integrated into forecasting models. The methodology consisted in adjusting predictors of the forecasting models into a Driving Force-Pressure-State-Impact-Response (DPSIR) framework and their sensitivity analysis enables their choice into forecasting models and their verification using appropriate data. DECWASTE supports decisions made by the Ministry of the Environment to improve the implementation of the national Waste Management Plan.

Keywords: waste generation, waste categories, forecasting, modelling, DPSIR, multi-linear regression models.

1 Introduction

Making decisions in waste management is not only very capital-intensive, but also difficult from the environmental, economic and social points of view. There is a need to develop, master and implement simple but reliable models and application software that will help decision-makers in public administration (PA) to analyse waste manage-

ment processes, follow national and European Union (EU) legislation, the implementation of the Waste Management Plan (WMP) for the respective PA level [1] and to consider also EU Circular Economy tasks [2].

Therefore, we have developed an improved version of the information system, DECWASTE, forecasting the quantity of waste generated for each waste category listed in Section 2(1) of Annex I of Regulation (EC) No 849/2010 amending Regulation (EC) No 2150/2002 of the European Parliament and of the Council on waste (WSR). We have updated the previous information system, DECWASTE [3], forecasting the quantity of waste generated for each category of waste on the European list of waste (LoW) established by the Commission Decision 2000/532/EC.

The new DECWASTE can forecast these quantities for the prescribed years 2016-2025. It uses linked open data of the Czech eGovernment system, i.e. historical waste data (2009-2015) of the waste information system (ISOH) [4] of the Czech Environmental Information Agency (CENIA) at the national level of Czech Republic and transforms them into quantities of waste categories of the WSR. It downloads and parses linked open data to populate appropriate sets of indicators (predictors) for the forecasting models. The predictors are then used in the Driving Force-Pressure-State-Impact-Response (DPSIR) framework [5] which also integrates linked open data from the Czech eGovernment system [6] and from the Czech Statistical Office (CSO) [7].

DPSIR analysis was carried out by Delphi method [10] undertaken by the partnership on the basis of an assessment of the underlying data and searching for causal links of waste category production to the management of "driving force" (economic sectors of human activity), through the "pressures" (emissions, waste), on the "states" (physical, chemical and biological) and "impacts" on ecosystems, human health and function, that eventually lead to political "responses" (the setting of priorities of waste management, setting goals of the WMP and its indicators [1]), see [8-9].

The Delphi method was carried out with the assistance of a chosen expert group of finding solutions to the survey group experts' opinions. The group of experts carried out, independently of each other, predictors designated mediator were summarized, also distributed for the next round. Use of standardized questionnaires and the procedure was repeated until the approximate match.

Each group was assessed from the perspective of the individual parts of the DPSIR cycles. From the perspective of the production of waste category and the environment impacts we understood the individual parts of the DPSIR cycle, the following [5]:

- Driving forces are human activities or activities caused by the lifestyle of society. They lead to pressures on natural resources and the environment, which disrupt the ecological stability and impair the quality of the environment (e.g., emissions and waste).
- State is what is usually measured in the environment: waste category generation, the quality of water, soil, air and the environment (nature), energy and material flows.
- Pressures and states effect impacts: health problems, change in ecosystems, e invasion of alien species etc.
- Responses are responses of society to the identified problems in the form of specific measures (e.g. legislation).

The result of the implementation of DPSIR analysis was the description of the predictors of the individual parts of the DPSIR cycle, including identification of driving forces of the 51 waste categories of the WSR, expressed as the time series input data for calculating the forecast model [11]. The most relevant predictors were chosen through a selection process that included experts' opinions and literature reviews based on the relevance and applicability to different waste categories of WSR settings [9].

A verified multilinear regression model is presented. The construction of the forecasting model consisted in the construction and definition of the predictors based on the DPSIR framework; a sensitivity analysis of the models; implementation of the forecasting models. The presented results follow the authors' results from the iEMSS 2014 [8] and iEMSS 2016 [3] conferences.

2 Material and Methods

The development of the new information system DECWASTE includes the following consequent steps:

1. Identification of each waste category listed in Section 2(1) (51 items) and waste generating activity listed in Section 8(1) (19 items) of Annex I of the WSR and development of computation formulae for their amounts of using the Table of equivalence of Annex III of the WSR between the European Waste Classification for Statistics, version 4 (EWC-Stat Rev 4) (substance oriented waste statistical nomenclature) and the European List of Waste (LoW) established by the Commission Decision 2000/532/EC.
2. Processing of the historical annual waste generation and treatment reports (2009–2015) based on the LoW provided by waste generators and facilities in the Czech Republic to the ISOH and creating their data sets of waste categories of the WSR.
3. Identification and development of socioeconomic and demographic predictors for waste categories and activities based on the DPSIR framework (which have influence on the generation of waste categories) using linked open government data (eGovernment systems) of the Czech Republic.
4. Development of the multi-linear regression model of waste category generation with predictors from the DPSIR framework analyses.
5. Forecasting of predictors from the DPSIR framework and the calculation of waste category forecasts.
6. Processing of sensitivity analyses of predictors of waste category generation models.
7. Visualization scenarios forecasting the quantity of waste categories for the prescribed period 2016-2025.

We briefly describe the above steps.

Firstly, we implemented simple meta-computation formulae in DECWASTE using the Table of Equivalence of Annex III of the WSR between the EWC-Stat Rev 4 and the LoW, which enabled us to calculate the amount of every item of waste category of the WSR from items in the LoW.

Data sets of the ISOH keep the quantity of waste generation and treatment by generators and data from facilities to treat, recover and dispose of waste [3,8] from 2009-2015. Every year, the ISOH records more than 70,000 different generators' reports in all 6,500 municipalities of the Czech Republic and more than 3,000 facilities' reports. The annual ISOH database contains more than 50,000 records of municipal waste generation and 10,000 records concerning their treatment. The database is available to the public for the years 2009-2015 and we calculated all appropriate waste categories for these years.

The Czech eGovernment systems provide sources of linked open data of necessary input data for the DPSIR framework predictors for different waste categories and activities [5-8]. Parsing of the linked open data of eGovernment systems is divided in the DECWASTE information system into five phases [8]: definition of the appropriate data; data export; data processing; data import and optimization.

Identification and development of socioeconomic and demographic predictors based on the DPSIR framework [5] of the third step have been described [3, 9]. In the next section, we discuss the fourth step.

2.1 Development of a multi-linear regression model of waste generation based on DPSIR

We need to predict the amount of waste generated on the national level of the Czech Republic for the given year t and the chosen waste category $w^f(t)$ of the WSR $f = 1, \dots, 51$. We have developed mathematical forecast models based on the DPSIR framework [5] predictors according to the work of [8, 9]. They use available linked open data on DPSIR predictors from eGovernment systems [4] and data of past waste generation from the ISOH (2009 - 2015).

Let us assume that for a given waste category f , the amount of waste \hat{w}_t^f and predictors $\hat{A}_{i,t}^f$, $i=1, \dots, K^f$; in years $t=2009, \dots, 2015$ are known, where K^f is the number of predictors for the waste category f of the WSR. Let waste category generation $w^f(t)$ at the given year t fulfil the equation

$$\log(w^f(t)) = a_0^f + \sum_{i=1}^{K^f} a_i^f \cdot \log(A_i^f(t)) + \varepsilon_t^f, \quad (1)$$

where

- $A_i^f(t)$, $i=1, \dots, K^f$ are predictors in the given year t derived from the DPSIR analysis of the waste category f ,
- $\varepsilon_t^f = \log(w^f(t)) - \log(\hat{w}_t^f)$, for $t=2009, \dots, 2015$ are approximation errors.
- Coefficients $a_0^f, \dots, a_{K^f}^f$ in (1) for each waste category f are calculated using multiple regression on the basis of the values of waste generation \hat{w}_t^f and predictors $\hat{A}_{i,t}^f$, $i=1, \dots, K^f$; $t=2009, \dots, 2015$. Approximation errors ε_t^f , $t=2009, \dots, 2015$, have the mean equal to 0 and the normal distribution.

If we want to establish the confidence interval of predictors $A_i^f(t)$, it is necessary to restrict their number to $K^f \leq 5$, since we only have a time series of six past known values. If we have the values of \hat{w}_t^f and $\hat{A}_{i,t}^f$ for the next years $t=2016, \dots$, the model (1) will be more accurate and the approximation error $\varepsilon_t^{f,ps}$ will be smaller.

Furthermore, we assume that the predictors $A_i^f(t)$, $i=1, \dots, K^f$, for $t=2016, \dots, 2025$ have either known values (e.g. GDP, population, household consumption, etc.) from the eGovernment systems [6, 7] or are determined by an appropriate extrapolation method or are chosen by decision makers using DECWASTE.

These models are implemented in DECWASTE, written in language R and they use predefined predictors $\hat{A}_{i,t}^f$, which were parsed from linked open data [6, 7]. The outputs $w^f(t)$ are time series of the amount of waste generated for the years $t=2016, \dots, 2025$ of waste category f .

An estimate of the waste category generation $w^f(t)$ can be expressed after treatment (1) as

$$w^f(t) = A_0^f \cdot \prod_{i=1}^{K^f} A_i^f(t) a_i^f, \quad (2)$$

where $A_0^f = \exp(a_0^f)$, while we have neglected the error ε_t^f .

2.2 The sensitivity of estimate of waste generation with respect to the statistical significance of predictors.

The local sensitivity of the model (1), (2) for the i^{th} predictor $A_i^f(t)$, in the year t and waste category f can be estimated using the partial derivatives of the waste category $w^f(t)$ in (2) according to the i^{th} predictor $A_i^f(t)$:

$$a_i^f \cdot \frac{w^f(t)}{A_i^f(t)}. \quad (3)$$

It follows from (3) that if the value of predictor $A_i^f(t)$ is increased by 1 percent then the amount of waste $w^f(t)$ in waste category f will increase or decrease by a_i^f percent, if $a_i^f > 0$ or $a_i^f < 0$, for $i=1, \dots, K^f$. This knowledge is important for users of the model. We continue in the further analysis of the developed model, i.e. assessment of the statistical significance of predictors.

The predictors $A_i^f(t)$, $i=1, \dots, K^f$ in (1), (2) were determined for each waste category f on the basis of the analysis of the DPSIR framework and experts' assessment. Therefore, we will conduct an assessment of their statistical significance in the multiple linear model (1) depending on the input values of the predictors in the years 2009 to 2015. Firstly, we consider in the model (1) all predictors $A_j^f(t)$, $j=1, \dots, K^f$ in (1), in the waste category f where we calculate its sample variance, i.e.:

$$s^2 = (\sum_{t=2009}^{2015} (\varepsilon_t^f)^2) / (K^f - 1). \quad (4)$$

The statistical significance of each predictor $A_i^f(t)$ is calculated using the F-test of the difference of two sample variances, i.e. the sample variance s of the model (1) with all predictors and sample variance s_i of the model (1) where the i^{th} predictor $A_i^f(t)$ is excluded and the rest of the predictors, i.e. K^f-1 , are recalculated and its sample variance is calculated

$$s_i^2 = (\sum_{t=2009}^{2015} (\varepsilon_{t,i}^f)^2) / (K^f - 2), \quad (5)$$

Where $\varepsilon_{t,i}^f$ is the approximation error of multiple linear regression of the model (1), where the i^{th} predictor was excluded.

Then we set the number of degrees of freedom for both samples: $n_1 = K^f - 1$ (for s^2) and $n_2 = K^f - 2$ (for s_i^2) and compute the value of the test criteria (statistics) F_i :

$$F_i = \frac{\text{the greater of } (s^2, s_i^2)}{\text{the smaller of } (s^2, s_i^2)}, \quad (6)$$

that has the Fischer probability density distribution

$$H(x) = \frac{\Gamma(\frac{n_1+n_2}{2})}{\Gamma(\frac{n_1}{2})\Gamma(\frac{n_2}{2})} \cdot n_1^{\frac{n_1}{2}} \cdot n_2^{\frac{n_2}{2}} \cdot \frac{x^{\frac{n_1-2}{2}}}{(n_2+n_1 \cdot x)^{\frac{n_1+n_2}{2}}}, \quad (7)$$

for $x > 0$ and is equal to 0 for $x \leq 0$.

We use statistical software where it is more common to calculate the test p -value, which we denote as p_i . This is the smallest level of the F-test in which we would reject the hypothesis $H_0: \{s^2 = s_i^2\}$. We set this value as $p_i = 1 - H(F_i)$.

This procedure is repeated gradually for other predictors $A_i^f(t)$, for which we calculate F_i statistics and the test p -values $p_i, i = 2, \dots, K^f$.

Denote $V = \{F_i, p_i, s_i\}_{i=1}^{K^f}$ the set of predictor significance. We can now simplify the model (1) and exclude non-significant predictors.

Let us choose the level of significance α (values 0.05 or 0.1 are usually selected) of the predictors. We calculate p -values $p_i, i = 1, \dots, K^f$ and compare them with this level of significance α :

- If $p_i > \alpha \Rightarrow$ the null hypothesis $H_0: s^2 = s_i^2$ is rejected. Conclusion: the variances of different models are statistically significant and the i^{th} predictor $A_i^f(t)$ is significant.
- If $p_i < \alpha \Rightarrow$ we cannot reject the hypothesis H_0 . Conclusion: the variances of both models are not statistically significantly different (i.e., the selections originated from the same basic model with the common variance s^2) and the i^{th} predictor is not significant.

In this way we can simplify the model (1), (2) if we exclude insignificant predictors $A_i^f(t)$, which we originally selected on the basis of the DPSIR analysis. In the developed model, only the statistically significant predictors $A_i^f(t)$ then remain. For some waste category f , however, it may happen that the original proposed predictors based on the DPSIR analysis do not remain in any model (1). The resulting forecast p of waste category generation is then constant and independent of the predicted values of the predictors of the future. In this case we can only reduce the statistical significance level α , i.e. reliability of the developed model and accept a greater risk of erroneous predictions.

2.3 Extrapolation of the predictors

Let us suppose that for each waste category f the amount of waste category generation $\hat{w}_{i,t}^f$ and predictors $\hat{A}_{i,t}^f, i=1, \dots, K^f; t=2009, \dots, 2015$ are known and we have calculated for each waste category f the coefficients $a_0^f, \dots, a_{K^f}^f$ in the model (1) by using multiple regression.

For the calculation of the forecast waste category generation $w^f(t)$ in the years $t=2016, \dots, 2025$ it is necessary to know the values of predictors $A_i^f(t)$, $i=1, \dots, K^f$, $t=2016, \dots, 2025$. These values, however, may not always be listed in the sources (linked open data in the eGovernment systems) from which we draw the data predictors $\hat{A}_{i,t}^f$, $i=1, \dots, K^f$; $t=2009, \dots, 2015$. In this case, the procedure is as follows:

- Enter the values of the predictors based on experts' estimates or other appropriate sources;
- On the basis of the values of the predictors $\hat{A}_{i,t}^f$, $i=1, \dots, K^f$; $t=2009, \dots, 2015$ the values of predictors $A_i^f(t)$, $i=1, \dots, K^f$, $t=2016, \dots, 2025$ are calculated using either linear or exponential extrapolation.

2.4 Modelling measures in waste management and scenarios

The developed models (1), (2) of forecast waste generation $w^f(t)$ for the waste category f should also reflect the trends in changes in this waste category as a result of the anticipated N measures o_j , $j = 1, \dots, N$ (WMP, EU and national legislative changes etc.), which will have an impact on the given waste category f in the year t .

Therefore, we introduce the function $P^f(t, o_1, o_2, \dots, o_N)$ in the form:

$$P^f(t, o_1, o_2, \dots, o_N) = \prod_{j=1}^N \left(1 - \frac{P_j^f(t)}{100}\right), \quad (8)$$

where functions $P_j^f(t)$ are time-dependent functions of the impact of given measures (e.g. prevention, kind of collection, permitted treatment, delivery distance to waste facilities, etc.) on the waste category f in the given year t (generally in the whole waste management of the Czech Republic) and for simplicity, we assume that

$$0 \leq P_j^f(t) < 100, j = 1, \dots, N; t = 2016, \dots, 2025. \quad (9)$$

The functions $P_j^f(t)$ for the year t are set as a percentage which should be achieved in years $t = 2016, \dots, 2025$, as a result of the measures o_j (on the basis of the WMP, WPP or other strategic documents). For the period $t = 2009, \dots, 2015$, zero values of these functions are considered. In most of the waste categories f simple function values $P_j^f(t)$ for the given year $t = 2016, \dots, 2025$ will not be provided, therefore, they will be estimated on the basis of expert assessment. If it is not possible to set them realistically, we consider them to be equal to 0.

The forecast of the waste generation $w_{prognosis}^f(t)$ for the given waste category f let us consider now as the product of the function $P^f(t, o_1, o_2, \dots, o_N)$ and estimated waste category generation $w^f(t)$ of (2):

$$w_{prognosis}^f(t) = \prod_{j=1}^N \left(1 - \frac{P_j^f(t)}{100}\right) \cdot A_0^f \cdot \prod_{i=1}^{K^f} A_i^f(t)^{a_i^f} \quad (10)$$

To compute logarithms (10) we obtain the resulting equation of our developed model for the forecast of the waste category generation f in the year $t = 2016, \dots, 2025$:

$$\log(w_{prognosis}^f(t)) = a_0^f + \sum_{i=1}^{K^f} a_i^f \log(A_i^f(t)) + \sum_{j=1}^N \log\left(1 - \frac{P_j^f(t)}{100}\right). \quad (11)$$

The values of the waste category generation \hat{w}_t^f and the values of predictors $\hat{A}_{i,t}^f$ are known for each waste category f and $P^f(t, o_1, o_2, \dots, o_N) = 0$ in the years $t = 2009,$

...,2015 and the coefficients a_0^f, \dots, a_K^f from (1), (2) are calculated using the method of multiple regression. The forecast of the amount of waste category generation $w_{prognosis}^f(t)$ in the years $t = 2016, \dots, 2025$ depends on the values of the predictors $A_i^f(t)$ and functions $P_j^f(t)$, $i=1, \dots, K^f$; $j = 1, \dots, N$ for each waste category f .

Using the functions $P_j^f(t)$, the decision makers at the PA could simulate different scenarios for the anticipated impact of the individual measures o_j , $j = 1, \dots, N$ on the total waste generation in the given waste category f and year t .

They could choose, for example, the value of $P_j^f(t)$, which was achieved for the year 2025 and the values of the function $P_j^f(t)$ in the previous years from 2016 to 2024, until the year 2015, where they choose this equal to zero and examine what the impact of the forecast value on the waste category generation $w_{prognosis}^f(t)$ is.

3 Results and Discussion

The DECWASTE information system enables the outputs of the above forecasting models for 51 waste categories listed in Section 2(1) of Annex I of Regulation (EC) No 849/2010. We present this for forecasting the household waste category.

3.1 Household waste forecast generation analysis outputs

Firstly, we describe the process of development of DPSIR analysis and the choice of predictors for the household waste category [11].

The basic *driving force* behind the production of household waste is the size of the population, which is also significantly affected by urbanization of the population, when residents with higher incomes and stronger consumer behaviour live in cities and their surroundings. High population density together with higher purchasing power also limits their own waste management options (e.g. composting) and creates requirements for faster replacement of goods, which affects household consumption and composition of the consumer basket. The driving force for the production of municipal waste is also the age structure of the population and the amount of the unemployment benefit, when families with small children, students, pensioners and partially unemployed remain throughout the day near the residence where their activities produce waste.

The main *pressure* influencing the amount of household waste generation is environmental education and enlightenment in the prevention of and management of household waste.

The *state* of household waste generation is annually evaluated by indicators of the WMP [1], data from information systems of the MoE relating to the amount of emissions/imissions from energy recovery waste, the landfilling of waste, the status of surface water and groundwater etc.

The *impacts* of the generation and treatment of household waste on the environment have not yet manifested themselves to distort the landscape; mainly landfills, incinerators and industrial areas with waste treatment plants, the potential threat to

groundwater and surface water and air pollution, etc. Furthermore, collection and transport of household waste and pollution in the case of collection containers, etc.

The *response* to generated household waste is increasing the number of collection places for sorting waste and streamlining the system of collection and treatment of household waste. There is, along with improving municipal waste management systems, the education of the population in the context of environmental education. Together with education legislative measures are used, which are operated both by the municipality as the originator of household waste, so the whole system for the collection and processing of waste, i.e. throughout the life cycle of the household waste. Legislative instruments are accompanied by economic instruments, in the form of charges (waste disposal in a landfill), and the amount of the tax rates and subsidies directed towards the development of systems for the municipal waste management.

The expert group choose the most significant four predictors for household waste for which past time series are available or can be estimated [11]:

- Driving forces: population; number of retired; unemployed rate; household expenditure on food, footwear and clothing.
- Press: prevention percentage according to environmental awareness (sorting rate, influence of ecological education etc.) which decreases the amount of household waste.
- State: household waste generation.

The outputs of every forecasting model consist of four basic output steps, which are presented for the household waste category forecasting, in the following figures.

Step 1: Input data for household waste generation model

Initial year of the model: <small>(first year with modelled generation).</small>	<input type="text" value="2016"/>
Statistical significance: <small>(percentage of probability that error not occurs when evaluate predictors).</small>	<input type="text" value="95"/>
Previous HW generation: <small>(metric tons in years, separated by comma; for unknown value insert NA).</small>	<input type="text" value="3846367.209,3732837.508,3598752.269,3424688.198,3348157.697,3398917.2,2"/>
Previous households expenditures in food, footwear and clothing: <small>(billions of CZK, values separated by comma; for unknown value insert NA).</small>	<input type="text" value="495.01,502.25,527.19,545.84,561.8,589.77,620.82"/>
Previous population: <small>(thousands of persons, values separated by comma; for unknown value insert NA).</small>	<input type="text" value="10491,10517,10497,10509,10510,10525,10543"/>
Previous number of retired: <small>(thousands of persons, values separated by comma; for unknown value insert NA).</small>	<input type="text" value="2790,2881,2873,2866,2858,2863,2874"/>
Previous unemployment rate: <small>(percentagess, values separated by comma; for unknown value insert NA).</small>	<input type="text" value="7.12,7.4,6.77,7.37,8.17,7.46,6.24"/>
Previous waste generation prevention: <small>(percentagess, values separated by comma; for unknown value insert NA).</small>	<input type="text" value="0,0,0,0,0,0"/>

Fig. 1. Step 1. Input data for MSW generation. Source: authors.

1. In the first step, Fig. 1, basic data inputs over the time $t=2009, \dots, 2015$ for modelling are provided: previous household waste generation and time series of the known values of the above DPSIR predictors (implicit values are parsed from linked open data) and forecasting model outputs.
2. In the second step, Fig. 2, users can specify the values of the predictors (pre-filled values are available with hyperlinks to the linked open data sources) expected in the model (1), (2) or choose their possible linear/exponential extrapolation. Users

can also input expected prevention measures (three possible scenarios of household waste prevention are available).

Step 2: Expected development up to 2025

	2016:	2017:	2018:	2019:	2020:	2021:	2022:	2023:	2024:	2025:	
Households expenditures in food, footwear and clothing: [bil. CZK]	632.82	653.79	674.76	695.73	716.69	737.66	758.63	779.59	800.56	821.53	Source: https://vdb.czso.cz/vdbvo2/faces/index.jspx?_af=statistiky#katalog=30847
Population [thds. persons]	10491	10439	10388	10336	10284	10271	10257	10244	10231	10217	Source: https://www.czso.cz/csu/czso/rnhabitantstvo_hu
Retired [thds. persons]	2944	3011	3078	3148	3219	3291	3365	3441	3518	3597	Source: http://www.cszs.cz/cz/cz/cz/informace/informacni-materialy/statisticke-roceniky.htm
Unemployment [%]	5.19	6.08	5.89	5.69	5.5	5.3	5.11	4.91	4.71	4.52	Source: https://vdb.czso.cz/vdbvo2/faces/cs/index.jspx?_af=vystup-objekt&pvo=ZAM06&z=N&z=T&F=TABULKA&verze=-1&nahled=N&sp=N&titr=G-F_M-F_Z-F_R-F_P-S-_null_null_&katalog=30853&str=v95&c=v3_RP2014
Prevention [%]	2	4	5	7	8	10	11	12	14	15	

[Input values into model](#)

Fig. 2. Step 2. Scenario of forecasting future development in household waste generation in 2016-2025 with expected prevention measures. Source: authors.

- In the third step, we conducted the assessment of predictor' statistical significance $\alpha=0.05$ in the model (1) depending on the input values of the predictors in the years 2009 to 2015 and we excluded insignificant predictors: number of retired; unemployed rate; household expenditure on food, footwear and clothing. We obtained a linear model with the predictor amount of population. In Fig. 3, results are shown in the form of a table, mathematical expression of (1) and a time-plot, showing the development of future household waste generation and the effects of the prevention measures taken (the curves represent the prediction interval and shaded areas show the confidence intervals).
- In the fourth step, Fig. 4, a sensitivity analysis is presented, showing decision makers the estimated effect of the individual predictors from the model (1) and the quality of forecasting.

4 Conclusion

The construction of the DECWASTE waste information system is presented. It uses historical waste data of the ISOH [4] and linked open data [6, 7] of the chosen predictors [8, 9] based on the DPSIR framework [5, 9]. Integration of the appropriate predictors into the forecasting model is discussed. The forecasting model was implemented as open source software DECWASTE and was verified using appropriate household waste data, whose outputs are presented.

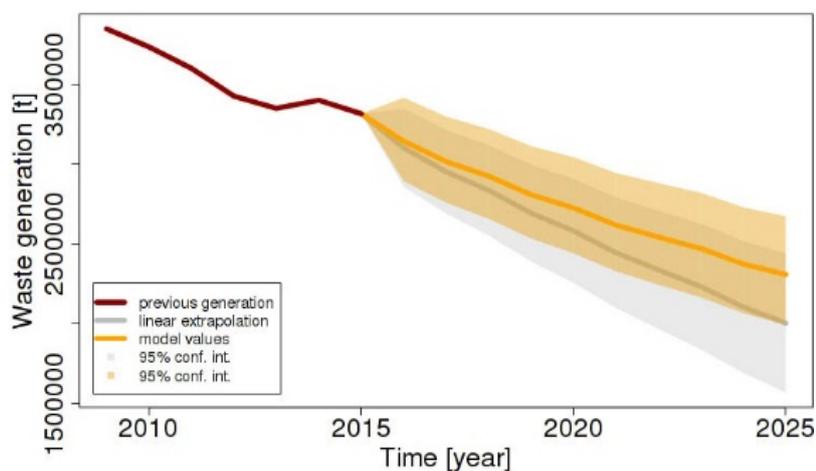
The new DECWASTE information system allows decision makers at the national level of the Czech Republic to make sustainable environmental decisions, focusing on waste management data requirements, national strategies for waste data acquisition, management and processing in a similar way as was done previously. They assisted in identifying alternative waste management strategies of the Czech Republic and support the national WMP that meet the objectives of EU legislation and Circular Economy principles [2].

Step 3: Prognosis and model equation

Show plot legend: Show confidence interval:

- nowhere yes
- bottom left no
- bottom right
- upper left
- upper right

[Refresh plot](#)



Extrapolation (time): $\text{generation} = (184202618.9 + -89800.8 \times \text{year}) \times (1 - \text{prevention})$
 Model: $\log(\text{generation}) = (19.102 + -0.6389 \times \log(\text{households})) \times (1 - \text{prevention})$

[Download table](#)

Extrapolation in time: $\text{generation} = (184202618.9 + -89800.8 \times \text{year}) \times (1 - \text{prevention})$
 Model generation: $\log(\text{generation}) = (19.102 + -0.6389 \times \log(\text{households})) \times (1 - \text{prevention})$

Table of model values:

	2016	2017	2018	2019	2020	2021	2022	2023	2024	2025
Model generation	3142702	3015105	2924115	2807124	2724773	2616872	2541862	2469916	2373191	2307163
Lower border	2892087	2757503	2656426	2532306	2440366	2326659	2243432	2164017	2064174	1992307
Upper border	3415033	3296770	3218778	3111766	3042326	2943285	2879990	2819056	2728469	2671778
Linear extrapolation	3100954	2951460	2835405	2692198	2580633	2443711	2336636	2231357	2103416	2002627
Lower border	2855472	2689412	2552168	2389578	2254724	2097809	1966937	1837807	1690913	1566925
Upper border	3346435	3213508	3118642	2994818	2906541	2789614	2706336	2624907	2515918	2438328

[Back to waste flow selection](#)

Fig. 3. Step 3. Final prediction of household waste generation. Source: authors.

Table of model results (in log form):

Predictor	Degrees of freedom	Sum of squared residuals	F statistics	p value
Variable: households	1	0.017	40.3	0.024
Variable: inhabitants	1	0	0.9	0.437
Variable: retired	1	0.001	2.6	0.25
Variable: unemployment	1	0.001	2.1	0.286
Residuals	2	0.001	NA	NA

Fig. 4. Step 4. Sensitivity analysis of forecasting household waste generation. Source: authors.

References

1. Preparing a Waste Management Plan. A methodological guidance note. http://ec.europa.eu/environment/waste/plans/pdf/2012_guidance_note.pdf
2. Closing the loop - An EU action plan for the Circular Economy. COM/2015/0614 final. <http://eur-lex.europa.eu/legal-content/EN/TXT/?qid=1453384154337&uri=CELEX:52015DC0614>
3. Hřebíček, J., Kalina, J., Soukopová, J., Valta, J., Prášek, J. (2016) Decision support system for waste management. In: Sauvage, S., Sánchez-Pérez, J.M., Rizzoli, A.E. (eds.) Proceedings of the 8th International Congress on Environmental Modelling and Software. Toulouse, France 10-14 July
4. Waste Management Information System. <http://www1.cenia.cz/www/odpady/iso>
5. Liao, M., Chen, P., Ma, H., Nakamura, S. (2015) Identification of the driving force of waste generation using a high-resolution waste input-output table. *Journal of Cleaner Production* 94, 294-303
6. Soukopová, J., Hřebíček, J., Valta, J. (2015) National Environmental Data Facilities and Services of the Czech Republic and Their Use in Environmental Economics. In: Denzer, R., Argent, R.M., Schimak, G., Hřebíček, J. (Eds.) *Environmental Software Systems: Infrastructures, Services and Applications*, 11th IFIP WG 5.11 International Symposium, ISESS 2015, Melbourne, Australia, Heidelberg: Springer, pp. 361-370
7. Czech Statistical Office. <https://www.czso.cz/csu/czso/home>
8. Kalina, J., Hřebíček, J., Bulková, G. (2014) Case study: Prognostic model of Czech municipal waste production and treatment. In: Ames, D.P., Quinn, N.W.T., Rizzoli, A.E. (eds.) *Proceedings of the 7th International Congress on Environmental Modelling and Software*. San Diego, 15-19 June. pp 227-332.
9. Hřebíček, J., Kalina, J., Soukopová, J., Prášek, J., Valta, J. (2016) The Forecasting Waste Generation Model based on Linked Open Data and the DPSIR Framework. Case study concerning municipal waste in the Czech Republic. In *Proceedings of the Cyprus 2016. 4th International Conference on Sustainable Solid Waste Management*. Limassol, 23–25 June.

10. Seker, S.E. (2015) Computerized Argument Delphi Technique. *IEEE Access*. 3, 368–380. doi:10.1109/ACCESS.2015.2424703
11. Horáková, E., Kalina, J., Hřebíček, J., Prášek, J., Soukopová, J., Buda Šepeřová, G., Valta, J. (2015) Vyhodnocení DPSIR rámce pro vybrané skupiny odpadů stanovené výzvou TB940MZP003. (Evaluation of the DPSIR framework for selected groups of waste laid down a challenge TB940MZP003). Internal report. CENIA, Praha