Waste Management 34 (2014) 2537–2547

Contents lists available at ScienceDirect

Waste Management

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

Using multivariate regression modeling for sampling and predicting chemical characteristics of mixed waste in old landfills

Christian Brandstätter [a,*], David Laner [a], Roman Prantl [b], Johann Fellner [a]

[a] Institute for Water Quality, Resource and Waste Management, Vienna University of Technology, Karlsplatz 13/226-2, 1040 Vienna, Austria
[b] blp GeoServices gmbh, Felberstrasse 24/1, 1150 Vienna, Austria

ARTICLE INFO

Article history:
Received 25 April 2014
Accepted 12 August 2014
Available online 11 September 2014

Keywords:
Bootstrapping
Landfill emission
Multivariate linear model
MSW landfill
Waste characterization

ABSTRACT

Municipal solid waste landfills pose a threat on environment and human health, especially old landfills which lack facilities for collection and treatment of landfill gas and leachate. Consequently, missing information about emission flows prevent site-specific environmental risk assessments. To overcome this gap, the combination of waste sampling and analysis with statistical modeling is one option for estimating present and future emission potentials. Optimizing the tradeoff between investigation costs and reliable results requires knowledge about both: the number of samples to be taken and variables to be analyzed.

This article aims to identify the optimized number of waste samples and variables in order to predict a larger set of variables. Therefore, we introduce a multivariate linear regression model and tested the applicability by usage of two case studies. Landfill A was used to set up and calibrate the model based on 50 waste samples and twelve variables. The calibrated model was applied to Landfill B including 36 waste samples and twelve variables with four predictor variables.

The case study results are twofold: first, the reliable and accurate prediction of the twelve variables can be achieved with the knowledge of four predictor variables (Loi, EC, pH and Cl). For the second Landfill B, only ten full measurements would be needed for a reliable prediction of most response variables. The four predictor variables would exhibit comparably low analytical costs in comparison to the full set of measurements. This cost reduction could be used to increase the number of samples yielding an improved understanding of the spatial waste heterogeneity in landfills.

Concluding, the future application of the developed model potentially improves the reliability of predicted emission potentials. The model could become a standard screening tool for old landfills if its applicability and reliability would be tested in additional case studies.

1. Introduction

Municipal solid waste (MSW) landfills pose a long-term threat to human health and environment via leachate and gaseous emissions. The source of these emissions, the solid body of landfills is typically not very well examined. This is especially the case for older MSW-landfills, where no documentation about the deposited material is available. For assessing the long-term emission potential, the highly heterogeneous solid body composition is of major interest. The emissions eventually derive from the solid body, which particularly applies for gas production (Barlaz et al., 1990).

The lack of information concerning the landfilled waste is related to high sampling and chemical analytical costs. Furthermore, the extracted information of such solid waste sampling campaigns is limited, since the heterogeneity of the landfill body typically overshadows the results (e.g. Östman et al., 2006). In comparison to solid waste sampling, leachate sampling campaigns are more common, mainly because leachate is easier to access and also directly related to groundwater pollution (Eggen et al., 2010). However, leachate also strongly varies in its composition within a landfill spatiotemporally (Kjeldsen et al., 2002) and at most older landfill sites leachate cannot be collected as they are not equipped with a base lining system.

Abbreviations: BOD_5, biological oxygen demand after 5 days; COD, chemical oxygen demand; DW, dry weight; EC, electrical conductivity; ICP-AES, inductively coupled plasma atomic emission spectroscopy; Loi, loss on ignition; MSW, municipal solid waste; RI_4, respiration index after 4 days; RMdSPE, root median square percentage error; SD, standard deviation; sqrt, square root; TN, total nitrogen; TOC, total organic carbon; WC, water content.

* Corresponding author. Tel.: +43 1 58801 22658.
E-mail addresses: bran.chri@gmail.com (C. Brandstätter), david.laner@tuwien.ac.at (D. Laner), r.prantl@blpgeo.at (R. Prantl), johann.fellner@tuwien.ac.at (J. Fellner).

http://dx.doi.org/10.1016/j.wasman.2014.08.009
0956-053X/

Typical sampling campaigns of MSW-landfills, as rarely as they are conducted, lead to 20–40 valid samples (Mor et al., 2006; Sormunen et al., 2008), where a wide range of analytical procedures in view of emission potential is applied. According to the Austrian standard ON S 2087, a guideline for the investigation of contaminated sites such as old landfills, in standard situations eight solid variables should be analyzed for as well as twelve variables after elution. Measuring twenty variables would lead to high analytical costs per sample. According to this guideline, it is recommended to take samples every 20 × 20 m. In view of landfill heterogeneity, we propose to take a larger number of samples and measuring less variables. This would ideally lead to an improved understanding of the emission source and enable site administrators to better identify for example hot spots of organic matter.

The aim of this article hence is to identify the optimized number of waste samples and chemical variables in order to predict a larger set of chemical variables. To achieve this goal, we applied a linear multivariate model based on four selected predictor variables with low analytical costs (Loi (loss on ignition), pH, EC (electrical conductivity) and Cl (Chloride)). By that we obtained predictions about twelve response variables for two fairly differing landfills.

2. Material and methods

In this work, we present a multivariate linear model. The model was formulated in the view of low analytical costs and with data from Landfill A. To show that the model can be applied to other MSW-landfills as well, the resulting model coefficients were applied on data from a second Landfill B. Afterwards, the model outcome was adjusted as will be explained below. For the sake of clarity we provide a short glossary with the most crucial statistical terms in the context of this article:

Variable: Method of measurement (e.g. Loi, TOC, WC); column in a typical dataset.
Observation: The measured values for every variable of one sample; row in a typical dataset.
Predictor variable: Variables used for prediction (here: Loi, pH, EC and Cl).
Response variable: The predicted variable (here by applying multilinear modeling).
Model adjustment criterion: Maximum 10% RMdSPE deviation from adjustment based on all measured variables.
Model evaluation criterion: Maximum 35% RMdSPE deviation predicted vs. measured.

2.1. Site characteristics and sampling

2.1.1. Landfill A

Landfill A contains approximately 220,000 m^3 (fresh matter) of waste with an average deposition height of 3.7 m. According to test pitting conducted the landfilled waste mainly consists of MSW (66% moist mass), excavated soil (18% moist mass) and construction and demolition waste (16% moist mass). Landfilling took place at the site from 1965 to 1974 and resulted in a total landfill volume of approximately 240,000 m^3 including top soil (Brandstätter et al., 2013).

2.1.2. Landfill B

At Landfill B about 210,000 m^3 (fresh matter) of untreated municipal and commercial solid waste have been deposited between 1976 and 1995 (Prantl, 2007; Prantl et al., 2006a,b). The site is characterized by an average waste depth of 7.7 m and is divided into two sections differing in reactivity and age. The waste samples analyzed for this article originate from both sections.

2.1.3. Sampling

Sampling and chemical analyses of the deposited waste at Landfill A were conducted by the authors of the present study, while information about Landfill B was obtained from the literature (Prantl, 2007; Prantl et al., 2006a,b). At both landfills the excavated material was sieved with a mesh width of 20 mm. For the sampling campaign in Landfill A an excavator and for Landfill B a grab-excavator with 600 mm diameter was used. The average sample size in both cases was ~20 kg. At Landfill A in total 56 samples from 17 pits were taken and at Landfill B 54 samples from 20 excavated holes. For the sampling campaign in Landfill A, a priori information about the distribution of organic matter in the landfill was available (unpublished study), performed after the evaluation of a risk assessment study from the Austrian environmental agency (Environmental Agency Austria, 2005). Based on this information, the herein described sampling campaign at Landfill A focused on zones rather rich in organics.

2.2. Chemical analyses

This section provides additional information to Table 1 which contains the applied analytical devices and/or the standardized methods.

2.2.1. Solids

For chemical extraction of Cr_s, Cu_s, Pb_s and Zn_s in Landfill A, a microwave oven (Start 1500, MLS GmbH, Leutkirch Germany) was used (200 °C). The applied solvent was aqua regia (HCl and HNO_3 in the volumetric ratio of 3:1). The analysis of the elemental content was performed with inductively coupled plasma atomic emission spectroscopy (ICP-AES Ultima 2, Horiba Jobin Yvon, Munich, Germany). The RI_4 (respiration index after 4 days) of solid waste samples taken from Landfill A was analyzed by Agrolab Austria GmbH and for samples of Landfill B it was determined according to Binner and Zach (1999).

2.2.2. Eluate

For eluting the fresh waste material a water/solid ratio of 10 l/kg was applied for both landfills. The material was eluted in an orbital shaker at 250 rpm for 24 h at room temperature. For Landfill A, the EC and pH were measured with a pH meter (Seven Excellence S470 kit, Mettler Toledo, Ohio, USA) and the variables Cl and SO_4 were analyzed using ionic chromatography (IC Dionex ICS 900, ThermoFisher Scientific Inc., Massachusetts, USA).

2.3. Statistical analyses

2.3.1. Data selection with plausibility tests

All statistical analyses were performed with the program R (ver. 3.0.1, R Core Team, 2013). For selecting plausible data from the full datasets we applied two reproducible criteria on data of each of the landfills: the first criterion was that the ratio of Loi and TOC_s (total organic carbon; total carbon – total inorganic carbon) should be lower than 2.4. The second criterion was that the ratio of TOC_s and TN should be higher than 8.

The reasoning behind the first criterion was that when investigating the ratio of Loi and TOC_s for organic substances present in landfills (such as glucose, cellulose, lignin, fats, protein or different plastic polymers), this ratio can theoretically vary between 2.5 for glucose (cellulose, hemicellulose and lignin show a ratio of 2.25, 2.31 and 1.5, respectively) and 1.08 for polystyrol, while polyethylene and polypropylene are characterized by a ratio of 1.17 (Kost, 2001)). A table indicating TOC:TN values of different polymers is given in supplementary material (Appendix 1). Taking possible measurement errors into account as well as the fact that there will be always a mixture of different organic substances in landfills