# An ANN-based Decision Support Tool for the Sustainable Performance Prediction of the Waste Management Systems

**Digiesi S.*, Facchini F.*, Mossa G.*, Monno M.*, Mummolo G.*, Verriello R.*,**

* *Dipartimento di Meccanica, Matematica e Management (DMMM) – Politecnico di Bari, Viale Japigia 182, 70126 - Bari – Italy (salvatore.digiesi@poliba.it, francesco.facchini@poliba.it, giorgio.mossa@poliba.it, mariacristinamonno@gmail.com, giovanni.mummolo@poliba.it, rossella.verriello@poliba.it )*

**Abstract**: Assessment and prediction of environmental and economic performance of waste management systems should be jointly considered in a sustainable planning of such a system. Complexity of the economic and the environmental evaluation of waste management system limits the adoption of analytical tools by policy makers. Consistently, the authors propose an Artificial Neural Network (ANN)-based decision support tool for the prediction of the optimized sustainable performance of an integrated waste management system. It stands as a user-friendly dashboard designed for a local-policy maker who seeks to have insights into potential effects of different waste management policy mainly on greenhouse gases (GHGs) emissions and monetary savings. Indeed, starting from both demographic and urban fabric features, the tool predicts the most suitable collection configuration, the flows managed as well as the amount of CO2eq emitted/avoided and relative financial flows. The ANN developed is trained through data deriving from an analytical optimization model. The modeling code is developed by Alyuda NeuroIntelligence™. Results show the good reliability of the ANN in the prediction of the optimized sustainable performance for the waste management system.

**Keywords**: ANN, Waste Management, DSS, Sustainability

## 1. Introduction

In the last decades increasing attention has been paid to the concept of sustainable production and a new one, the "Green Impact", has been introduced to indicate all activities aiming at measuring and minimizing the negative effects on the environment of industrial activities (Digiesi et al., 2012). Creating an economy based on an efficient resource use, on a low greenhouse gases (GHGs) emissions level as well as on a climate resilience future starting the city level, is a milestone recognized by (IPPC, UNEP). Among the different urban action areas, the waste management systems demand for substantial investigation. Its strategic position allows to effectively counteract GHGs emissions increasing level through good practices such as prevention, reuse, recycling and *'waste to energy'* processes (EEA, 2011; European Commission). A huge contribution originate from the above mentioned practices also from the economic point of view by the production of secondary raw materials as well as of 'green energy source'.

Assessment and prediction of environmental and economic performance of waste management systems should be jointly considered in a sustainable planning of such a system. In literature, several contributions can be recognized in that field. In (Pires et al., 2011), systems analysis models are grouped into two categories: systems engineering models and system assessment tools. Systems engineering models help technicians and expert in identifying most suitable technologies and practices as well as site facility and performance prediction for a system that should be 'engineered' and that does not yet exist tailored with the particular analysis target.

They are typically carried out by optimization model (Gnoni et al., 2008; Nguyen-Trong et al., 2017; Anghinolfi et al., 2013), Analytic Hierarchy Model, simulation model (Younes et al., 2016; Sukholthaman and Sharp, 2016). On the other hand, system assessment tool helps in the performance evaluation of a system that already exists. They are typically carried out by scenarios analysis and life cycle assessment (Muñoz et al., 2004; den Boer et al., 2007; Wang et al., 2014). Complexity of the economic and the environmental evaluation of waste management system limits the adoption of analytical tools by policy makers. Indeed, as evident by the number of works developed in recent years, a particular research area concerns the development of tools for sustainable performance evaluation for public organization or private companies. The Waste Reduction Model (WARM) developed by U.S. Environmental Protection Agency (EPA), starting from a collection system defined, allows the assessment of GHGs emissions resulting from different waste management practices. In (Sevignè Itoiz et al., 2013), a tool is developed for the monitoring of emissions starting from the total amount of waste generated, waste composition, waste fraction collected, biogas captured in landfill. In the frame of the European project "RES NOVAE", a web-app named "Smart Waste - Carbon Footprint Calculator (SW-CFC)" is conceived as a service planning tool as well as an environmental assessment tool for a pivotal phase, which is that of the municipal waste collection (Digiesi et al., 2016). For a better and comprehensive understanding, the evaluation should address all phases of the integrated waste management system even in an economic perspective.

### 1.1 Aim of the study

To overcome the drawbacks highlighted previously, the authors propose an Artificial Neural Network (ANN)-based decision support tool for the prediction of the optimized sustainable performance of an integrated waste management system. It stands as a user-friendly dashboard designed for a local-policy maker who seeks to have insights into potential effects of different waste management policies mainly on GHGs and monetary savings. The ANN allows a non-technical user an easier planning of the waste integrated system and easy extensibility to different contexts of application, compared to the optimization model. In the Section 2, both the analytical optimization model and the ANN model are described. Results obtained and insights into the model validation are shown in Section 3, while conclusions are presented in Section 4.

## 2. Materials and Method

Waste management is a really complex process task. Its complexity relies on the high number of variables involved and that must be considered not only as individual elements of a macro system but chiefly in their interrelated relationships and influences on environmental and economic performance. Searching for a technological option that fits an assigned objective function requires the adoption of a systematic approach capable to quantify technical (e.g. management and / or environmental ) and economic variables.

### 2.1 The analytical optimization model

The analytical optimization model used to train the ANN described in the following section, is a Mixed Integer Non Linear Programming Model (MINLP). The aim of the MINLP model is to find out the municipal waste integrated management system with the lowest carbon footprint. The model works on both operative and strategic levels. At the operative level, the model allows the identification of the collection system configuration: grouping modality, collection system ad weekly collection density. In an urban center, the collection systems largely depend on both areas' building features and on urban fabric constraints (D'Alessandro et al., 2012). The 'door-to-door' is the best choice for quality of collected materials and users monitoring feasible in case of urban center and old town with a widespread presence of independent houses with private backyards and low urban density. For small urban aggregates in old town or urban center with private common waste storage areas, an aggregate curbside collection system can be adopted. The proximity collection system is feasible in case of users living in high, but not strictly contiguous buildings in urban center or residential area. Otherwise, the street collection system can be adopted for a large number of users living in city suburbs showing large streets and contiguous residential building (usually blocks without common open areas). At the strategic level, the model allows the selection of the most suitable treatment from the environmental point of view for each waste fraction collected. The treatments foreseen for the dry waste fraction separated collected (paper, plastics and glass), aim to separate materials that have enough value from the remaining flow of impurities to make their recovery worthwhile. For the organic fraction, the treatments foreseen can be the anaerobic or aerobic digestion to respectively produce energy from biogas and/or compost The mechanical biological treatment is envisaged as a preliminary treatment for the remaining mixed waste. Downstream, the biostabilized organic flow is addresses to landfill, while the dry remaining flows is addressed to waste to energy processes.

The objective function is defined as the algebraic sum of the emissions (Em) due to municipal integrated waste management system phases as in (1):

$$Min\ (Em_{collection} + Em_{DrywasteTreatment} + Em_{organicTreatment} + Em_{MixedWasteTreatment} + Em_{landfill} + Em_{energy.recovery} + Em_{material.recovery}) \tag{1}$$

With subscripts in table 1, the decision variables of the model are:

- $Y_{i,k}$: boolean variable to infer if the i-th grouping system is adopted for the k-th waste fraction;
- $N_{j,i,k}$ :number of users served by the j-th collection system and the i-th waste stream grouping for the k-th waste fraction;
- $W_{j,i,k}$: weekly collection frequency of the k-th fraction, by the j-th collection system according to the i-th waste stream grouping.

The model is powered by distinctive socio-demographic data of the application context. The determination of the optimal value of the variables results in:

- organization of the collection service at operational level. It determines the number of users to be served with each of the listed modality in Table 1. Collection emissions are therefore dependent on the number of shifts needed and the distances spent to ensure the service.
- evaluation of the amount of the flows to be treated at a strategic level. The amount of streams to deal with is strictly dependent on the organization of the collection service.

Emissions are evaluated following IPCC Guidelines (2006). In particular, for the collection a distance-based approach for estimating CO2eq emissions for each type of vehicle, payload and average speed is followed. For what concern the treatments phase, both emissions resulting from the fuel consumption according to a fuel-based approach, and indirect emissions due to energy use is accounted for. Avoided emissions derive from both the substitution of energy and of materials derived from waste as an alternative source. Materials recovery from waste and subsequent recycling leads to avoided GHG emissions Avoided emissions are calculated as difference between emissions associated with the production of a product from virgin raw materials and emissions associated with the production of the same from secondary raw materials. For further details relating the optimization models readers can refer to (Digiesi et al., 2015).

**Table 1.** Subscripts legend

| Symbol | Description |
|---|---|
| *k* | Waste fractions k = 1 organic; k = 2 glass; k=3 paper; k=4 plastics, metal cans; k=5 cardboard; k=6 wood; k=7 textile; k=8 un-recyclable; k=9 others. |
| *i* | Grouping Systems i =1 mono material; i=2 multi stream. |
| *j* | Collection Systems j=1door-to-door; j=2 aggregate; j=3 proximity; j=4 street |

The economic evaluation is conducted to determine the revenue for the municipalities resulting from an efficient integrated waste management system. The fees ensured by the waste consortia have been considered paper, plastics, aluminium cans and glass (CONAI, 2014). The energy price per kilowatt-hour for the electricity produced. The optimization model is used to train the ANN described in the following section.

### 2.2 The ANN Simulation Model

The ANN simulation model is developed in order to establish a relationship between the municipal Solid waste-flow and the general characteristics of the city/town. The prediction model is based on the adoption of ANN characterized by an inherent ability to learn and recognize highly the nonlinear relationships and then organize dispersed data into a nonlinear model (Xiao et al., 2009). This learning technique mimics the biological learning process occurring in the brain, the neural networks present a robust way to predict actual-value after learning from a supplied sample set. The networks connect are based on a number of individual elements, each of which take a set of inputs and produce a single real number. The learning algorithm determines the numeric weights to be applied between each of these neurons to obtain the desired output (Facchini et al., 2013).

Different ANNs are adopted, they are used for identify the following targets:

- Waste collection arrangements;
- Quantity of waste separated collected and resulting mixed waste;
- Net yearly emissions due to both direct contribution (collection, treatments and valorization) and avoided emissions (materials and energy recovery );
- Economic evaluation due to monetary valorization of waste separated collected in consortium and as well as of energy produced by resulting mixed waste.

**Table 2.** List of output parameters for each ANN

| Target | ANN ID | Output |
|---|---|---|
| Waste collection arrangements | #1 | Collection typology ($C_t$): door-to-door, aggregate, proximity and street collection system |
| Quantity of waste in separate collections and mixed waste | #2 | Bio-waste ($B_w$) [kg/year] |
| | #3 | Paper ($Pa_w$) [kg/year] |
| | #4 | Plastic ($PL_w$) [kg/year] |
| | #5 | Glass ($G_w$) [kg/year] |
| | #6 | Mixed waste ($Mix_w$) [kg/year] |
| Net Yearly emission | #7 | Emission avoided ($E_a$) [tCO2eq/year] |
| Economic evaluation | #8 | Separated Collection Income (CI) [k€/year] |

For each ANN is identified a set of different output parameters (tab. 2) on the bases of four input variables:

- District typology: residential, urban center and old town;
- Population density: low, medium and high;
- Number of citizens;
- Percentage level of separate collections it needs to be ensured.

Two main phases are identified for the design of the prediction model ANN-based, the first phase consisting of data splitting and the second phase consisting of design of the ANN architecture followed by the identification of the best learning algorithm.

#### 2.2.1 Data splitting

The appropriate data splitting can be handled as a statistical sampling problem. Therefore, various classical sampling techniques can be adopted in order to split the data in three subset for training, validation and testing of ANN, most commons are: Simple random sampling (SRS), Trial-and-error methods, Systematic sampling, and Convenience sampling. In this case, Trial-and-error method is adopted, the splitting strategy try to overcome the high variance of the SRS by repeating the random sampling several times in order to minimize the Mean Square Error (MSE) of the ANN. This technique is high time-consuming and requires significant computational costs. For this analysis the data splitting phase is developed on Core i7-4510U with 8GB RAM.

On the basis of best MSE identified, a subset as big as 60% of the available experimental data, which was composed by 72 inputs/output pairs, is used for the ANN training. In this phase, the synaptic weights, which are the links between neurons, have a synaptic weight attached. They are updated repeatedly in order to reduce the error between the experimental outputs and the associated forecasts.

A subset of 20% of the sample was adopted for validation. In particular the adoption of validation sets allows to identifying the underlying trend of the training data subset, avoiding the overfitting phenomenon. As far as concern the testing phase, a sample of 20% is used for testing the forecast reliability of the ANN in the learning phase.

In order to deal with the overfitting problem,the training phase is stopped when the mean square error (MSE) assumed values lower than 0.01. In most cases are required about 20.000 iterations.

#### 2.2.2 ANN Design

The design of the ANN architecture consists in identifying the number of hidden layers and the number of neurons for each layer. On one hand, many neurons can lead to memorize the training sets with lost of the ANN's capability to generalise. On the other hand, a lack of neurons can inhibit the appropriate pattern classification. In this work, different networks are tested varying the number of hidden layers and the number of neurons in the hidden layer. For each architecture, the software provide to compute a "fitness bar" based on the inverse of the mean absolute error (MAE) on the testing set.

In every cases the best accuracy is achieved adopting ANNs characterized by architecture with only one hidden layers. The number of input nodes (N), hidden (K), and output (M) nodes for every network, are given in table 3.

Two different training algorithms namely Quick Propagation (QP) and Conjugate Gradient descent (CG) are adopted.

**Table 3.** Learning algorithm and architecture adopted for each ANN

| ANN | Architecture (N-K-M) | Learning Algorithm |
|---|---|---|
| #1 | 4-2-1 | Quick Propagation |
| #2 | 4-18-1 | Quick Propagation |
| #3 | 4-14-1 | Quick Propagation |
| #4 | 4-11-1 | Quick Propagation |
| #5 | 4-28-1 | Conjugate Gradient Descent |
| #6 | 4-17-1 | Quick Propagation |
| #7 | 4-11-1 | Quick Propagation |
| #8 | 4-18-1 | Quick Propagation |

QP is a heuristic modification of the standard back propagation, the output of the *m*-th output node for the *p*-th input pattern is given by $o_{pm}$ (eq. 2).

$$o_{pm} = f\left(\sum_{k=1}^{K} \overline{\omega}_{km} o_{pk}\right) \qquad (2)$$

Where *f* is the activation sigmoidal function (eq. 2), $\overline{\omega}_{km}$ is the weight between the *m*-th output neuron and the *k*-th hidden neuron. The value of $o_{pk}$ depends by two parameters: the first is given by the weight between *k*-th hidden neuron and the *n*-th input neuron ($\overline{\omega}_{nk}$). The second parameter is $x_{pn}$ given by *p*-th input pattern of *n*-th neuron.

$$f(x) = \frac{1}{(1 + e^{-x})} \qquad (3)$$

All network weights are updated after presenting each pattern from the learning data set.

The CG learning algorithm starts with a random weight vector that is iteratively updated according the direction of the greatest rate of decrease of the error evaluated as $\omega^{(\tau)}$ in equation 4.

$$\Delta\omega^{(\tau)} = -\eta \nabla E_{\omega(\tau)} \qquad (4)$$

Where $E$ is the error function evaluated at $\omega^{(\tau)}$ and $\eta$ is the arbitrary learning rate parameter. For each step $(\tau)$ the gradient is re-evaluated in order to reduce $E$.

The performance of the gradient descent algorithm is very sensitive to the proper setting of the learning rate, in case $\eta$ is too high the algorithm can oscillate and become unstable, for $\eta$ too small the algorithm takes too long to converge. In this case an adaptive learning rate allows to keep the learning step size as large as possible, ensuring, in this way, the learning rate stable.

## 3. Results and model validation

The parameters predicted by the model are compared to the actual outputs (fig. 1). In order to evaluate the reliability of the model the Mean Absolute Percentage Error (MEPA) and p-values are computed for each ANN.

It is very interesting noted that for all cases the MAPE values are less than 10% and the p-values are very low (tab. 4). In other words the reliability of the ANN simulation model is very high.

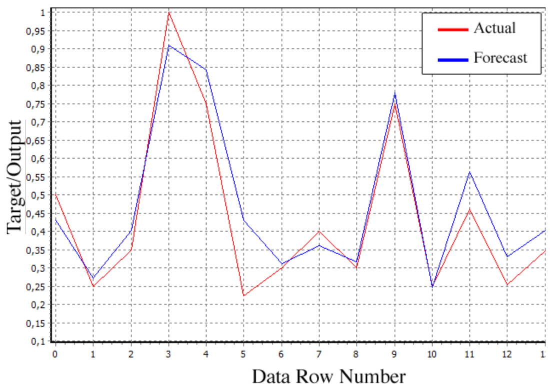


**Figure 1.** Actual vs. forecast in case of $Pl_w$ value prediction

As far as concern the output provided by ANNs #7 and #8 are identified as 'categorical' and not 'numerical' output (see tab. 3). In these cases the high variability of the actual values not allowed to obtain a good estimation of the predicted parameters. Therefore five different levels of cost due to waste collection strategies adopted and to the amount of emission avoided are defined (see tab. 4).

**Table 4.** MAPE and p-value computed for each output provided by ANN simulation model

| Output | Categorical/ Numerical | MAPE [%] | p-value |
|---|---|---|---|
| $C_t$ | Categorical | 4.55 | - |
| $B_w$ | Numeric | 3.43 | 2.74E-20 |
| $Pa_w$ | Numeric | 6.59 | 4.72E-15 |
| $Pl_w$ | Numeric | 8.24 | 1.55E-12 |
| $G_w$ | Numeric | 5.52 | 7.03E-18 |
| $Mix_w$ | Numeric | 3.70 | 6.39E-21 |
| $E_a$ | Categorical | 9.09 | - |
| CI | Categorical | 9.09 | - |

The forecast values are in excellent agreement with the actual ones, showing that the developed model is very accurate and has a greater aptitude for evaluating the environmental and economic performance of waste management systems. In order to predict the amount of the Ea and C the increase of the data set used for the training of the ANN is required. Furthermore the inclusion of the new experimental dates allows to enhance the reliability of the forecast and would improve the control of the overfitting phenomenon in phase of the ANN training.

**Table 5.** Range for each category in case of Ea and C parameters

| Output | Level | Range parameters |
|---|---|---|
| $E_a$ [tCO2eq/year] | A | $E_a \leq 1000$ |
| | B | $1000 < E_a \leq 2000$ |
| | C | $2000 < E_a \leq 3000$ |
| | D | $3000 < E_a \leq 4000$ |
| | E | $4000 < E_a \leq 5000$ |
| | F | $E_a > 5000$ |
| C [k€/year] | Very cheap | $C \leq 300$ |
| | Cheap | $300 < C \leq 500$ |
| | Medium | $500 < C \leq 700$ |
| | Expensive | $C > 700$ |

## 4. Conclusions

The complexity and relative difficulty to plan waste management systems and consequently to evaluate economic and environmental performance, is satisfactory solved by adopting an ANN. Indeed the ANN, accurately trained by an analytical optimization model, allows to gather relationships existing among the key features of waste management system, determining the sustainable performance usually matter of concern of a policy maker. Consistently, the developed ANN stands as a user-friendly dashboard designed for a local-policy maker who seeks to have insights into potential effects of different waste management policy mainly on GHGs and monetary savings. Results show the ANN effectiveness in the prediction of the waste flows separated, collected, the GHGs emissions avoided as well as the economic income. The reliability of the model is assessed through the Mean Absolute Percentage Error and p-values. While the MAPE values are less than 10%, the p-values are very low showing that the forecast values are in excellent agreement with the actual ones.

Finally, this work suggests that the full integration of analysis, prediction and controlling with a continuous learning in that user-friendly tool esteeming both economic and environmental results is promising. Further developments will be oriented towards the inclusions in the above mentioned tools of waste management costs for both collection and treatments and the challenging assessment of the sustainable social dimension.

## References

Anghinolfi, D., Paolucci, M., Robba, M., Taramasso, A.C., 2013. A dynamic optimization model for solid waste recycling. Waste Manag. 33, 287–296. doi:10.1016/j.wasman.2012.10.006

CONAI,"Accordo ANCI-CONAI 2014-2019, Allegato Tecnico", available at: [http://www.conai.org/enti-locali/accordo-quadro-anci-conai/](http://www.conai.org/enti-locali/accordo-quadro-anci-conai/) (Accessed on 21/02/2017).

Digiesi, S., Mossa, G., Mummolo, G., Verriello, R. (2016), A web-app for implementing a 'Carbon Footprint Calculator' for smart waste management systems, Published on Procedia Environmental Science, Engineering and Management 3 (3) 165-172.

Digiesi, S., Facchini, F. , Mossa, G., Mummolo, G., Verriello, R. (2015). A Cyber – based DSS for a Low Carbon Integrated Waste Management System in a Smart City. IFAC Papers Online, vol. 48, p. 2356-2361, Elsevier Ltd.Pub., ISSN: 1474-6670;

Digiesi, S., Mossa, G., Rubino, S. (2012). Sustainable Order Quantity of Repairable Spare Parts. IFAC Papers Online, vol. 45 (31), p. 181-186, Elsevier Ltd.Pub.

EPA-Waste Reduction Model, online at: http://epa.gov/epawaste/conserve/tools/warm/index.html.

European Commission, EC (available at: http://ec.europa.eu/environment/waste/);

European Environment Agency (EEA), 2011, Waste opportunities. Past and future climate benefits from better municipal waste management in Europe. Technical report No 3/2011 ISSN 1725-9177, doi:10.2800/79764;

D'alessandro G, Gnoni G, Mummolo G (2012). Sustainable Municipal Solid Waste Management and Recovery. In: Editors: C.N. Madu and C. Kuei. Handbook of Sustainability Management, Imperial College Press 2012, London. ISBN: 978-981-4354-81-3

den Boer J., den Boer E., Jager J., 2007. LCA-IWM: a decision support tool for sustainability assessment of waste management systems. Waste Management, 27, 1032–1045;

Facchini, F., Mossa, G., and Mummolo, G. (2013). A model based on artificial neural network for risk assessment to polycyclic aromatic hydrocarbons in workplace. 25th European Modeling and Simulation Symposium, EMSS 2013; Athens; Greece; 25 September 2013, p. 282-289.

Gnoni M., Mummolo G., Ranieri L. , 2008. A mixed integer linear programming model for optimisation of organics management in an integrated solid waste system. Journal of Environmental Planning and Managemen, vol. 51, p. 833-845, ISSN: 0964-0568;

Intergovernmental Panel on Climate Change (IPCC), 2006, Guidelines for National Greenhouse Gas Inventory, available at: http://www.ipcc-nggip.iges.or.jp/public/2006gl/vol1.html;

Muñoz, I., Rieradevall, J., Domènech, X., Milà, L., 2004. LCA application to integrated waste management planning in Gipuzkoa. The International Journal of Life Cycle Assessment, 9 (4), 272-280;

Mummolo G., 1995, Analytic hierarchy process model for landfill site selection, Journal of environmental system, Volume 24, Issue 4, 1995, Pages 445-465;

Nguyen-Trong, K., Nguyen-Thi-Ngoc, A., Nguyen-Ngoc, D., Dinh-Thi-Hai, V., 2017. Optimization of municipal solid waste transportation by integrating GIS analysis, equation-based, and agent-based model. Waste Manag. 59, 14–22. doi:10.1016/j.wasman.2016.10.048

Pires A, Martinho G., Chang N.B, 2011, "Solid waste management in European countries: A review of systems analysis techniques," Journal of Environmental Management, vol. 92, no. 4, pp. 1033-1050.

Sevignè E.I, Gasol C.M, Farreny R., Rieradevall J., Gabarell X., (2013), CO2ZW: Carbon footprint tool for municipal solid waste management for policy options in Europe. Inventory of Mediterranean countries, Energy policy 56, 623-632.

Sukholthaman, P., Sharp, A., 2016. A system dynamics model to evaluate effects of source separation of municipal solid waste management: A case of Bangkok, Thailand. Waste Manag. 52, 50–61. doi:10.1016/j.wasman.2016.03.026

Wang, X., Chen, Y., Sui, P., Gao, W., Qin, F., Wu, X., Xiong, J., 2014. Efficiency and sustainability analysis of biogas and electricity production from a large-scale biogas project in China: an emergy evaluation based on LCA. Journal of Cleaner Production 65, 234-245.

Younes, M.K., Nopiah, Z.M., Basri, N.E.A., Basri, H., Abushammala, M.F.M., Younes, M.Y., 2016. Landfill area estimation based on integrated waste disposal options and solid waste forecasting using modified ANFIS model. Waste Manag. 55, 3–11. doi:10.1016/j.wasman.2015.10.020

United Nations Environmental Program (UNEP, 2010), Waste and climate change: Global Trends and strategy framework, Osaka, 2010.

Xiao, G., Ni, M., Chi, Y., Jin, B., Xiao, R., Zhong, Z., and Huang, Y. (2009). Gasification characteristics of MSW and an ANN prediction model. *Waste Management,* 29, 240-244.