



**HAL**  
open science

# Dynamic Mode Decomposition with Control for Data-driven Modeling of Anaerobic Digestion Process

Benaissa Dekhici, Boumediene Benyahia, Brahim Cherki

► **To cite this version:**

Benaissa Dekhici, Boumediene Benyahia, Brahim Cherki. Dynamic Mode Decomposition with Control for Data-driven Modeling of Anaerobic Digestion Process. CARI 2022, Oct 2022, Tunis, Tunisia. hal-03696038

**HAL Id: hal-03696038**

**<https://hal.science/hal-03696038>**

Submitted on 15 Jun 2022

**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.

# Dynamic Mode Decomposition with Control for Data-driven Modeling of Anaerobic Digestion Process

Benaïssa DEKHICI<sup>\*1</sup>, Boumediene BENYAHIA<sup>1</sup>, Brahim CHERKI<sup>1</sup>

<sup>1</sup>Laboratoire d'Automatique de Tlemcen, Université de Tlemcen , 13000, Tlemcen, Algeria ,

\*E-mail : [benaïssa.dekhici@gmail.com](mailto:benaïssa.dekhici@gmail.com)

---

## Abstract

This paper proposes a data-driven modeling approach for complex Anaerobic Digestion (AD) systems. This method is called Dynamic Mode Decomposition with Control (DMDc), which is an emerging equation-free technique for deducing global linear state-space input-output models with actuation for complex systems. DMDc is applied to a set of data generated from simulating the Ordinary Differential Equations (ODEs) of the Anaerobic Model 2 (AM2) using MATLAB. The simulation results demonstrate the prediction accuracy of the linear state-space model generated from the DMDc algorithm.

## Keywords

Anaerobic digestion; DMD ; DMDc; Modeling, System identification.

---

## I INTRODUCTION

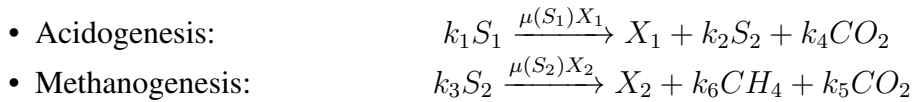
Mathematical modeling and simulation are flexible and economical means to analyze and predict the dynamics of bioprocesses [1]. Anaerobic Digestion (AD) is a biological process through which several microorganisms degrade organic components without the presence of oxygen. Depending on the number of biochemical processes taken into account, there are more or less complex models in the literature of AD processes[2]. The Anaerobic Digestion Model number 1 (ADM1) [3], developed by researchers from the International Water Association (IWA), is the most complete phenomenological model for simulating AD. At its base, it consists of more than 32 state variables and more than 80 parameters. However, this model is not devoted to process control. On the other hand, the Anaerobic Model 2 (AM2) [4] is a low-order model which describes the AD in only two main steps (Acidogenesis-Methanogenesis). Nevertheless, the high nonlinearity and complexity of the description of the dynamics are still present. The aim of this work is to derive a data-driven linear model for the AD process based on the states (substrate and biomass) and control (dilution rate) measurements. The proposed approach is the Dynamic Mode Decomposition with Control (DMDc) algorithm [5]. DMD was first introduced in [6] as a data-driven algorithm for the modeling and order-reduction of the high complex fluid flow. Since then, the DMD is used in a broad range of applications [7] including the dimensionality reduction of ADM1 model [8]. The DMDc technique [5] used in this study is an extension of the DMD algorithm incorporating the impact of actuation to generate global linear state-space input-output models from nonlinear complex systems [9]. In this paper, first, the AM2 model is described, then, the DMDc algorithm is introduced. Simulation results are given by the application of the DMDc algorithm to model the AD process based on the data generated from a virtual

AD system modeled by the AM2 using MATLAB. Finally, some conclusions and perspectives are drawn.

## II ANAEROBIC DIGESTION SYSTEM

The Anaerobic Model 2 (AM2) is a reduced model which represents AD in two stages (Acidogenesis-Methanogenesis). It was developed within the framework of the European Research Project called AMOCO [4] and it is derived from the law of mass balance [1]. This paper first presents the principle of modeling bioprocesses with regards to the law of mass balance of the AM2 model. Then, the mathematical model of AM2 is introduced as set of ODEs. AM2 describes the dynamics of four main state variables: two consortia of bacteria  $X_1$  and  $X_2$  and two types of substrates  $S_1$  and  $S_2$ . In the first stage, the consortium of acidogenic bacteria  $X_1$  degrades the organic substance ( $S_1$ ) into Volatile Fatty Acids (VFA) ( $S_2$ ) and carbon dioxide ( $CO_2$ ), at the reaction rate  $\mu(S_1)X_1$ . In the second stage, the consortium of methanogenic bacteria  $X_2$  transforms  $S_2$  into methane ( $CH_4$ ) and  $CO_2$ , at the reaction rate  $\mu(S_2)X_2$ .

These two steps are represented by the following reaction schemes:



Where  $X_1$ : is the population concentration of acidogenic bacteria ( $g/L$ ).  $X_2$ : is the population concentration of methanogenic bacteria ( $g/L$ ).  $S_1$ : is the substrate concentration of carbonic substance ( $g/L$ ).  $S_2$ : is the substrate concentration of VFA ( $mmol/L$ ).  $\mu(S_1)$ : is the specific growth rate of  $X_1$  on  $S_1$ .  $\mu(S_1)X_1$ : is the Acidogenesis reaction rate.  $\mu(S_2)$ : is the specific growth rate of  $X_2$  on  $S_2$ .  $\mu(S_2)X_2$ : is the Methanogenesis reaction rate.  $k_i$  are the stoichiometric coefficients associated with the two reactions. In Figure 1, a simplified diagram of an anaerobic Continuous Stirred Tank Reactor (CSTR) modeled by AM2 is shown, where we consider a continuous mode of functioning (inflow rate = outflow rate) of our bioreactor. We proceed to

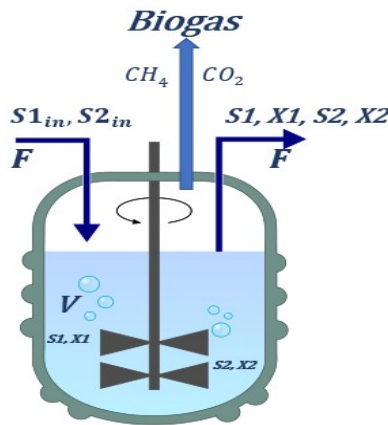


Figure 1: Schematic diagram of a CSTR modeled by AM2 model.

deduce the AM2 model, using the mass balance law. The ODEs of the AM2 model as originally

proposed in [4] are:

$$\begin{cases} \dot{S}_1 = D(S_{1in} - S_1) - k_1\mu_1(S_1)X_1 \\ \dot{X}_1 = (\mu_1(S_1) - D)X_1 \\ \dot{S}_2 = D(S_{2in} - S_2) + k_2\mu_1(S_1)X_1 - k_3\mu_2(S_2)X_2 \\ \dot{X}_2 = (\mu_2(S_2) - D)X_2 \end{cases} \quad (1)$$

Where  $D = \frac{F}{V}$  is the dilution rate considered as the control input ( $F$  is the flow rate and  $V$  is the volume of the vessel),  $S_{1in}$  and  $S_{2in}$  are respectively the input concentrations of the substrates  $S_1$  and  $S_2$ . The kinetics  $\mu_1$  and  $\mu_2$  are Monod[10] and Haldane[11] functions, respectively, given by:

$$\mu_1(S_1) = \mu_{1max} \frac{S_1}{S_1 + K_1}, \quad \mu_2(S_2) = \mu_{2max} \frac{S_2}{\frac{S_2^2}{K_i} + S_2 + K_2} \quad (2)$$

where  $\mu_{1max}$  and  $\mu_{2max}$  are the maximum growth rates ( $\mu_{2max}$  without inhibition).  $K_1$  and  $K_2$  are the half-saturation constant related to the substrates  $S_1$  and  $S_2$ , respectively.  $K_i$  is the inhibition constant associated with the substrate  $S_2$ . The main contribution of this work is to develop a data-driven strategy for the modeling of the CSTR bioreactor modeled with AM2 by using only the data measurements of its states and the control input.

### III METHODS

#### 3.1 Dynamic mode decomposition (DMD)

In this section the DMD procedure is presented briefly. The standard DMD algorithm [6] can be described by assuming the existence of a linear relationship between the recorded snapshots of data, therefore, a companion matrix is adopted to project a linear operator onto a Proper Orthogonal Decomposition (POD) basis obtained by applying a Singular Value Decomposition (SVD) on the snapshot of data matrix. A collection of snapshots of data with  $K$  samples from either simulation or experiments is organised in a table given by  $\{x_1, x_2, x_3, \dots, x_K\}$ , where  $x_i$  is the  $i$ th snapshot (vector of all possible measurements that could be taken for the system) and  $\Delta t$  is the time step between two samples. The DMD procedure can generate the following linear dynamical system

$$x_{i+1} = Ax_i, \quad (3)$$

where  $A$  is a constant coefficient matrix that represents the linear dependency of the system. Therefore, the matrix  $A$  contains on its eigenvalues the dynamical characteristics of the nonlinear system. If the system is high-dimensional, the matrix  $A$  can be reduced in order to deduce the leading eigenvalues of the linear operator  $A$ . The DMD procedure has, as an input, the following matrices:

$$X_1 = \begin{bmatrix} | & | & | & \dots & | \\ x_1 & x_2 & x_3 & \dots & x_{K-1} \\ | & | & | & \dots & | \end{bmatrix} \quad (4)$$

$$X_2 = \begin{bmatrix} | & | & | & \dots & | \\ x_2 & x_3 & x_4 & \dots & x_K \\ | & | & | & \dots & | \end{bmatrix} \quad (5)$$

The system in (3) can be rewritten as follows:

$$X_2 \approx AX_1, \quad (6)$$

where  $A$  is the best fit linear operator that advances  $X_1$  to  $X_2$  one  $\Delta t$  in the future. The main aim of DMD method is to provide the DMD modes along with the leading eigenvalues of the underlying system included in the matrix  $A$ . A reduced matrix  $\tilde{A}$  is obtained if the DMD algorithm is attained by a similarity transformation of the system matrix. In addition, the matrix  $\tilde{A}$  is built to substitute the tall and skinny matrix  $A$  of the high-dimensional system. First, a SVD is performed on the snapshot of data matrix  $X_1$

$$X_1 = U\Sigma V^*, \quad (7)$$

$$A = U^* \tilde{A} U, \quad (8)$$

where  $\Sigma$  is a diagonal singular values matrix and  $U, V$  are the POD modes and the right singular vectors respectively. The matrix  $A$  can be calculated as the solution of the following minimization problem :

$$\|X_2 - AX_1\|_F \quad (9)$$

Where  $\|\cdot\|_F$  is the Frobenius norm given the difference between  $X_2$  and  $AX_1$ .

We can also deduce from (7), (8) and (9) the following

$$\|X_2 - U\tilde{A}\Sigma V^*\|_F \quad (10)$$

where  $A$  can be replaced by

$$\tilde{A} = U^* X_2 V \Sigma^{-1}, \quad (11)$$

where  $\tilde{A}$  contains the leading eigenvalues of  $A$ . The DMD modes  $\Phi$  then can be calculated with the following equation:

$$\Phi = X_2 V \Sigma^{-1} W, \quad (12)$$

where  $W$  is the matrix of eigenvectors of the reduced matrix  $\tilde{A}$ . We should stress that in our case we are only using the full matrix  $A$  since our system is already low-dimensional and has just four states to measure.

### 3.2 Dynamic mode decomposition with control (DMDc)

DMDc is a data-driven method used for the identification and modeling of high-dimensional complex dynamical systems. It was first introduced in [5] where authors incorporate the influence of the control to generate models of complex dynamical systems from data. DMDc is an extension of the DMD method where the effect of the external forcing is included in the classical DMD algorithm. DMDc algorithm only requires the snapshot data of the state along with the actuation measurements.

A controlled linear dynamical system can be described as follows:

$$x_{i+1} = Ax_i + Bu_i, \quad (13)$$

where  $x$  is the state of the system,  $u$  is the control input and  $A, B$  are the system and control matrices, respectively. The data snapshot matrices of the measurements of the system (13) are given as in (4) and (5) while a data snapshot matrix of the control input is given as follows:

$$Y_0 = \begin{bmatrix} | & | & | & \dots & | \\ u_1 & u_2 & u_3 & \dots & u_{K-1} \\ | & | & | & \dots & | \end{bmatrix} \quad (14)$$

The system (13) can then be rewritten as

$$X_2 \approx AX_1 + BY_0, \quad (15)$$

where the matrices  $A, B$  should be identified together by the DMDc algorithm. We can reformulate the relationship between the data matrices  $X_1, X_2$  and  $Y_0$  as follows:

$$X_2 \approx [A \ B] \begin{bmatrix} X_1 \\ Y_0 \end{bmatrix} = G\Omega \quad (16)$$

where  $G = [A \ B]$  is the augmented operator matrix, and  $\Omega = \begin{bmatrix} X_1 \\ Y_0 \end{bmatrix}$  is the augmented data matrix. By following the same standard DMD procedure, the matrix  $G$  can be calculated as

$$G = X_2\Omega^\dagger \quad (17)$$

$$[A \ B] = X_2 \begin{bmatrix} X_1 \\ Y_0 \end{bmatrix}^\dagger \quad (18)$$

where  $\dagger$  denotes the Moore-Penrose pseudo-inverse. We should stress that the matrix  $\Omega$  contains all the information about the state and control measurements given in the snapshot of data. The best fit of the matrix  $G$  containing the process dynamics  $A$  and the control input  $B$  is given by minimizing the following Frobenius norm :

$$\|X_2 - G\Omega\|_F \quad (19)$$

For computing  $G$ , the SVD is performed on both  $\Omega$  and  $X_2$  matrices. The DMDc algorithm is shown below:

1. The data matrices  $X_2, X_1$  and  $Y_0$  are collected and constructed. Then, we construct the augmented data matrix  $\Omega$  by stacking the matrices  $X_1, Y_0$  together.
2. Perform the SVD on the augmented data matrix  $\Omega$ , resulting the decomposition  $\Omega \approx \tilde{U}\tilde{\Sigma}\tilde{V}^*$ , where  $\tilde{U}$  can be written as  $\tilde{U} = \begin{bmatrix} \tilde{U}_1 \\ \tilde{U}_2 \end{bmatrix}$  (separating the space of the state with the space of the input).
3. Compute another SVD on the output space  $X_2$ , thereby resulting the decomposition  $X_2 \approx \hat{U}\hat{\Sigma}\hat{V}^*$ .
4. Compute the approximate of the augmented operator matrix  $G = [A \ B] = [\hat{U}\tilde{A}\hat{U}^* \ \hat{U}\tilde{B}]$ . It can be also rewritten as follows:

$$\tilde{A} = \hat{U}^* X_2 \tilde{V} \tilde{\Sigma}^{-1} \tilde{U}_1^* \hat{U} \quad (20)$$

$$\tilde{B} = \hat{U}^* X_2 \tilde{V} \tilde{\Sigma}^{-1} \tilde{U}_2^* \quad (21)$$

5. Apply the eigen-decomposition of the matrix  $\tilde{A}$ , given by  $\tilde{A}W = W\lambda$ .
6. Deduce the DMD modes of the operator  $A$ , given by  $\Phi = X_2 \tilde{V} \tilde{\Sigma}^{-1} \tilde{U}_1^* \hat{U} W$ .

After the identification of the matrices  $A$  and  $B$ , the real system can be reconstructed and predicted by the linear dynamical model in (13). In this work, we apply the native DMDc algorithm to our system that has only four states and one possible control input to measure.

## IV SIMULATION RESULTS AND DISCUSSION

In this section, the DMDc method proposed in this study is applied for modeling the AD process simulated by the CSTR reactor modeled with AM2. First, we consider the AM2 model in (1) as a virtual system generating the snapshots of data as in (4),(5) and (14). Then, the data matrices are used as an input for the DMDc algorithm presented in Section III. We run the simulation for a period of 120 days with a sample step of one day and applying a variable control input  $u = D$  ( $D$  is the dilution rate considered as the only control input of the CSTR reactor). The snapshots of data matrices used for the data-driven modeling of the AD system are given below:

$$X_1 = \begin{bmatrix} S_1^1 & S_1^2 & , \dots , & S_1^{120-1} \\ X_1^1 & X_1^2 & , \dots , & X_1^{120-1} \\ S_2^1 & S_2^2 & , \dots , & S_2^{120-1} \\ X_2^1 & X_2^2 & , \dots , & X_2^{120-1} \end{bmatrix}, X_2 = \begin{bmatrix} S_1^2 & S_1^3 & , \dots , & S_1^{120} \\ X_1^2 & X_1^3 & , \dots , & X_1^{120} \\ S_2^2 & S_2^3 & , \dots , & S_2^{120} \\ X_2^2 & X_2^3 & , \dots , & X_2^{120} \end{bmatrix} \quad (22)$$

$$Y_0 = [D_1 \ D_2 \ , \dots , \ D_{120-1}] \quad (23)$$

where the first column of  $X_1$  is the states measurement of the AM2 model (1) at day one generated by the control input ( $D_1$ ) applied at the same day. In addition,  $X_2$  is the same as  $X_1$  but shifted one  $\Delta t$  ( $\Delta t = 1day$ ) in the future. Parameters used in the simulation of the AM2 model (1) are described in Table 1.

Parameter	Value	Unit
$S_{1in}$	20	$g/L$
$S_{2in}$	150	$mmol/g$
$\mu_{1max}$	0.5	$d^{-1}$
$\mu_{2max}$	0.74	$d^{-1}$
$k_1$	42.14	—
$k_2$	116.5	$mmol/g$
$k_3$	268	$mmol/g$
$K_1$	7.1	$g/L$
$K_2$	9.28	$mmol/g$
$K_i$	80	$mmol/g$

Table 1: Parameters used in the simulation of the AM2 model (1) using MATLAB as in [4].

Simulation results are shown in Figure 2, where we can remark a good fit from the *DMDc global linear state-space input-output model* to the real data of the AD system. In Figure 2, we have applied a variable control input in order to derive the DMDc model; the dilution rate was a bit tough as we wanted to force the system to be unstable (starting from a low dilution rate and suddenly switching to a higher value) and see if the DMDc linear model can respond accurately. We observe that the DMDc model (the matrices  $A$  and  $B$  of the DMDc linear model are given below with (24)) is following exactly the real system and is tracking any sudden changes in the behavior of the AM2 dynamics with respect to the variable control input. Importantly, the DMDc global linear model fits exactly the real system with impressive accuracy. We should emphasize that DMDc method is totally data-driven and only input-output data is needed for

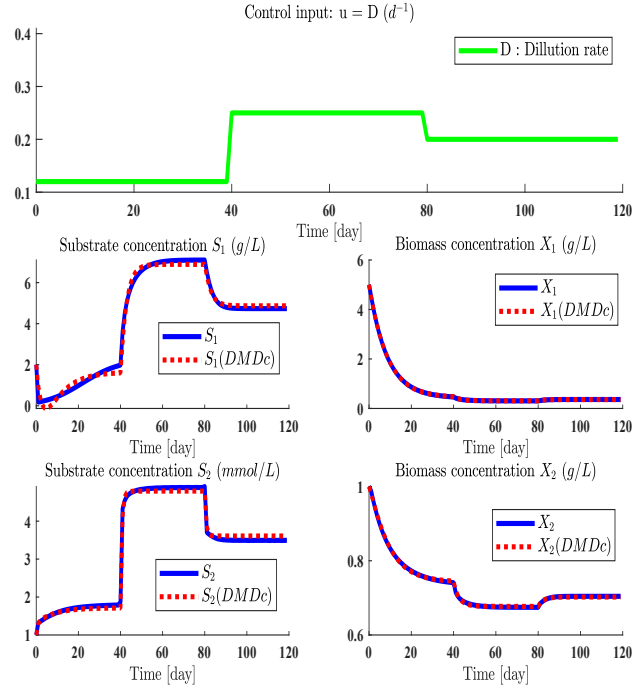


Figure 2: DMDc applied on the AM2 model for modeling and identification by applying a variable dilution rate.

the construction of a global linear state-space model that can be used easily for tasks like state-estimation and control.

$$A = \begin{bmatrix} 0.909 & -0.446 & -0.017 & -1.011 \\ -0.003 & 0.669 & 0.001 & 0.235 \\ -0.065 & 0.029 & 0.508 & -3.866 \\ 0 & -0.014 & 0.002 & 1.036 \end{bmatrix}, B = \begin{bmatrix} 6.357 \\ -0.15 \\ 22.849 \\ -0.15 \end{bmatrix} \quad (24)$$

## V CONCLUSION

In this work, we have presented a data-driven strategy for modeling the nonlinear complex AD system. The DMDc algorithm used the data from simulating the AM2 model in MATLAB. A global linear state-space input-output model was obtained for the forecast of the controlled AD process. Simulation results have shown the accuracy of the DMDc model in predicting the AM2 dynamics even when a variable persistent control input is applied. Future work will focus on the use of the DMDc method for data-driven control of the AD systems.

## REFERENCES

- [1] J. Harmand, C. Lobry, A. Rapaport, and T. Sari. *The Chemostat: Mathematical Theory of Microorganism Cultures*. Wiley, 2017. ISBN: 978-1-119-43712-3.
- [2] Bastin and G. Dochain. *On-line estimation and adaptive control of bioreactors*. Volume 1. Process Measurement and Control. Elsevier, 1990.
- [3] D. Batstone, J. Keller, I. Angelidaki, S. Kalyuzhnyi, S. Pavlostathis, A. Rozzi, W. Sanders, H. Siegrist, and V. Vavilin. “The IWA Anaerobic Digestion Model No 1 (ADM1)”. In: *Water Sci. Technol.* 45.10 (2002), pages 65–73.



- [4] O. Bernard, Z. Hadj-Sadok, D. Dochain, A. Genovesi, and J.-P. Steyer. “Dynamical model development and parameter identification for an anaerobic wastewater treatment process”. In: *Biotechnology and bioengineering* 75.4 (2001), pages 424–438.
- [5] J. L. Proctor, S. L. Brunton, and J. N. Kutz. “Dynamic Mode Decomposition with Control”. In: *SIAM Journal on Applied Dynamical Systems* 15.1 (2016), pages 142–161.
- [6] P. J. Schmid. “Application of the dynamic mode decomposition to experimental data”. In: *Exp. Fluids* 50.4 (2011), pages 1123–1130.
- [7] J. N. Kutz, S. L. Brunton, B. W. Brunton, and J. L. Proctor. *Dynamic mode decomposition: data-driven modeling of complex systems*. SIAM, 2016.
- [8] B. Dekhici, B. Benyahia, and C. Brahim. “Model order reduction using Dynamic Mode Decomposition: Application to the Anaerobic Digestion Model Number 1(ADM1)”. In: *CARI 2020 - 15th African Conference on Research in Computer Science and Applied Mathematics*. Thiès, Senegal, Oct. 2020.
- [9] S. L. Brunton, J. L. Proctor, and J. N. Kutz. “Discovering governing equations from data by sparse identification of nonlinear dynamical systems”. In: *Proceedings of the national academy of sciences* 113.15 (2016), pages 3932–3937.
- [10] J. Monod. “La technique de culture continue theorie et application”. In: *Ann. Inst.Pasteur* 79 (1950), pages 390–410.
- [11] J. B. S. Haldane. *Enzymes*. Longmans, London, 1930, page 80.