Waste production classification and analysis: a PCA-induced methodology

Christelle Hatik, Jean-Claude Gatina

Available online at www.sciencedirect.com

ScienceDirect

Energy Procedia 136 (2017) 488–494

Energy Procedia

www.elsevier.com/locate/procedia

4th International Conference on Energy and Environment Research, ICEER 2017, 17-20 July 2017, Porto, Portugal

Waste production classification and analysis: a PCA-induced methodology

Christelle Hatik[a,*], Jean-Claude Gatina[a]

[a]*Laboratory PIMENT of the University of Reunion Island, 117 rue Général Ailleret, 97430 Le Tampon, France*

Abstract

Knowledge of waste composition and production is a requirement to build an efficient waste management scenario. Analysis of this data at a detailed level of observation (regional or communal) is useful to create adapted local scenarios, thus optimizing the overall waste management. However, working at a detailed level of observation multiplies the number of scenarios to build. In this article, we use Principal Components Analysis (PCA) to identify similarities between local administrative areas. By grouping administrative areas based on their waste production, this analysis is an efficient way to reduce the number of local waste management scenarios to define and it also favors cooperation among similar administrative areas. To illustrate our methodology, we focus on the specific case of Reunion Island, which is composed of 24 municipalities. The PCA analysis enabled to identify 5 groups of municipalities, thus reducing the number of required scenarios to build drastically.

Peer-review under responsibility of the scientific committee of the 4th International Conference on Energy and Environment Research.

Keywords: Decision-making; Principal Component Analysis; Town's typology and territories analysis; Waste management scenarios

1. Introduction

Waste management is often defined at a territorial level. The overall waste production is determined and a waste management scenario is built accordingly, covering the entire territory. Although this approach is interesting to size and deploy massive treatment plants, it lacks adaptation to sub-territorial specific waste flux characterization and treatment. In fact, when waste management is studied at small-scale like municipalities, towns or districts, waste production can be analyzed in depth in order to capture specificities and optimize its management. In practice, organizing waste management at a sub-territorial level is possible but multiplies the number of scenarios to build

* Corresponding author. Tel.: +262-262-579245
E-mail address: christelle.hatik@univ-reunion.fr

1876-6102 © 2017 The Authors. Published by Elsevier Ltd.
Peer-review under responsibility of the scientific committee of the 4th International Conference on Energy and Environment Research.
10.1016/j.egypro.2017.10.308

before even analyzing their correlation at the territorial level.

In this article, we propose a methodology to optimize waste management at small-scale by identifying correlation between small-scale areas' waste production, thus enabling mutualisation and co-creation of local scenarios. This approach is effective to leverage the number of scenarios to be built and favor collaborative decision-making between all waste management actors. Based on waste flux analysis, this approach is useful to better anticipate waste and urban development, thus improving forecasting decision-making [1]. PCA is used as the main analysis tool. It enables to highlight similarities between local administrative areas. So, this methodology can be used as follow:

- To compare behaviors into different administrative area (regions, municipalities, towns, etc.);
- To extract some similarities between studied areas;
- And to classify areas with the same profile.

We will present the general two-steps methodology in a first part. Then we will use the particular case of Reunion Island to apply the methodology and analyze its results. Lastly, we will discuss about the advantages of this methodology before concluding.

2. Methodology

Our approach is directly based on the quantitative and qualitative evaluation of wastes at the desired level of observation. Therefore, the first step is to choose waste categories (municipal solid wastes, organic wastes, recyclable wastes, e-wastes, etc.) and to determine the associated waste production for each studied area. At this step, we build a database with the classified wastes streams and their quantities. At the second step, we use Principal Component Analysis as our main analysis tool. With this method, we want to determine: i) if there is any correlation or similar behaviors in terms of composition and tonnages between two or more areas; ii) if there are indeed some similarities, which waste management scenarios can be established and for which towns.

2.1. Methodology to quantify the estimation of the waste generation

To quantify waste production, several methods can be used depending on the considered waste stream. As stated above, before quantifying waste production it is important to clearly identify the appropriate waste streams to be integrated into the analysis. Once the waste streams identified, we have to adapt the quantifying method for each stream. From our experience, we can distinguish two ways: either we use existing data that is made available by some institutions (municipalities for example), either we calculate from waste production ratio. To illustrate the later, a general example is given to quantify biowastes.

Annual biowaste estimation (Q_{iy}) generated by a type of activity (i) over one year (y), is obtained by multiplying either the number of meals supplied (Nb_R), the number of merchants, (Nb_{Ex}), the workforce (ES) or the area S for a year by biowaste production ratios (R_{DO}) produced per meal or production ratios of DIB (R_{DIB}) and then by biowaste percentage ($\%_{DO}$). The sum of biowaste generation for each i activity for the y year is equal to total biowaste generation (Q_{ty}). And lastly, total biowaste generation (Q_{iy}) is multiplied by the organic matter percentage ($\%_{OM}$) specific to the wastes streams considered to obtain total biowaste generation into tons per organic matter Q_{tyOM}.

The equations are:

$$Q_{ty}\ \textit{(tonne/an)} \geq \sum_{i=1}^{n} Q_{iy}\ \ \textit{(tonne/an)}$$

$$Q_{iy}\textit{(t/an)} = (Nb_{Ex}\ ;\ Nb_R\ ;\ ES\ ;\ S) \times R_{DO} \tag{1}$$

$$Q_{iy}\textit{(t/an)} = (Nb_{Ex}\ ;\ Nb_R\ ;\ ES\ ;\ S) \times R_{DIB} \times \%_{DO}$$

$$Q_{tyOM} = Q_{ty} \times \%_{OM}$$

At this step, we created a database containing tonnage for each waste stream and each studied local administrative area. The dataset is usually composed of many variables that make the analysis difficult with classic graphs (diagram, histogram, cloud point, etc.). In this case the PCA is a very efficient method.

2.2. Principal Component Analysis and database analysis

PCA is a factorial dimension reduction method for statistical analysis of complex quantitative data ([2], [3]). The PCA provides an optimal two-dimensional or three-dimensional graph on which all the administrative areas are classified. The axes are based on the waste characteristics that are the most discriminant in the dataset. PCA is used to reduce the dimensions of dataset to usually two dimensions while remaining close to reality (initial dataset). PCA permits to visualize graphically the database and to have a global view. So, two administrative areas close the one from another in the graph have similar waste production. Applying PCA to our data implies representing the dataset as a matrix of real quantitative variables X^j observed on n individuals i and assigned weight w_i. These data are classified in a matrix of dimensions $(n \times p)$ where each variable $X_j = (X_{1,p}, \dots, X_{n,p})$ has mean $\bar{x}_p$ and standard deviation σ_{Xj}.

$$X = \begin{bmatrix} x_1^1 & x_1^j & x_1^p \\ x_i^1 & x_i^j & x_i^p \\ x_n^1 & x_n^j & x_n^p \end{bmatrix} \tag{2}$$

The aim of PCA is to propose a projection of each variables on different axes. The projection axes consist in that the first components represent the most information level of the origin variables. In general, the first two axes are sufficient to represent the set of original data.

$$\begin{aligned} C^1 &= a_1^1 X^1 + a_1^2 X^2 + \cdots + a_1^p X^p \\ C^2 &= a_2^1 X^1 + a_2^2 X^2 + \cdots + a_2^p X^p \\ &\dots \end{aligned} \tag{3}$$

It is important to note that C^1 and C^2 are non-correlated linear functions of the origin variables so that the information carried by C^2 is completely new and complementary to that of C^1 ([4]). For a two-dimensional analysis, C^1 and C^2 are the axes on which our dataset will be represented: these axes bring out correlations between areas but also between wastes streams. Now that we have presented the general methodology, and why this approach is important for an integrated waste management, we focus on an application example.

3. Case study on Reunion Island

Our methodology has been applied to Reunion Island and considering a specific waste stream: Biowastes. Reunion Island is a French oversea territory, located in the Indian Ocean, between Mauritius and Madagascar. Its population reached 833 000 inhabitants in 2010. RI is structured with 24 municipalities. With landfills having reached their maximal capacity and an ever-growing population, the waste management scenario must be revised. An optimal global scenario is particularly difficult to build due to the specificities of waste composition and production of each municipality.

As we presented previously, we will follow the two-steps process of our methodology. We will first focus on the database creation, with all waste generation data, and then apply the PCA on the produced dataset.

In this work, the estimation of biowaste generation in Reunion Island is estimated at 206 336 tOM/y. The results show that some wastes streams are predominant, like farming waste with about 101 331 tOM/y (Fig. 1). Therefore, it is logical that the towns producing the most waste should be the towns with the most farms like Saint-Joseph, Le Tampon, Saint-Pierre and Saint-Paul (Fig. 2).

3.1. Principal Components Analysis results

The database is composed of 24 entries that represent towns and 15 variables that represent wastes streams. Our goal is to determine similar profiles among the towns. In our example, the PCA is able to represent 84% of the data with only two axes, which confirm that two axes can already be sufficient to correctly analyze the database.

To interpret PCA results, we analyze firstly all graphics results and secondarily numerical results simply because graphics are easier to interpret. The numerical results permit to confirm conclusions made from graphic observation.

We analyze simultaneously results of each variable (wastes stream) and entrie (towns) with the aim of the characterization of different typologies of towns out of wastes streams.

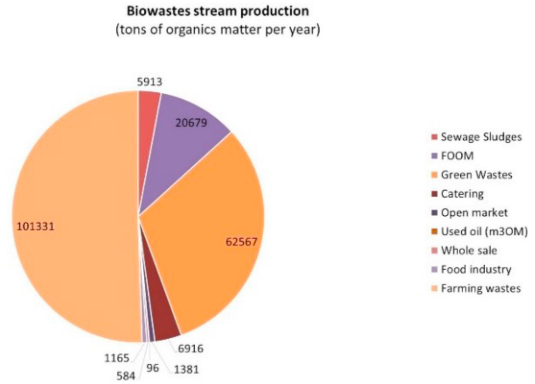

Fig. 1. Estimation of biowastes production on Reunion Island

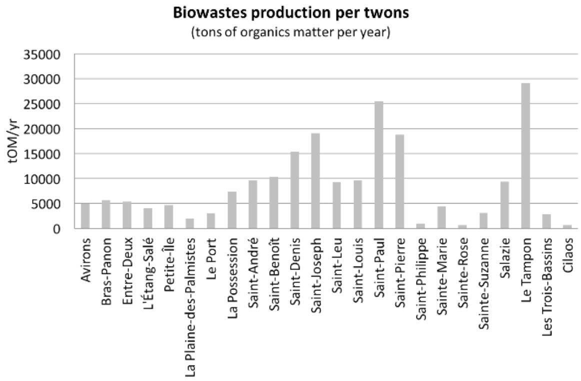

Fig. 2. Biowaste production per towns on Reunion Island

3.2. Exploiting the waste streams graphic representation

To interpret PCA graphics, it is important to identify the axes in function of the towns and wastes streams location (Fig. 3). We suppose by graphic observations that axis F2 represents represent variables (wastes streams). So, we can see that food market waste (FMW) and FOOM are positively correlated. This means that more GMS waste are product by a town, more FOOM wastes will produce, and conversely. And more FMW and FOOM wastes is produced, less institutional catering wastes, fast food waste and open market wastes will produce. The same behavior is observed for wholesale stores wastes and retail stores wastes. And farming sludge are principally represented the axis F1 and are positively correlated (quasi align on axis 1). So, we can conclude that the axis F2 is representing anthropic activities and the axe F1 farming activities.

3.3. Exploiting the towns graphic representation

To interpret the towns graphic location, we used the same approach that for waste streams location (Fig. 4). Previously, we suppose by graphic observations that axis F1 represents wastes streams and that axis F 2 represents waste quantities. So, with this hypothesis, we can say that Saint-Denis, Saint-Pierre and Saint-Paul on the right of the graph produce high quantity of waste (location at the right of axis 2). Saint-Pierre and Saint-Denis seem to produce all types of biowaste (location in the middle of axis F1, close to axis F2). Saint-Paul, Le Tampon and Le Port seem to have a more specific waste production (localization at the ends of axis F1). Moreover, we suppose that Le Tampon and Saint-Paul have a waste production totally different from Le Port (location at the opposite). We can conclude that the top of axis F1 correlates with the type of waste production related to farming with Le Tampon and Saint Paul producing mainly this kind of waste; whether the bottom of axis F1 correlates with industrial-related waste production with Le Port. For the other towns, we can observe two groups, one group in the middle of axes 1 and 2 and another one on the left of the graph. The group with Saint-Benoît, Saint-Louis, Saint-André, Sainte-Marie, Saint-Leu, Cilaos, Bras-Panon and La Possession, seems to be balanced. It means that their production of wastes is globally the same for each stream (location on axis 2). In addition, the second group, the less productive one (left on axis 2) but equally balanced in their wastes production (middle of axis 1). These two groups of towns show that their waste productions are similar in quantity and in quality. We can say that towns of a same group can be managed by a same scheme of waste management.

To conclude, we can say that axis F1 organizes towns according to the composition of their waste (farming character to industrial character). The PCA analysis enabled to identify 5 groups of municipalities, thus reducing the number of required scenarios to build drastically:

- One group with Saint-Denis, Saint-Paul et Saint-Pierre,
- One group with Le Tampon and Saint-Joseph,
- One group with Le Port,
- One group with Salazie, Petite-Île, Les Avirons, Saint-Philippe, Sainte-Rose, Les trois Bassins, La Plaine des Palmistes and Sainte-Suzanne,
- One group with Saint-Benoît, Cilaos, Bras-Panon, Saint-Louis, L'Etang-Salé, La Possession, Sainte-Marie and Saint-André.

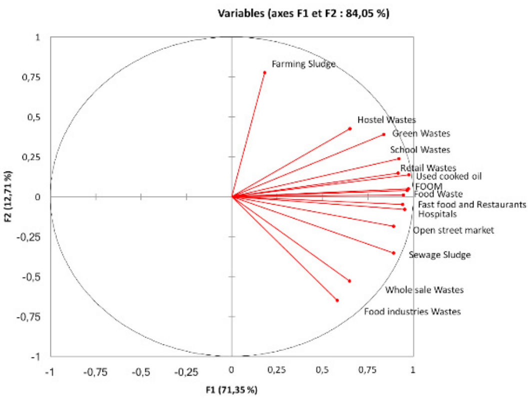

Fig. 3. PCA analysis result for the 24 municipalities of Reunion Island

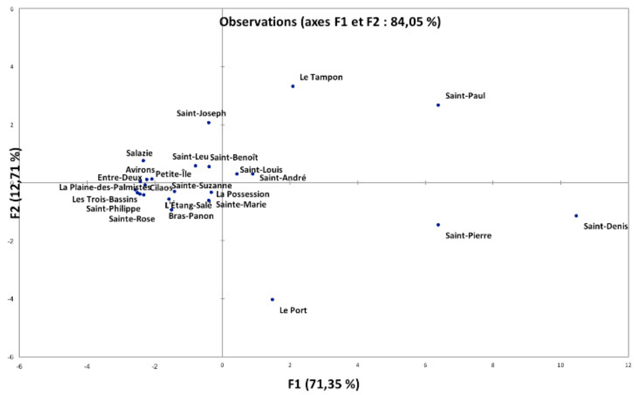

Fig. 4. Graphical analysis of axis composition

In this approach, the overall waste management scenario emerges from all the local areas waste management systems that are each adapted in terms of volumes and waste streams. Likewise, if the results show that areas with the same typology are geographically close, the waste management structures could be put in common by towns.

These conclusions can be reached by the analysis of numerical results that show coordinates of towns and variables (waste stream) on the axes F1 and F2, to explain their contribution to building axes (Table 1).

We can see that the first axis is correlated to FOOM (0,972), to green wastes (0,836), to sewage sludge (0,890), to fast food (0,942), to traditional restauration (0,945), to hospital catering (0,952), to school catering (0,920), to open market food wastes (0,891), to used cooking oil (0,976) and to wholesale (0,916).

While, the second axe is principally described by farming sludges (0,775) and organic industrial wastes with a trust level of 95%.

All these numerical observations confirm the graphical analysis. The PCA analysis enabled to identify 5 groups of municipalities, thus reducing the number of required scenarios to build drastically. Each town in a group produces roughly the same quantity and quality of waste as the others.

To summarize, we demonstrated that towns activity (urban, farming or industrial characteristics) can be qualified out of their wastes streams with the example of Le Port and IAA stream or Le Tampon and farming waste. This method can be applied for all type of wastes, all type of waste treatment and all territories.

Table 1. Numerical data of coordinates for towns (on left) and wastes stream (on right)

Observation	Coordinates	
	F1	F2
Avirons	-2,243	0,117
Bras-Panon	-1,506	-0,936
Entre-Deux	-2,433	0,040
L'Étang-Salé	-1,578	-0,569
Petite-Île	-2,084	0,127
La Plaine-des-Palmistes	-2,578	-0,243
Le Port	1,479	-4,030
La Possession	-0,332	-0,330
Saint-André	0,889	0,298
Saint-Benoît	-0,393	0,550
Saint-Denis	10,464	-1,138
Saint-Joseph	-0,408	2,063
Saint-Leu	-0,796	0,584
Saint-Louis	0,424	0,304
Saint-Paul	6,383	2,674
Saint-Pierre	6,380	-1,450
Saint-Philippe	-2,519	-0,349
Sainte-Marie	-0,403	-0,604
Sainte-Rose	-2,444	-0,389
Sainte-Suzanne	-1,417	-0,296
Salazie	-2,341	0,765
Le Tampon	2,080	3,314
Les Trois-Bassins	-2,332	-0,425
Cilaos	-2,293	-0,077

Variables	Coefficient	
	F1	F2
Farming sludge	0,181	0,775
FOOM	0,972	0,049
Green wastes	0,836	0,389
Sewage sludges	0,890	-0,353
Fast food	0,942	-0,050
Restaurants	0,945	0,011
Hostel wastes	0,651	0,426
Hospital wastes	0,952	-0,081
School wastes	0,920	0,238
Open market	0,891	-0,186
Used oil	0,976	0,136
Food wastes	0,966	0,039
Whole sale	0,649	-0,530
Retail wastes	0,916	0,146
Food industries	0,581	-0,651

4. Discussion

In practice, organize waste management at small-scale for each region is possible but requires to provide and to manage waste scenarios on long-term. It is a hard work of anticipation. So generally, it is more simply to establish one scenario for all regions on the same country. Moreover, dataset is generally consequently with many numerical data not easy to analyze ([5], [6], [7], [8]). In this article, we have use PCA method to analyze waste production in each towns of a country with this aim of towns to identify similarities. We demonstrate that it is possible to analyse a dataset of wastes production for many towns with greater ease that the habitual analytical techniques like classic plots.

5. Conclusion

Through the example of Reunion Island, we demonstrate effectiveness of PCA analysis to identify waste composition and production patterns among different administrative areas. One of the assets of PCA is the fact that it features graphical representation of the dataset, enabling anyone to visualize correlation existing between the different administrative areas. Principal Components Analysis of wastes production can then be used for different purposes: waste management scenarios pooling, cooperation among administrative authorities, location and optimization of waste treatment plants, and more generally waste management decision-making. This study will be particularly useful toward making towns of a same group working together to mutualize solutions. We think is important to propose more specific waste management scenario adapted to the territory. For this, we have developed a methodology using PCA to identify some typologies of local administrative areas in a territory. This simple method can be used for all type of wastes and countries in function of the objectives of the analysis. The final aim is to reduce the number of waste scenarios but more adjusted to a local context at small-scale.

References

[1] Coronado, M., Dosal, E., Coz, A., Viguri, J., Andrés, A. (2011). Estimation of construction and demolition waste (c&dw) generation and multicriteria analysis of c&dw management alternatives: a case study in spain, Waste and Biomass Valorization 2 (2) (2011) 209{225}.

[2] Pearson, K. (1901)Principal components analysis, The London, Edinburgh and Dublin Philosophical Magazine and Journal 6 (2) (1901) 566.

[3] Hotelling, H. (1933). Analysis of a complex of statistical variables into principal components., Journal of educational psychology 24 (6) (1933) 417.

[4] Jolliffe, I. (2002). Principal component analysis, Wiley Online Library, 2002.

[5] Asase, M., Yanful, E. K., Mensah, M., Stanford, J., & Amponsah, S. (2009). Comparison of municipal solid waste management systems in Canada and Ghana: A case study of the cities of London, Ontario, and Kumasi, Ghana. *Waste Management*, *29*(10), 2779-2786.

[6] Sharholy, M., Ahmad, K., Mahmood, G., & Trivedi, R. C. (2008). Municipal solid waste management in Indian cities–A review. *Waste* management, *28*(2), 459-467.

[7] Wilson, D. C., Rodic, L., Scheinberg, A., Velis, C. A., & Alabaster, G. (2012). Comparative analysis of solid waste management in 20 cities. *Waste Management & Research*, *30*(3), 237-254.

[8] Zhang, L. X., Chen, B., Yang, Z. F., Chen, G. Q., Jiang, M. M., & Liu, G. Y. (2009). Comparison of typical mega cities in China using emergy synthesis. *Communications in Nonlinear Science and Numerical Simulation*, *14*(6), 2827-2836.