Journal of Power Sources xxx (2017) 1–11

Contents lists available at ScienceDirect

Journal of Power Sources

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

Ensemble engineering and statistical modeling for parameter calibration towards optimal design of microbial fuel cells

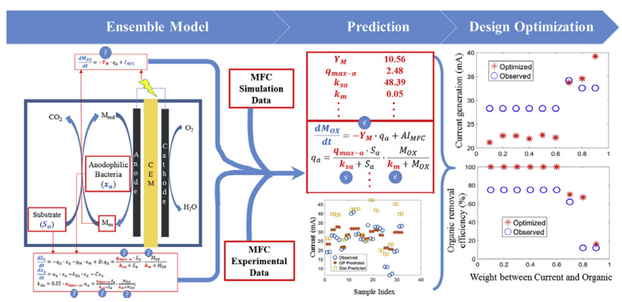
Hongyue Sun [a, 1], Shuai Luo [b, 1], Ran Jin [a, **], Zhen He [b, *]

[a] Grado Department of Industrial and Systems Engineering, Virginia Polytechnic Institute and State University, Blacksburg, VA, 24061, USA
[b] Department of Civil and Environmental Engineering, Virginia Polytechnic Institute and State University, Blacksburg, VA, 24061, USA

HIGHLIGHTS

- The ensemble model for MFC integrates engineering and statistical principles.
- The ensemble model can reduce laboring effort in parameter calibration.
- Comparable accuracy to the traditional statistical model is achieved.
- Input factors can be optimized for best current generation and organic removal.

GRAPHICAL ABSTRACT



ARTICLE INFO

Article history:
Received 22 November 2016
Received in revised form
12 February 2017
Accepted 13 February 2017
Available online xxx

Keywords:
Bioenergy
Ensemble model
Extrapolation
Microbial fuel cell
Optimal design
Wastewater treatment

ABSTRACT

Mathematical modeling is an important tool to investigate the performance of microbial fuel cell (MFC) towards its optimized design. To overcome the shortcoming of traditional MFC models, an ensemble model is developed through integrating both engineering model and statistical analytics for the extrapolation scenarios in this study. Such an ensemble model can reduce laboring effort in parameter calibration and require fewer measurement data to achieve comparable accuracy to traditional statistical model under both the normal and extreme operation regions. Based on different weight between current generation and organic removal efficiency, the ensemble model can give recommended input factor settings to achieve the best current generation and organic removal efficiency. The model predicts a set of optimal design factors for the present tubular MFCs including the anode flow rate of 3.47 mL min^{-1}, organic concentration of 0.71 g L^{-1}, and catholyte pumping flow rate of 14.74 mL min^{-1} to achieve the peak current at 39.2 mA. To maintain 100% organic removal efficiency, the anode flow rate and organic concentration should be controlled lower than 1.04 mL min^{-1} and 0.22 g L^{-1}, respectively. The developed ensemble model can be potentially modified to model other types of MFCs or bioelectrochemical systems.

1. Introduction

Recovering useful energy from wastewater is important to achieve environmental sustainability. Microbial fuel cells (MFCs) have become a promising treatment technology that uses microorganisms as biological catalysts to convert the organics in

* Corresponding author.
** Corresponding author.
E-mail addresses: jran5@vt.edu (R. Jin), zhenhe@vt.edu (Z. He).
[1] Those authors contributed equally to this work.

http://dx.doi.org/10.1016/j.jpowsour.2017.02.051

wastewater directly to electricity [1,2]. Compared with the widely used anaerobic digester, MFCs can be operated at a relatively low temperature with less sludge production and lower energy demand [3]. Tremendous efforts have been made to improve MFC input factor settings for maximizing electricity generation and wastewater treatment efficiency [4]. A major challenge for MFC development is system scaling up [5], and there is a strong need for demonstrating the feasibility of this technology at a scale significantly larger than laboratory systems. At this moment, the design of MFCs is highly empirical, and the MFC performance is greatly affected by the differences in materials, design and operation. Enlarging MFC reactors lacks systematical guidance and optimization. Therefore, how to link bench scale MFCs to large system development and use the results obtained from bench studies to provide implications or guidance to design large scale MFCs becomes of strong interest. In this aspect, mathematical modeling could play a critical role.

Mathematical modeling can be categorized into engineering modeling and statistical modeling, according to different mathematical principles [6]. Engineering modeling for MFCs is based on the first principle governing the chemical and biological processes, which will remain unchanged under different conditions, except that some calibration parameters are sensitive to operation conditions [6]. Engineering modeling is capable of simulating MFC performance under various input factor settings [7–9]. However, due to the assumptions in modeling and numerical computation, engineering modeling may not accurately predict the MFC performance variables in real cases [6,10,11]. Moreover, many calibration parameters in engineering modeling need to be determined, and this will involve time-consuming experiments and computation for parameter calibration [10,11]. Statistical modeling, on the other hand, can be built based on data set from MFCs [6,12], and usually has higher accuracy when the prediction is made in the neighborhood of the measurement data sets [11]. Some statistical techniques such as principal component analysis (PCA) is helpful to identify the major factors affecting MFC performance [13]. However, in extrapolation scenarios, the model estimated based on a data set from one input factor space is used for prediction at another input factor space, where input factor settings deviates from original input factor space of the training data set [12,14,15]. Such an extrapolation may be needed to explore a large input variable space to optimize the design of MFCs, but the extrapolation usually has a large prediction uncertainty (e.g., wide prediction intervals). To solve this problem, Bayesian methods are usually adopted to preserve information from historical measurement data sets, and update the model for extrapolation scenarios [16]. For an accurate model update, a large data set from the new scenarios will be required, but the large data set is usually hard to obtain.

Because engineering modeling can capture the first principles, while statistical modeling can reflect the actual measurements, we envision that the modeling prediction performance for extrapolation can be improved by integrating both the engineering model and data analytics into one model. An ensemble modeling approach is proposed for the model integration [17]. The "ensemble modeling" refers to modeling with different data sources, such as the simulation data from engineering model and measurement data from MFC system. In particular, a Gaussian Process (GP) model can serve as a surrogate model to represent the engineering model, and also be used to capture the discrepancy between the surrogate model and measurement data [18–20]. Such an approach will facilitate the engineering model parameter calibration and improve the model performance in extrapolation scenarios [19,21,22]. In addition, the number of samples required to construct an accurate model is significantly reduced compared with a statistical model which reaches a comparable prediction accuracy. This is important as the smaller sample size will result in less cost and efforts in experiments for MFC design and optimization. The traditional GP models have been developed for a single output variable [19,21], but are not capable of modeling a system with multiple output variables. Several works were performed to construct surrogate models based on engineering model [23,24], but the joint modeling with an engineering model and measurement data was rare. To solve the problem, a framework was proposed for multiple output variables, and focused on model estimation and uncertainty analysis [20].

In this study, we proposed an ensemble modeling framework focusing on MFC extrapolation and design optimization. In particular, the effects of various sample sizes on the modeling performance were investigated under normal operation conditions of laboratory MFCs to reduce the efforts and cost during data collection for obtaining an accurate model. Rather than predicting at the original input factor space of the training data set [20], the ensemble model was evaluated for extrapolation. We further applied the ensemble model for MFC optimal design under extreme operation conditions that may provide guidelines for large system development.

2. Experimental and modeling

2.1. MFCs setup and operation

The experimental MFC system consisted of five tubular MFC modules that were placed horizontally in serially hydraulic connection, which had a working anode volume of ~200 mL/each MFC (diameter: 4 cm, length: 15 cm), using synthetic wastewater to continuously and sequentially feed each MFC (Fig. 1). The data of current generation and organic removal efficiency were collected from the entire MFC system. The reason why five MFCs were treated as "one MFC" was because future MFC development will likely have MFC systems consisting of multiple MFC modules and modeling those systems (instead of individual unit) will be more important to understand MFC technology/treatment. The anode electrode material was carbon brush (diameter: 2.6 cm, length: 11 cm), pretreated by the acetone and heated under 450 °C for 30 min, while cathode electrode was carbon cloth coated with activated carbon (5 mg cm^{-2}), using 5% polytetrafluoroethylene

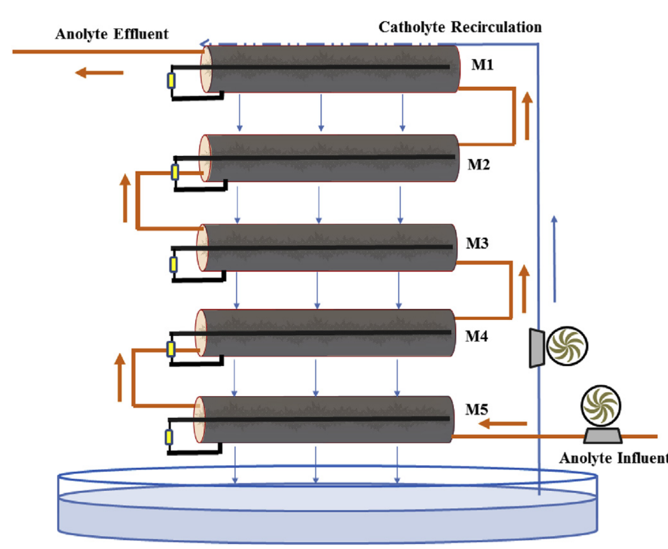


Fig. 1. A schematic of the MFC system consisting of five tubular modules through hydraulic connection.