INTERNATIONAL JOURNAL OF HYDROGEN ENERGY 41 (2016) 12929–12940

Available online at www.sciencedirect.com

ScienceDirect

journal homepage: www.elsevier.com/locate/he

Intelligent models to predict hydrogen yield in dark microbial fermentations using existing knowledge

Yeshona Sewsynker, Evariste Bosco Gueguim Kana *

Discipline of Microbiology, School of Life Sciences, University of KwaZulu-Natal, Private Bag X01, Scottsville, 3209, Pietermaritzburg, South Africa

ARTICLE INFO

Article history:
Received 1 February 2016
Received in revised form
26 April 2016
Accepted 28 May 2016
Available online 16 June 2016

Keywords:
Modeling
Biohydrogen production
Artificial Neural Network
Sensitivity analysis

ABSTRACT

Bioprocess development for hydrogen production requires an excellent understanding of the influence of key operational parameters on hydrogen yields at early stages of the innovation chain. Knowledge on the impact of these key input parameters on fermentative hydrogen yields exist in the public domain. This study builds on this knowledge to implement intelligent models that could predict the hydrogen response on new physicochemical input values. Two Artificial Neural Network (ANN) models for hydrogen production were implemented and assessed using published data from 64 selected studies. For both models the multilayer perceptron (MLP) class of neural network was used with a topology of 5-7-7-1, corresponding to the number of neurons of input, hidden (two) and output layers. The input variables consisted of inoculum type, substrate type, substrate concentration, pH and temperature. The output was the hydrogen yield expressed as mole of hydrogen per mole of substrate (Mol_Model) or as cumulative volume of hydrogen per gram substrate (Vol_Model). A high coefficient of determination (R^2) was obtained for Vol_Model (0.90) whereas a low value was observed with Mol_Model (0.46). These findings showed that the Vol_Model efficiently abstracted the non-linear relationship between the considered inputs and biohydrogen yield with a higher prediction accuracy on new physicochemical parameters. Thus, these ANN derived models could be used to navigate the optimization space and shorten the biohydrogen process development time.

Introduction

Fossil fuel depletion and the steep increase in greenhouse gas emissions have driven research towards renewable energy methods. Despite the limitations of fossil fuel resources, several factors influence its continuity for future use such as existing infrastructure and the current low price of crude oil. However the negative environmental impact associated to its consumption has prompted research towards renewable and sustainable biofuels [1].

Biohydrogen has proven to be an excellent potential alternative to fossil fuel sources, due to its high gravimetric energy density of 122 kJ/g, which is approximately 2.9 times higher than conventional fossil fuels [2]. Additionally, the combustion of hydrogen results in water as the only by-product [1]. Various biological methods exist for hydrogen production and include: photo-fermentation, dark fermentation and microbial electrolysis. Microbial production of bio-hydrogen via dark fermentation entails the use of microorganisms for the degradation of organic substrates

* *Corresponding author*. Tel.: +27 332605527.
E-mail address: kanag@ukzn.ac.za (E.B. Gueguim Kana).

http://dx.doi.org/10.1016/j.ijhydene.2016.05.250
0360-3199/

under anaerobic conditions [1]. Dark fermentation has shown to generate superior hydrogen production rates compared to other processes in terms of energy efficiency, and can exploit a wide array of renewable organic matter. It proceeds via the acetate (4 mol H_2/mol hexose) or butyrate pathway (2 mol H_2/mol hexose) as shown in Equations (1) and (2), respectively, however; practical yields do not reach theoretical values due to metabolic limitations [2].

$$C_6H_{12}O_6 + 2H_2O \rightarrow 2CH_3COOH + 2CO_2 + 4H_2 \quad (1)$$

$$C_6H_{12}O_6 \rightarrow CH_3CH_2CH_2COOH + 2CO_2 + 2H_2 \quad (2)$$

Bioprocess development and scale up requires modeling and optimization of the key parameters that impact the output at the initial stages of process development [3]. A bioprocess model provides insight into the individual as well as the interactive effects of the various input parameters on the corresponding output. Nevertheless, the non-linearities associated with microbial fermentations increase the complexity in model development. Non-linear systems as opposed to linear systems are not standardized thus resulting in deviations between results obtained. The implementation of bioprocess models that efficiently encapsulate these non-linearities are of paramount importance for optimization and scale up of the process [3,4]. Some studies have attempted to provide models that relate these physico-chemical input parameters to the corresponding hydrogen yield [3–6].

Several factors have shown to impact the hydrogen production process and include: inoculum type, substrate type and concentration, temperature, and pH [3–7]. These factors affect the microbial community composition, impact the metabolic fluxes and ultimately the amount of hydrogen produced in the system, thus selecting the metabolic pathway for biohydrogen production [7]. A slight change from the optimum set point may have a significant impact on the process yield [7–18].

Studies have revealed that pH values below 4.5 inhibit the hydrogenase activity and thus will influence the overall yield [8]. Both pure and mixed cultures may be used for hydrogen production. The latter are cheaper to operate, simpler to control at large scale without contamination and have a broader choice of substrate [2].

With regards to substrate type, pure glucose has been mostly used for biohydrogen research [3]. It is easily metabolized by most microorganisms. However, the availability and costs associated with glucose substrate have restricted its potential use for biofuel research. Alternatively, sucrose and xylose have been used to a lesser extent [9,11]. Sucrose, a disaccharide is more resistant for microbes to degrade; however, the more versatile the culture, as in the case of a mixed consortium, the less challenging it becomes. Synergistic interactions between microbial communities permit simultaneous carbohydrate degradation and biohydrogen production using complex substrates. This is advantageous when considering substrates such as lignocellulosic biomass that is mainly comprised of xylose, lignin and cellulose. Reports on pure xylose as a substrate are scarce since it is commonly accessible from waste plant matter that may be pretreated for the fermentation process.

Optimum substrate concentration has been reported within the range of 10–30 g/L [3,6,9,11,19–21]. Wang and Wan [3] reported a maximum yield of 305.3 mL H_2/g glucose. This result was consistent with Wang and Wan [6]. Contrariwise, Mu et al. [9] obtained a maximum yield of 252 mL H_2/g sucrose. Significant variations exist between the reported optimum set points of input parameters for fermentative hydrogen production [3,6,9,11,19–22].

Different bioprocess modeling algorithms have been employed in biohydrogen research. These include the Response Surface Methodology (RSM), fractional factorial design and Artificial Neural Networks (ANN) [3,4,6,23,24]. Sekoai and Gueguim Kana [4] used the RSM to model the effect of substrate concentration, pH, temperature and hydraulic retention time on the hydrogen production process and indicated that this model was able to adequately relate the inputs to the hydrogen output. Likewise, Venkata-Mohan et al. [23] modeled the effect of inoculum type and pretreatment, inlet pH and feed composition on the hydrogen production and substrate degradation efficiency using a fractional factorial design (Taguchi method). Results showed that the developed model was able to determine the optimum conditions for hydrogen production and substrate degradation.

ANNs are described as mathematical representations of the neurological functioning of the human brain. They imitate the brain's learning process by mathematically modeling the network structure of interconnected nerve cells and can be used for bioprocess model development without prior knowledge of the kinetics of metabolic fluxes within the cell and the cultural environment. The effectiveness of ANN in bioprocess development has been reported in various studies [3,5,6,22,24–26]. The ability of ANN models to accurately capture the non-linear relationships in hydrogen fermentation processes were illustrated by the high correlation between the observed and predicted data in the above-mentioned studies.

For instance, Prakasham et al. [22] developed an ANN model on hydrogen production with inputs of pH, glucose to xylose ratio, inoculum size and inoculum age. Whiteman and Gueguim Kana [5] implemented an ANN model of the effect of temperature, initial pH, substrate concentration and inoculum size on hydrogen yield. Both models showed a high level of correlation between the predicted and observed with R^2 values above 0.90 [4,22]. Similarly, Nikhil et al. [25] investigated the influence of hydraulic retention time (HRT), recycle ratio, sucrose concentration and degradation, biomass concentrations, pH, alkalinity, oxidation-reduction potential (ORP), acid and alcohol concentrations on hydrogen production rate and acquired a coefficient of determination of 0.90. These models were implemented with data sample sizes below 50, as large numbers of bioprocess experimentations are costly and time consuming. The predictive accuracy of ANN may be enhanced with an increase in data size [27]. With the exception of the study by Nasr et al. [24], the application of ANN on biofuel bioprocess modeling with a data set beyond 30 has been scantily reported [26].

Biohydrogen yields have been typically expressed using the specific hydrogen production potential (mL H_2/g substrate)