Documentation on Anaerobic Digestion models

1 Modification to original AM2 description

The original description of AM2 defines the concentration of CO_2 as $CO_2 = C + S_2 - Z$ [equation 19 from Bernard et al., 2001]. This is implied from equations $Z \simeq B + S_2$ and $C = CO_2 + B$ (respectively equations 10 and 3).

Since accurate pH measurements are assumed to be available, an alternative formulation, $CO_2 = C/(1 + 10^{pH-pKb})$, was implemented. It is a straightforward consequence of equations $K_b = \frac{[H^+]B}{CO_2}$ and $C = CO_2 + B$ (respectively equations 5 and 3).

The growth rate formulas [equations 33 and 34 in Bernard et al., 2001] were also modified to account for mortality, as described in Hassam et al. [2015]. Equations 6 and 7 from the latter source were implemented.

Code for AM2 is provided in 'aduq' library ('run_AM2' in 'aduq.pyam2' module).

2 Reparametrisation of AD models

For order constrained parameters, the lower and upper values parametrisation is replaced by the lower value and the difference between upper and lower value. As such, the only constraint on the parameters is that they should be positive, and log gaussian distributions can be safely used. For the AD models considered here, only parameters specifying pH inhibition for ADM1 are ordered. The new parameters are denoted pH_{LL} and $pH_{UL:LL} := pH_{UL} - pH_{LL}$.

3 Construction of the priors

For both AD models, the prior π_0 is constructed on the log of the parameters using Gaussian distributions with a diagonal covariance matrix. As such, the prior draws parameters independently from one another.

For ADM1, the set of default parameter values specified in Rosén and Jeppsson [2006] is used to define, after applying the log transform, the means of the univariate gaussian distributions. The standard deviations are inferred from the uncertainty specified in section 6 of Batstone et al. [2002]. The degree of variability specified is interpreted as a standard deviation in log space, using the equivalences specified in the

publication. Since these measure variability in terms of intervals, these intervals are transformed into standard deviation by applying a corrective factor of .4, which ensure that approximatively 99% of the draws are contained in the specified interval. For pH inhibition related parameters, a similar procedure is used, considering the logarithms of the difference between upper and lower limit, and using the higher degree of variability between upper and lower limit to obtain the prior.

For AM2, the prior is centered around the set of default parameter values specified in Bernard et al. [2001], and the standard deviations are inferred from the same source, using the same .4 corrective factor.

The default values and standard deviations used in the prior distribution are tabulated in Table 2a for ADM1 and Table 2b for AM2.

4 Construction of the intrant

280 days in total were modelled, with four feed information per day (i.e. a description of the influent every 6 hours). Of these, 196 days (70%) are used during the calibration phase.

For ADM1, the substrate is supposed to contain only carbohydrates, proteins and lipids. The concentrations of the remaining quantities are set to 0. Each of three concentration time series in the intrant is constructed independently using sums of 200 random sinusoids. The amplitudes and phase are drawn using uniform distributions between 0 and 1 and 0 and 2π respectively, while the frequencies are drawn using uniform distributions between $f_{\min}$ and $f_{\max}$. The sum is then renormalised to have a given mean and an amplitude (measured as the maximum difference between the maximum or minimum and the mean) of Amp times the mean value.

The feed rate is constructed in a like manner, with only a slight change to the renormalisation procedure, which is implemented in such a way that the maximum and minimum values $Q_{\max}, Q_{\min}$ are known. For the dataset exhibiting distribution shift, the feeding rate of the L dataset is multiplied by

$$c_t = 1 + \left(\frac{\tanh\left(\frac{t-215}{10}\right) + 1}{2} \right)^6 \frac{Q_{\max,H} - Q_{\max,L}}{Q_{\max,L}}.$$

This ensures that the maximum feeding rate for the test set remains below $Q_{\max,H}$. The maximum and minimum feed rates are inferred from minimum and maximum HRT.

For AM2, the intrant description is translated from a ADM1 intrant description generated in the manner described above, using the conversion methodology proposed by Hassam et al. [2015]. The pH of the influent, necessary for the conversion, is set to 6.0.

The values of the hyperparameters used to generate L and H datasets are tabulated in Table 1.

Table 1: Influent dataset construction hyperparameters

	X_{ch}	X_{pr}	X_{li}	$f_{\max}$	$f_{\min}$	Amp	$\mathrm{HRT}_{\min}$	$\mathrm{HRT}_{\min}/\mathrm{HRT}_{\max}$
L	10	20	3	0.01	0.002	0.2	30	0.9
H	10	20	3	0.02	0.002	0.4	10	0.7

The concentrations of carbohydrates, proteins and lipids (X_{ch}, X_{pr}, X_{li}) are in kgCOD m^{-3}. The frequencies $f_{\max}$ and $f_{\min}$ are in Day^{-1} and the HRT in Day.

References

D. J. Batstone, J. Keller, I. Angelidaki, S. V. Kalyuzhnyi, S. G. Pavlostathis, A. Rozzi, W. T. M. Sanders, H. Siegrist, V. A. Vavilin, The IWA Anaerobic Digestion Model No 1 (ADM1), Water Science and Technology 45 (2002) 65–73. URL: https://doi.org/10.2166/wst.2002.0292.

O. Bernard, Z. Hadj-Sadok, D. Dochain, A. Genovesi, J.-P. Steyer, Dynamical model development and parameter identification for an anaerobic wastewater treatment process, Biotechnology and Bioengineering 75 (2001) 424–438. URL: https://doi.org/10.1002/bit.10036.

S. Hassam, E. Ficara, A. Leva, J. Harmand, A generic and systematic procedure to derive a simplified model from the anaerobic digestion model no. 1 (ADM1), Biochemical Engineering Journal 99 (2015) 193–203. URL: https://doi.org/10.1016/j.bej.2015.03.007.

C. Rosén, U. Jeppsson, Aspects on ADM1 implementation within the BSM2 framework, Department of Industrial Electrical Engineering and Automation, Lund University, Lund, Sweden (2006) 1–35.

Table 2: Default parameter values (θ_0), uncertainty ($\sigma_{\log\theta}$) and generated parameters (θ_N, θ_F) for ADM1 and AM2 models

(a) ADM1

	θ_0	$\sigma_{\log\theta}$	θ_N	θ_F
k_{dis} (Day^{-1})	0.5	0.24	0.35	0.23
k_{hyd_CH} (Day^{-1})	10	0.12	10.0	5.2
k_{hyd_PR} (Day^{-1})	10	0.12	8.6	11.9
k_{hyd_LI} (Day^{-1})	10	0.24	23.4	2.4
k_{m_su} (Day^{-1})	30	0.12	37.7	47.7
k_{m_aa} (Day^{-1})	50	0.12	44.6	48.7
k_{m_fa} (Day^{-1})	6	0.24	5.2	4.47
k_{m_c4+} (Day^{-1})	20	0.12	21.2	17.6
k_{m_pro} (Day^{-1})	13	0.12	12.7	9.9
k_{m_ac} (Day^{-1})	8	0.12	10.7	10.2
k_{m_h2} (Day^{-1})	35	0.12	27.9	28.8
k_{dec} (Day^{-1})	0.02	0.12	0.0157	0.0145
K_{S_IN} (kMole m^{-3})	1e-4	0.046	9.3e-5	9.46e-5
K_{S_su} (kgCoD m^{-3})	0.5	0.12	0.427	0.29
K_{S_aa} (kgCoD m^{-3})	0.3	0.046	0.31	0.29
K_{S_fa} (kgCoD m^{-3})	0.4	0.24	0.389	0.58
K_{S_c4} (kgCoD m^{-3})	0.2	0.24	0.300	0.399
K_{S_pro} (kgCoD m^{-3})	0.1	0.12	0.115	0.065
K_{S_ac} (kgCoD m^{-3})	0.15	0.12	0.147	0.24
K_{S_H2} (kgCoD m^{-3})	7e-6	0.12	6.6e-6	6.1e-6
$K_{I,H2\_fa}$ (kgCoD m^{-3})	5e-6	0.046	5.01e-6	5.603e-6
$K_{I,H2\_c4+}$ (kgCoD m^{-3})	1e-5	0.046	1.06e-5	1.1e-5
$K_{I,H2\_pro}$ (kgCoD m^{-3})	3.5e-6	0.046	3.15e-6	3.4e-6
$K_{I,NH3}$ (kMole m^{-3})	1.8e-3	0.046	1.88e-3	2.1e-3
$pH_{UL:LL\ aa}$	1.5	0.12	1.16	1.36
$pH_{LL\ aa}$	4.0	0.12	5.11	4.4
$pH_{UL:LL\ ac}$	1.0	0.046	1.02	1.0
$pH_{LL\ ac}$	5.0	0.046	4.53	4.8
$pH_{UL:LL\ h2}$	1.0	0.12	0.52	1.6
$pH_{LL\ h2}$	5.0	0.046	5.5	6.1

(b) AM2

	θ_0	$\sigma_{\log\theta}$	θ_N	θ_F
μ_{1max} (Day^{-1})	1.2	0.2	1.23	0.99
μ_{2max} (Day^{-1})	0.74	0.48	1.05	1.32
K_{S1} (gCOD L^{-1})	7.1	0.32	7.51	15.98
K_{S2} (mmol L^{-1})	9.28	0.48	11.0	4.97
K_{I2} (mmol L^{-1})	256	0.4	196	345