FUZZY MODELLING OF THE COMPOSTING PROCESS

E. Giusti*, S. Marsili-Libelli *, A. Burchi**

* Department of Systems and Computers, University of Florence, Via S. Marta 3, 50139, Florence, Italy (E-mail: *marsili@dsi.unifi.it*)

** Ecomas srl, Via Caduti di Cefalonia 76, 50127 Florence, Italy

Abstract

Successful composting implies a careful monitoring of the batch process, which includes two phases: active composting and curing. In the first phase, a succession of microbial activities transform the organic material into progressively biodegradable matter, whereas during the second phase, stabilization ensues. The key to efficient composting lies in the definition of an appropriate temperature batch curve and in an effective control system to track such a profile. The control system should rely on a process model in order to predict the batch behaviour, in terms of temperature, and adjust the air supply accordingly.

Keywords. *Composting, Fuzzy models, Process control*

INTRODUCTION

Composting is a biological treatment process that is largely applied to the biodegradable fraction of solid waste to dispose of rubbish and obtain useful fertilizer. This process is attractive for its low energy consumption, limited technological requirements and commercial value of end-product. The composting process consists of the biochemical degradation of organic materials through a controlled microbial aerobic decomposition process with the formation of stabilized organic materials that may be used as soil conditioners and organic fertilizers.

Composting is the biochemical degradation of organic materials to a safe, humus-like material through a three-phase microbial aerobic process with the formation of stabilized organic materials that may be used in agriculture as an organic fertilizers (Haug, 1980; Negro et al., 1999). The main factors in the control of a composting process include environmental parameters, such as temperature, moisture content, pH and aeration (Liang et al., 2003; Kulcu and Yaldiz, 2004). The main products of biological metabolism are carbon dioxide, water, and heat and aeration has the twofold effect of providing oxygen for the biological reactions and removing the heath which they produce. For this reason temperature is regarded as the single most important parameter being both an indicator of process condition and the main reaction conditioner (Kulcu and Yaldiz, 2004). For this reason temperature profiles during the batch are highly important and the control system is designed to follow a profile which is considered to yield the most effective reaction in terms of disinfection, moisture content and degradation (Mason and Milke, 2005; Mason, 2006). Two typical temperature profiles are shown in Figure 1, where two main phases can be seen: the initial mesophilic phase when disinfection occurs, followed by a termophilic phase, when the biological reaction in the compost are completed. The former refers to the initial high-temperature period when the biological reactions are very fast a require a considerable amount of oxygen. The latter refers to the degradation of the less reactive substances, with a decreased oxygen requirements. It is important that during the first phase the temperature is high enough to guarantee disinfection. For this reason the

ECOMAS reference curve is higher in the first few days. The purpose of the control system is to follow the reference temperature profile of Figure 1 as closely as possible during the whole batch. For this reason a process model is required, in order to predict the future temperature evolution and modulate the air flow accordingly.

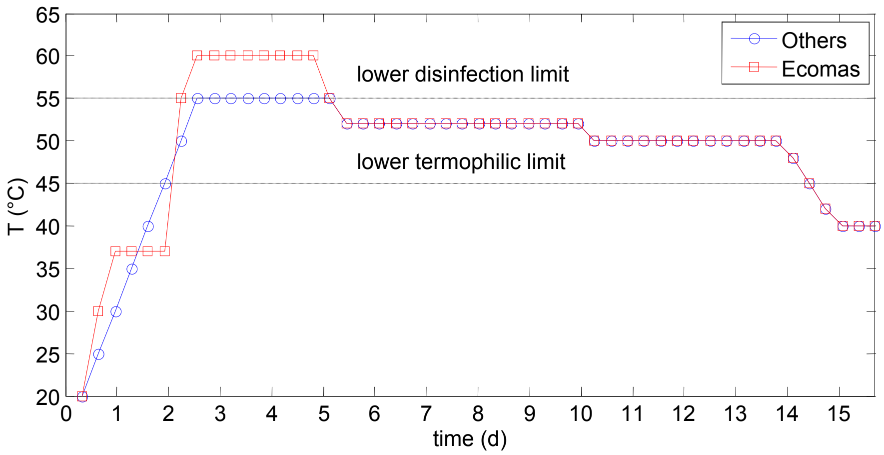

Figure 1. Typical temperature profile during the composting batch.

FUZZY MODELING OF THE COMPOSTINGPROCESS

A fuzzy approach is pursued in developing a dynamical model of the composting process and its performance is evaluated with batch data from full-scale composting plants. The reason for preferring a fuzzy model to a deterministic one (see e.g. Sole-Mauri et al., 2007) lies in the inherent complexity of the composting process, with a multi-stage succession of fungi and bacteria colonizing the organic material. This succession is difficult to model in deterministic terms (Sole-Mauri et al., 2007), whereas identifying a fuzzy model from operational data is reasonably simple, though it provides a moderate degree of insight into the process biochemistry. Nevertheless, a preliminary exploratory phase is useful to gain a knowledge of the basic process behaviour in terms of temperature profile.

The approach used in this paper follows the guidelines of Babuska (1998) and Abony (2003) in setting up a Sugeno fuzzy model with clustered antecedents and least-squares estimation of consequents. This approach proves very useful in modelling complex systems, for which rule-based conventional fuzzy models would result in an explosive proliferation of rules. A multi input single output (MISO) model was set-up using the product space fuzzy clustering based on the Gustafson-Kessel algorithm using data from a 400 m^3 biotunnel for which airflow/temperature coupled data were available for the whole batch duration, with the former being expressed as the ON/OFF setting measured at the air valve and the output being the temperature measured inside the compost pile.

Structure of the fuzzy model

When approximating a complex nonlinear dynamical system, a single linear model is often insufficient, especially when the nonlinear system has several operating regimes with significantly differing characteristics. Examples of this behaviour can be found, for example, in biotechnical processes, where the system behaviour strongly depends on biomass viability. This problem can be solved by the decomposition of the system operating range into a set of regimes to which local linear models are associated. The global model is obtained through a smooth interpolation of the local models, giving the most weight to the one which is thought to best approximate the system in the current operating point and less weight to those which are considered to be far from the present condition. A classifier is needed to decide which linear model is best approximating the current system operation and select the interpolating weights accordingly. This classifier should provide a set of smooth interpolating functions to combine the individual local models. This approach (regime identification and multi-model blending) can be naturally recast into the fuzzy set theory which provides an easy way for solving both the regime identification (fuzzy clustering) and the model blending problem (consequent models identification).

Model development

The Nonlinear ARMAX model (NARMAX) has often been used to approximate nonlinear systems representing its behaviour by nonlinear mapping of past inputs, outputs, and noise terms to obtain future outputs (Chen and Billings, 1989). For this application a multi input single output deterministic model with measurement noise is used to describe the relationship between airflow and temperature in the biotunnel, namely

$$y(k) = f\left(y(k-1),....,y(k-n_y),u(k-1),...,u(k-n_u)\right) + e(k) \tag{1}$$

where *y*, *u* and *e* represent the process output, input and measurement noise respectively.

The problem of decomposing the system behaviour into a number of significant operational modes has been considered from the fuzzy point of view by Hathaway and Bezdek (1993) and then by Babuska (1998) and Abony (2003). In many instances, like in this application, the features ***z****(k)* describing the operational state may differ from the input/output data, though they are generally strictly related to them. Let ***z****(k)* represent the current operating point in the *operational space Z* defined as a function of the past input/output samples,

$$z(k) = f\left(y(k),y(k-1),...,y(k-n_{z_y}),u^T(k),u^T(k-1),...,u^T(k-n_{Z_u})\right) \in \boldsymbol{Z} \tag{2}$$

where in general $n_y \neq n_{Zy}$ and $n_u \neq n_{Zu}$. Once a suitable features space ***Z*** has been defined, the problem of mode recognition consists of determining the most appropriate complete partition $\boldsymbol{Z}_i \subset \boldsymbol{Z}$ of all the process modes and attach a local linear model to each of them. This approximation problem, as posed by eqs. (1-2), can be recast into a fuzzy context in a quite natural way. From the structural point of view, TS models (Takagi and Sugeno, 1985) can indeed be viewed as a fuzzy extension of the deterministic interpolation concept. Further, a fuzzy decomposition yields a higher computational efficiency since the subsets Z_i are no longer disjoint, but contribute to the global model through their degrees of matching. Lastly, partitioning of the operational space ***Z*** into $\boldsymbol{Z}_i \quad i = 1...M$ can be obtained in a systematic way through fuzzy clustering (Babuska, 1998; Abony, 2003). In addition to requiring that the resulting fuzzy memberships must be smooth functions, for each point $z(k) \in \boldsymbol{Z}$ defined by eq. (2), to ensure model completeness there must exists at least one cluster $Z_i \quad i = 1...M$ for which the corresponding membership function $\mu_{Z_i}(z_k) > 0$. The current operational state of the system can then be expressed by the set $\mu_{Z_i}(z_k) \quad i = 1,....,M$ of the condition memberships.

Though the basic model structure adopted here is based on TS models, the implication-based antecedent is substituted by a cluster structure to reduce model complexity and maintain the concept of regime decomposition. Thus the Fuzzy Clustering theory of Bezdek (1981), further developed by Babuska (1998) and Abony (2003) can be applied. With these premises, let the process be approximated in the generic i-th regime by a linear discrete-time model

$$\hat{y}_i(k+1) = \alpha_{o,i}y(k) + ... + \alpha_{n_y,i}y(k-n_y) + \beta_{o,i}u(k) + ... + \beta_{n_u,i}u(t-n_u) + \gamma_i \tag{3}$$

Then the overall model response can be obtained by combining the collection of models (3) through the degree of membership of the clusters defined in the operational space ***Z***: $z \subset Z_i \quad \rightarrow \quad \mu_{i,k}$, namely

$$IF\ z_k \subset Z_i\ THEN\quad \hat{y}_i(k+1) = \alpha_{o,i}y(k) + ... + \alpha_{n_y,i}y(k-n_y) + \beta_{o,i}u(k) + ... + \beta_{n_u,i}u(t-n_u) + \gamma_i \qquad (4)$$

where n_y and n_u are integers defining the dimensions of the consequent linear models (it is assumed that all models have the same dimensions).

The model output is obtained a fuzzy combination of the individual linear model outputs resulting in a *nonlinear* interpolation, or defuzzification

$$\hat{y}(k+1) = \frac{\sum_{i=1}^{M} \hat{y}_i(k+1)\mu_{Z_i}(z_k)}{\sum_{i=1}^{M} \mu_{Z_i}(z_k)} \qquad (5)$$

The complete fuzzy modelling structure is shown in Figure 2, showing the preliminary feature extraction block, followed by clustering, which produces the required degree of truth to be applied to each consequent model. After defuzzification according to eq. (5) the model prediction is obtained.

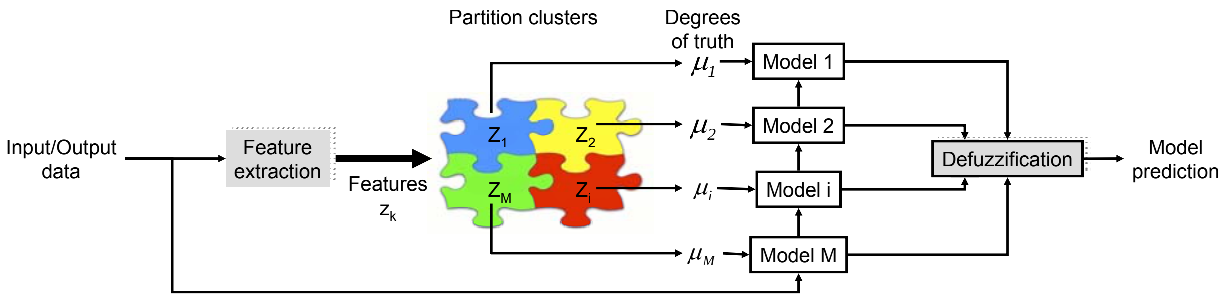

Figure 2. Overall structure of the fuzzy model. Each clustered antecedent yields the degree of truth which activates the corresponding model according to the clustering of the predefined features z_k.

Features extraction

Assuming that the basic process element is the OFF/ON aeration cycle, there are three basic modes in which the temperature may respond to the aeration, as described in Figure 3. The relevant variables defining each pattern are the temperatures at the beginning (T_a) and at the end (T_b) of the OFF portion of the cycle and at the end of the ON cycle (T_c), the maximum temperature swing (Td) and the difference between the initial and final temperatures, i.e. $\Delta = T_c - T_a$. The three possible responses are shown in Figure 3 and represent the features upon which the input/output data are clustered.

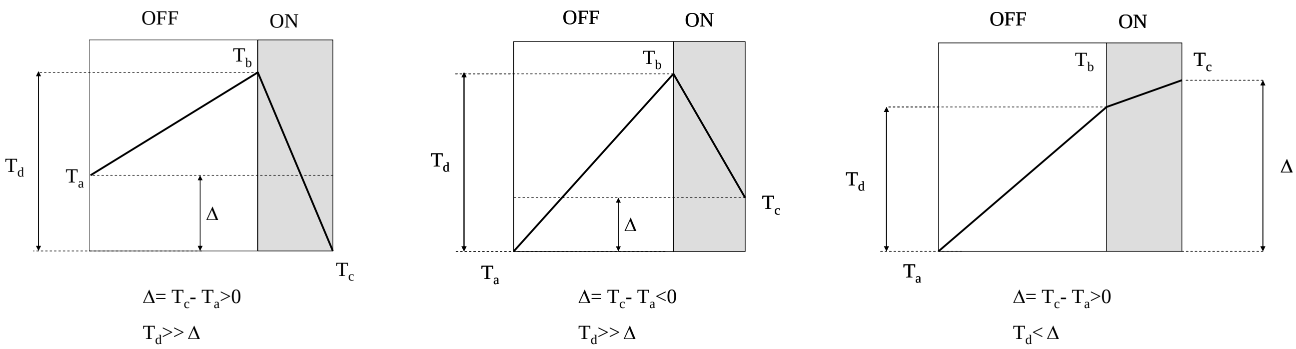

Figure 3. The three possible temperature patterns defining the operational models to be clustered.

In addition to the direct observations, suggesting that the modes of Figure 3 are indeed exhaustive, the optimal number of clusters has been searched with the XB criterion (Xie and

Beni, 1991) which considers the ratio of the total within clusters variation and separation of cluster. The optimal number of cluster should minimize the value of the index

$$XB(c)=\frac{\sum_{i=1}^{c}\sum_{j=1}^{N}(\mu_{ij})^{m}\left\|x_j-v_i\right\|^2}{N\,min_{ij}\left\|x_j-v_i\right\|^2} \tag{6}$$

where c is the number of clusters, $\boldsymbol{v}$ are their centroids and N the number of input/output process data. The selection obtained with the XB criterion was confirmed also by other clustering criteria such as the fuzzy hypervolume, the average partition density and the average cluster flatness (Babuska, 1998).

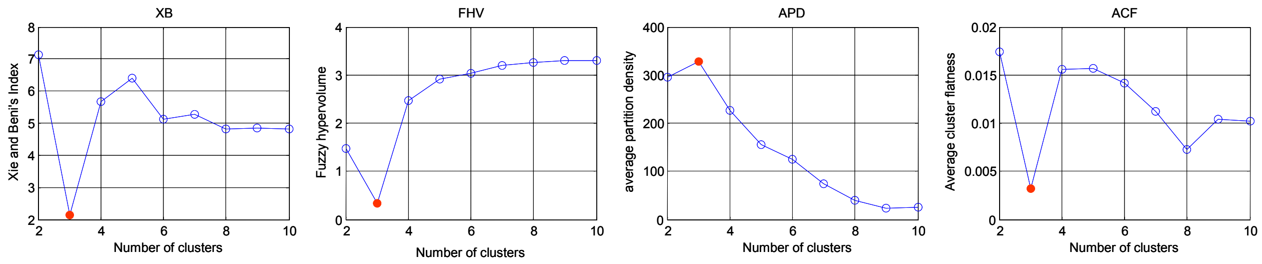

Figure 4. Four differing partition indicators agree in indicating an optimal partition into three clusters.

Model building

Several operational records were clustered using the Gustafson-Kessel fuzzy algorithm (Babuska, 1998) using the features of Figure 3 in order to extract information for the model-building process. Clustering the process data with the Gustafson-Kessel algorithm (Babuska, 1998) provided the following values of the cluster centers (centroids).

Features

$$\delta T_k = T_b - min(T_a, T_c)$$

$$\Delta_k = T_c - T_a \tag{7}$$

$$T_{max}(k)$$

Centroids

Table 1. Centroids of the three clusters describing the operational modes.

cluster	δT_k	Δ_k	$T_{max}(k)$
Z_1	1.90	-0.04	58.7
Z_2	1.81	0.02	49.9
Z_3	1.14	0.02	66.01

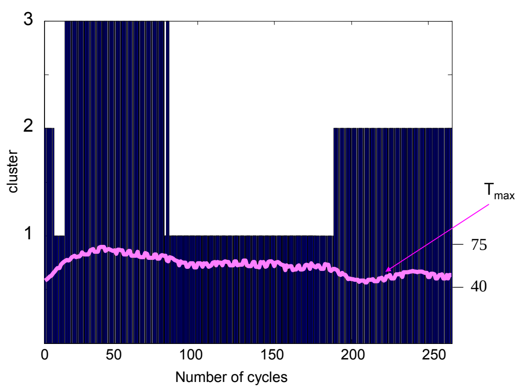

Figure 5. Cluster selection along the normal composting batch.

The original temperature and pressure data were split in two separate data streams. The pressure data provided information about the aeration cycle, duration and on/off ratio (duty cycle). The temperature data were first processed with a denoising wavelet decomposition; from its second approximation A_2 the cluster features of Figure 3 were extracted according to the following definitions eq. (7), with k being the cycle number, which is assumed as the basic time unit. The model inputs are composed of information on the aeration cycle (duration and duty cycle) and features extracted from the process (δT_k and Δ_k). These data are fed into the model identifier which, given the model order, calibrates the consequent sub-models, as in eq. (4) minimizing the sum of squared difference between the observed and computed $T_{max}(k)$ at each cycle.

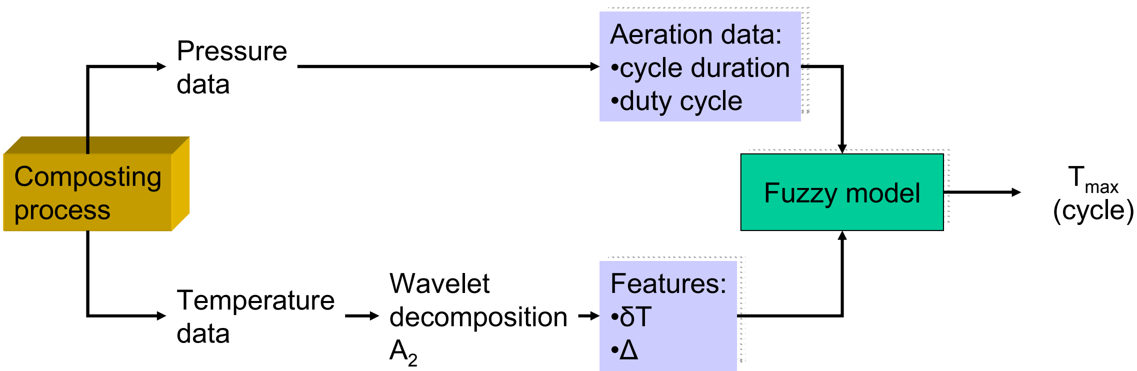

Figure 6. Structure of the model-building process to infer the maximum temperature reached in each cycle.

The complete model is shown in eq. (8) and its calibrated parameters are reported in Table 1. The actual model output is obtained from the three model outputs via defuzzification, according to eq. (4)

$$\begin{aligned}
&if \;\; z_k \subset Z_1 \; then \;\; \hat{y}_1(k+1) = \alpha_{1,1} y(k) + \alpha_{2,1} y(k-1) + \beta_{1,1} u_1(k-2) + \beta_{2,1} u_1(k-3) \\
&\qquad + \beta_{3,1} u_1(k-4) + \beta_{4,1} u_2(k-2) + \beta_{5,1} u_2(k-3) + \beta_{6,1} u_2(k-4) \\
&\qquad + \beta_{7,1} u_3(k-2) + \beta_{8,1} u_3(k-2) + \beta_{9,1} u_4(k-1) + \beta_{10,1} u_4(k-2) + \gamma_{1,1} \\
&if \;\; z_k \subset Z_2 \; then \;\; \hat{y}_2(k+1) = \alpha_{1,2} y(k) + \alpha_{2,2} y(k-1) + \beta_{1,2} u_1(k-2) + \beta_{2,2} u_1(k-3) \\
&\qquad + \beta_{3,2} u_1(k-4) + \beta_{4,2} u_2(k-2) + \beta_{5,2} u_2(k-3) + \beta_{6,2} u_2(k-4) \\
&\qquad + \beta_{7,2} u_3(k-2) + \beta_{8,2} u_3(k-2) + \beta_{9,2} u_4(k-1) + \beta_{10,2} u_4(k-2) + \gamma_{1,2} \\
&if \;\; z_k \subset Z_3 \; then \;\; \hat{y}_3(k+1) = \alpha_{1,3} y(k) + \alpha_{2,3} y(k-1) + \beta_{1,3} u_1(k-2) + \beta_{2,3} u_1(k-3) \\
&\qquad + \beta_{3,3} u_1(k-4) + \beta_{4,3} u_2(k-2) + \beta_{5,3} u_2(k-3) + \beta_{6,3} u_2(k-4) \\
&\qquad + \beta_{7,3} u_3(k-2) + \beta_{8,3} u_3(k-2) + \beta_{9,3} u_4(k-1) + \beta_{10,3} u_4(k-2) + \gamma_{1,3}
\end{aligned} \tag{8}$$

The model output is obtained by defuzzification of the three model outputs, according to eq. (4)

$$\hat{T}_{max}(k+1) = \frac{\sum_{i=1}^{3} \hat{y}_i(k+1)\mu_{Z_i}(z_k)}{\sum_{i=1}^{3} \mu_{Z_i}(z_k)} \tag{9}$$

Table 2. Model parameters after calibration.

regimes	Model parameters												
	α₁	α₂	β₁	β₂	β₃	β₄	β₅	β₆	β₇	β₈	β₉	β₁₀	γ₁
Z_1	-0.032	0.892	-0.145	-0.174	1.321	-2.511	2.505	-0.964	-6.856	6.720	-4.003	5.203	-0.913
Z_2	1.224	-0.248	0.335	-0.126	-0.141	3.656	-1.731	0.848	7.125	-6.987	6.352	-5.004	-8.889
Z_3	0.792	0.676	-3.715	2.047	-1.372	19.232	7.120	-5.944	-0.134	-0.010	-2.416	-0.202	2.733

The model identification was performed using the algorithm just outlined and the results shown in **Figure 7** were obtained. Figure 8 shows the agreement between the experimental data and the calibrated model response.

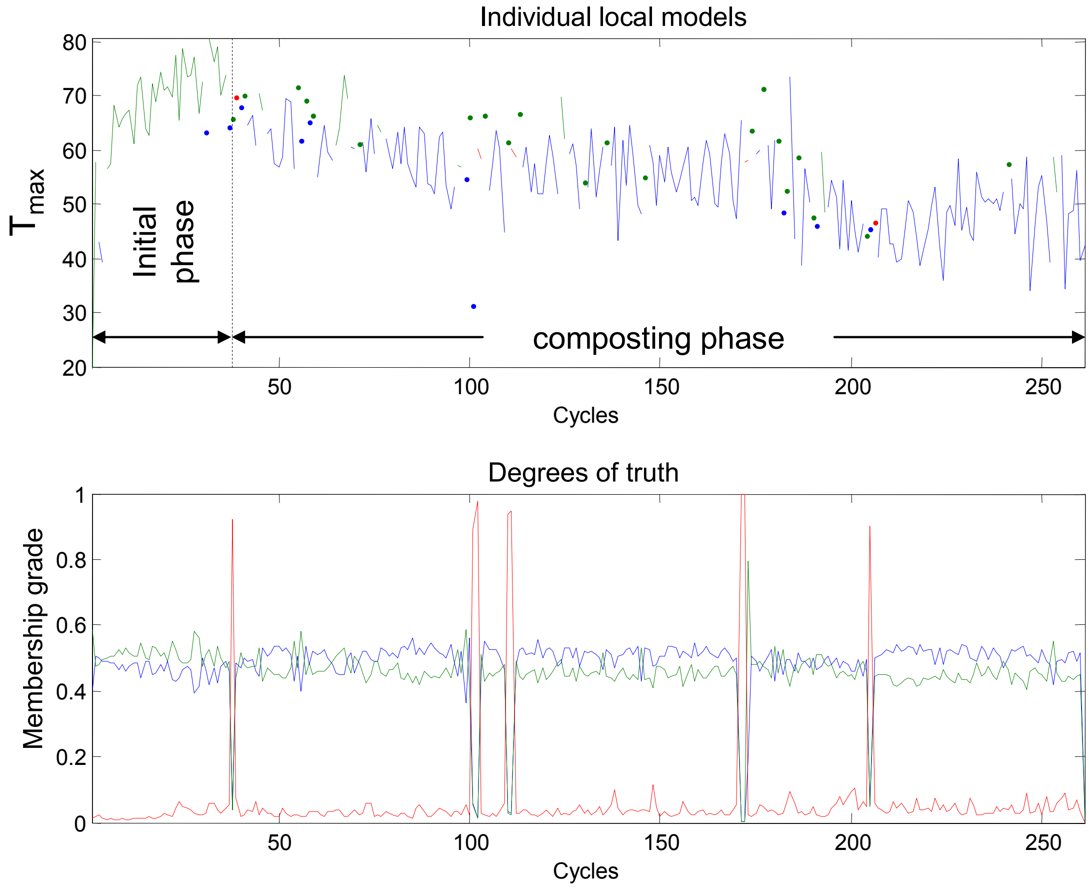

Figure 7. Fuzzy model identification results. The succession of prevailing memberships is consistent with the process evolution during the batch.

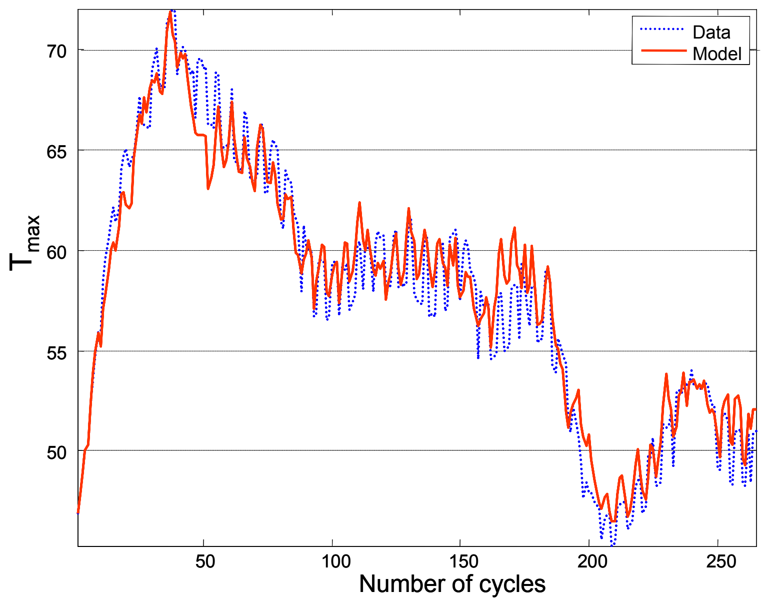

Figure 8. Agreement between the experimental temperature data and the identified model output.

CONCLUSIONS

This paper has presented a dynamical model of the composting process based on a hybrid fuzzy model with clustered antecedents and linear deterministic consequents. The model is based on air flow input data and temperature output data drawn from operating composting cells. The resulting data clusters are in agreement with the main process phases, including active composting and curing. The reason behind the model is that efficient composting requires accurate tracking of an appropriate temperature curve for which a model-based control system is required to adjust the air supply accordingly. The model inputs are the duration and duty cycle of the aeration cycle and features extracted from the process. With these data the model identifier determines the consequent sub-models by minimizing the sum of squared difference between the observed and computed maximum temperature at each cycle. The resulting model is in good agreement with the observed behaviour and is the first building block towards the required automatic control system.

REFERENCES

Abony, J. (2003). *Fuzzy model identification for control.* Birkhäuser, Boston, 273 pp.

Babuska, R., (1998a). *Fuzzy modeling for control.* Kluver Publ. Co., Amsterdam, 260 pp.

Babuska, R. (1998b). Fuzzy modeling and identification toolbox for use with MATLAB. Lecture notes, available on the net at http://lcewww.et.tudelft.nl/babuska

Bezdek, J.C. (1981). Pattern Recognition with Fuzzy Objective Function Algorithms, Plenum Press, New York, pp. 256.

Chen, S. and Billings, S.A. (1989). Representation of Nonlinear Systems, the NARMAX model. *Int. J. of Control*, **41**, 303 – 344.

Kulcu, R and Yaldiz, O. (2004). Determination of aeration rate and kinetics of composting some agricultural wastes. *Bioresources Technology* **93**, 49–57

Hathaway, R. J. and Bezdek, J. C., (1993). Switching Regression Models and Fuzzy Clustering. *IEEE Trans. on Fuzzy Systems*, **1**, 195-204.

Haug, R.T. (1980). *Compost engineering: principles and practices*. Technomic Publ. Co., Pennsylvania, pp. 655.

Liang, C., Das, K.C., McClendon, R.W., (2003). The influence of temperature and moisture contents regimes on the aerobic microbial activity of a biosolids composting blend. *Bioresource Technology* **86**, 131–137.

Mason, I.G. and Milke, M.W. (2005). Physical modelling of the composting environment: A review. Part 1: Reactor systems. *Waste Management* **25**, 481–500.

Mason, I.G. (2006). Mathematical modelling of the composting process: A review. *Waste Management* **26**, 3–21.

Negro, M.J., Solano, P.C., Carrasco, J., (1999). Composting of sweet sorghum bagasse with other wastes. *Bioresource Technology* **67**, 89–92.

Sole-Mauri F., Illa J., Magrı A., Prenafeta-Boldu F.X. and Flotats X., (2007). An integrated biochemical and physical model for the composting process. *Bioresource Technology* **98**, 3278-3293.

Takagi, T. and Sugeno, M., (1985). Fuzzy Identification of Systems and its Applications to Modeling and Control. *IEEE Trans. on Systems, Man, and Cybernetics*, **15**, 116-132.

Xie, X. and Beni, G. (1991). A validity measure for fuzzy clustering. *IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI)*, **13** (8), 841.847.