

# Prediction of MBR Membrane Flux Based on Differential Evolution Algorithm and Recurrent Neural Network

Rujian Huang, Chunqing Li

Computer Science and Technology Institution, Tian Gong University, Tianjin  
Email: [huangrujian123@163.com](mailto:huangrujian123@163.com), [franklcq@163.com](mailto:franklcq@163.com)

Received: Jun. 22<sup>nd</sup>, 2020; accepted: Jul. 8<sup>th</sup>, 2020; published: Jul. 15<sup>th</sup>, 2020

---

## Abstract

Wastewater treatment by membrane bio-reactor (MBR) is a complex dynamic process, which is difficult to be modeled directly by mathematical model. To solve this problem, this paper uses the recurrent neural network (RNN) based on differential evolution algorithm (DE) to predict membrane flux in wastewater treatment. First, authors used principal component analysis to determine the process variables affecting membrane fouling; and then used differential evolution algorithm to optimize initial value of RNN model; finally, the trained DE-RNN model was used to predict and compare with the sample data. The results show that the model has high accuracy, good adaptability and achieves the expected goal.

## Keywords

Membrane Bio-Reactor (MBR), Membrane Flux, Differential Evolution (DE) Algorithm, Recurrent Neural Network (RNN)

---

# 基于差分进化算法与循环神经网络的MBR膜通量预测

黄儒剑, 李春青

天津工业大学计算机科学与技术学院, 天津  
Email: [huangrujian123@163.com](mailto:huangrujian123@163.com), [franklcq@163.com](mailto:franklcq@163.com)

收稿日期: 2020年6月22日; 录用日期: 2020年7月8日; 发布日期: 2020年7月15日

## 摘要

膜生物反应器(Membrane Bio-Reactor, MBR)处理污水是一个复杂的动态过程, 难以用数学模型直接建模。针对该问题, 本文利用差分进化算法(Differential Evolution Algorithm, DE)优化的循环神经网络(Recurrent Neural Network, RNN)对污水处理过程的膜通量进行预测。首先运用主成分分析法确定影响膜通量的相关过程变量; 然后用DE算法优化RNN的初始权值和阈值; 最后用训练好的DE-RNN模型进行预测并与样本数据对比。结果显示, 该模型对膜通量地预测有着较高的准确率, 具有很好的自适应性, 达到了预期目标。

## 关键词

膜生物反应器, 膜通量, 差分进化算法, 循环神经网络

Copyright © 2020 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

## 1. 引言

随着我国工业的不断发展, 水污染问题日益严重[1]。污水的处理与再利用, 已成为“十三五”规划中, 国家环境保护的重点工作之一[2]。膜生物反应器(MBR)是一种由膜分离单元与生物处理单元相结合的新型态、高效能污水处理系统[3]。膜生物反应器具有污水处理效率高、占地面积小、经济效益高等优点[4], 已广泛应用于城市污水处理、工业污水处理等领域。

在膜生物反应器的使用过程中, 膜污染问题不可避免。膜污染会导致污水处理降低, 工艺能耗增加, 缩短中空纤维膜构件的使用寿命[5]。正确预测膜通量的大小, 从而判断膜污染的程度, 及时对膜构件进行清洗、替换, 能够提高膜生物反应器的工作效率、能耗比与使用寿命, 具有极大的实际意义与应用价值[6]。

在工业生产环境中, 衡量膜污染程度的膜通量很难直接测量, 需要根据膜生物反应器的其他运行参数进行预测。因为膜生物反应器处理污水是一个复杂的动态过程, 各污染因子之间相互交叉相互影响, 利用经典的数学模型难以直接建模并进行精确预测[7]。将人工神经网络用于 MBR 污水处理过程仿真, 是近年来进行膜污染预测的新途径[8]。本文采用循环神经网络为算法主体, 膜通量变化过程是一个随时间变化的连续过程, 循环神经网络对时序变量的预测有着更好的适应性。同时利用差分进化算法对 RNN 的初始权值和阈值进行优化, 可以提高模型的训练效率。研究发现, 利用此模型可以准确地预测膜生物反应器的膜通量。

## 2. RNN 神经网络(Recurrent Neural Networks)

循环神经网络(Recurrent Neural Network, RNN)是一类以序列数据作为输入变量, 在序列的演进方向进行递归运算且所有节点(循环单元)按链式连接的递归神经网络[9]。RNN 在 20 世纪 90 年代提出, 经过不断的发展与优化, 在 21 世纪初成为深度学习的核心算法之一。

循环神经网络具有记忆性、参数共享性, 因此在对序列的非线性特征进行学习时具有一定优势, 其简单结构如图 1 所示。其中  $x$  是一个向量, 它代表输入层的值;  $s$  是一个向量, 它代表隐藏层的值;  $o$  是

一个向量, 它代表输出层的值;  $U$  是输入层到隐藏层的权重矩阵;  $V$  是隐藏层到输出层的权重矩阵;  $w$  是一个权重矩阵, 将隐藏层上一次的值作为这一次的输入。RNN 神经网络与全连接神经网络的区别在于, 上一时刻隐藏层的输出, 作为下一时刻隐藏层的输入变量之一参与运算。

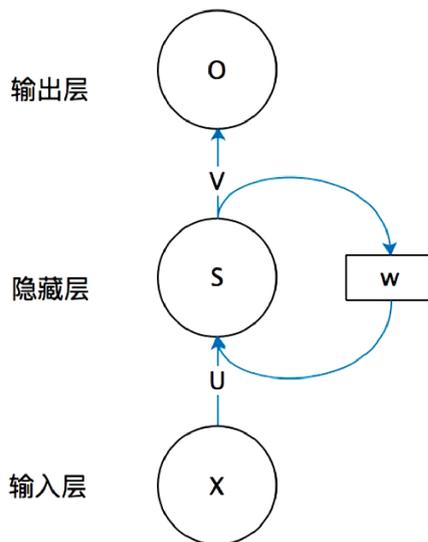


Figure 1. Schematic diagram of RNN  
图 1. RNN 结构示意图

其运算公式如下:

$$O_t = g(V * S_t) \quad (1)$$

$$S_t = f(U * X_t + W * S_{t-1}) \quad (2)$$

RNN 网络在  $t$  时刻接收到  $x_t$  之后, 隐藏层  $S_t$  的值取决于当前时刻输入的变量  $x_t$  和上一时刻隐藏层的值  $S_{t-1}$ 。

### 3. 差分进化算法(Differential Evolution Algorithm)

差分进化算法(Differential Evolution Algorithm, DE)是一种采用浮点矢量编码进行随机搜索的高效的全局优化算法, 对于具体问题条件限制的依赖性低。差分进化算法与遗传算法和粒子群算法相比, 具有工作原理简单、控制参数少、容易实现等优点。

RNN 神经网络的权重矩阵初始值的选择对于模型训练有着重要的意义。初始值选择不好, 容易形成局部最优化, 从而影响预测模型的精确度。用差分进化算法进行初始值的优化, 有利于模型的训练。

DE 算法采用实数编码, 每一代的种群包含  $NP$  个维数为  $D$  的实数向量, 其中个体表示为:

$$x_{i,G} (i = 1, 2, \dots, NP) \quad (3)$$

式中:  $NP$  表示种群规模;  $G$  表示当前进化代数;  $i$  表示个体在第  $G$  代种群中的位置标号。差分进化算法计算步骤如下:

#### 1) 种群初始化

初始化种群, 在边界约束条件下随机选择, 采用下面表达式随机产生初始种群:

$$x = rand(D, NP) * (x^U - x^L) + x^L \quad (4)$$

式中  $x$  表示存放初始种群的向量,  $D$  表示种群中个体的维数,  $NP$  表示种群规模,  $x^U$  表示种群中个体向量的上限,  $x^L$  表示种群中个体向量的下限,  $rand(D, NP)$  表示分布在  $(0,1)$  区间内  $D$  行  $NP$  列的随机矩阵。

#### 2) 变异操作

随机选取当代种群中两个不同向量生成变差向量, 经变异算子  $F$  作用后与第三个向量结合生成变异向量, 如公式(5)

$$v_{i,G+1} = x_{r_1,G} + F * (x_{r_2,G} - x_{r_3,G}) \quad (5)$$

其中  $r_1$ 、 $r_2$ 、 $r_3$  为随机选择的序号且互不相同,  $F$  为变异算子, 变异算子  $F$  通常在 0 到 2 之间取值。

#### 3) 交叉操作

通过将第  $G$  代种群中的目标向量个体  $x_{i,G}$  与其变异向量个体  $v_{i,G+1}$  部分基因进行交叉操作产生试验向量  $u_{i,G+1}$  个体。

#### 4) 选择操作

比较试验向量个体  $u_{i,G+1}$  和当前种群中的目标向量个体  $x_{i,G}$ , 分析不同向量个体的适应度, 若  $u_{i,G+1}$  的适应度更高, 则选其作为子代, 否则选  $x_{i,G}$  作为子代。

## 4. 基于 DE-RNN 的膜通量预测模型

### 4.1. 采用 PCA 进行 MBR 系统参数选取

主成分分析法(Principal Component Analysis, PCA)是一种统计分析方法, 用于掌握事物主要矛盾变量, 它通过对原始数据集进行线性空间投影来分析矩阵特征值, 从多维数据集中解析出主要影响因素, 从而达到对数据集降维, 简化复杂的问题的目的。

膜生物反应器的运行参数众多, 如果将全部运行参数作为输入变量通过 RNN 运算, 不仅计算量庞大, 而且影响预测模型的准确率。通过 PCA 算法, 减少 RNN 输入变量的个数, 从而提升模型的性能。

主成分分析法步骤为:

Step1: 把初选的影响因子组成矩阵  $X$ 。

Step2: 对矩阵  $X$  进行标准化处理, 得到标准化矩阵  $A$ 。

Step3: 计算出协方差矩阵  $S$ 。

Step4: 计算协方差矩阵  $S$  的特征值  $\lambda$  与相应的特征向量  $u$ , 然后由特征值降序排列得到特征值矩阵  $V$  和特征向量矩阵  $U$ 。

Step5: 由特征值矩阵  $V$  计算出贡献率, 确定主成分。

初选 MBR 系统的六个参数作为膜污染影响因子, 分别为总阻力、MLSS、COD、PH、温度、操作压力, 并以此作为 RNN 输入层的六个神经元。膜通量为 RNN 输出层的唯一神经元。

### 4.2. 建立基于 DE-RNN 的膜通量预测模型

循环神经网络 RNN 在应用时, 网络权值和阈值的初始值选取不当, 容易导致算法的局部最优化, 从而影响网络识别的精度。本文采用差分进化算法对 RNN 的权值和阈值进行优化, 以加快 RNN 的收敛速度和识别精度。

建立基于 DE-RNN 的预测模型。将 RNN 的参数作为种群的个体, RNN 的计算结果作为判断标准, 执行 DE 算法, 得到满足精确度的种群个体, 及经过优化的 RNN 权值和阈值的初始值。然后对经过优化的 RNN 进行训练, 得到用于预测的 DE-RNN 模型。算法的程序流程图如图 2 所示。

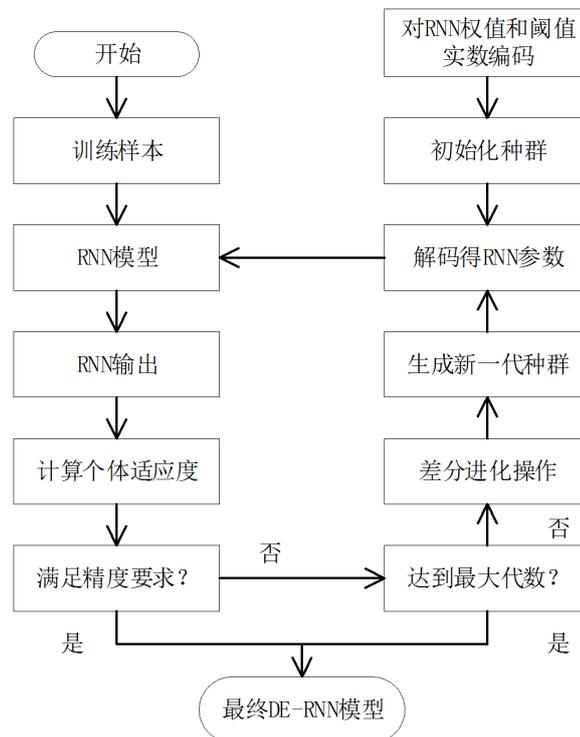


Figure 2. Procedure flow chart of DE-RNN  
图 2. DE-RNN 算法的程序流程图

具体实验步骤如下:

- Step1: 对 RNN 的参数编码, 所有编码得到的基因位串称为一个个体。  
 Step2: 初始种群随机产生, 个体数量为 20, 构建初始的交配池。  
 Step3: 执行 RNN 算法, 将误差百分比作为衡量个体适应度的标准。个体间基因进行交叉和变异。  
 Step4: 保留适应度高的个体, 淘汰适应度低的个体。  
 Step5: 回到(3)进行循环, 直到得出最佳结果结束循环。  
 Step6: 训练完毕, 得到 RNN 权值和阈值的最优值, 进行仿真实验。

## 5. 预测结果与实验结果对比分析

实验数据来自于石家庄市某污水处理厂, 此污水处理厂采用 MBR 污水处理工艺, 过滤膜为孔径为  $0.2\ \mu\text{m}$  的聚偏氟乙烯微滤膜。25 组实验数据作为训练样本, 5 组实验数据作为校验用样本。

通过 PCA 算法选取的水质指标为 6 个: 总阻力、MLSS、COD、PH、温度、操作压力, 作为输入变量。本文使用的 RNN 的结构为 6-10-1, 使用快速梯度下降算法调整参数, 输出单元为膜通量。进行实验, 用 RNN 模型和 DE-RNN 模型分别预测膜通量并计算拟合度, DE-RNN 模型预测数据和实际数据的对比结果如图 3 所示, RNN 模型的预测拟合度如图 4 所示, DE-RNN 模型的预测拟合度如图 5 所示。

对比实际数据和预测数据可知, 预测数据的准确度在 0.9 以上, 该模型具有较高的精确度, 达到了准确预测膜通量的预期目标。对比 RNN 模型和 DE-RNN 模型的拟合度, RNN 模型的拟合度最低为 0.87, 平均为 0.94; DE-RNN 模型的拟合度最低为 0.91, 平均为 0.97。DE-RNN 模型比 RNN 模型拟合度更高, 误差更小。用 DE 算法优化 RNN, 对预测模型的准确度有着明显的提升。

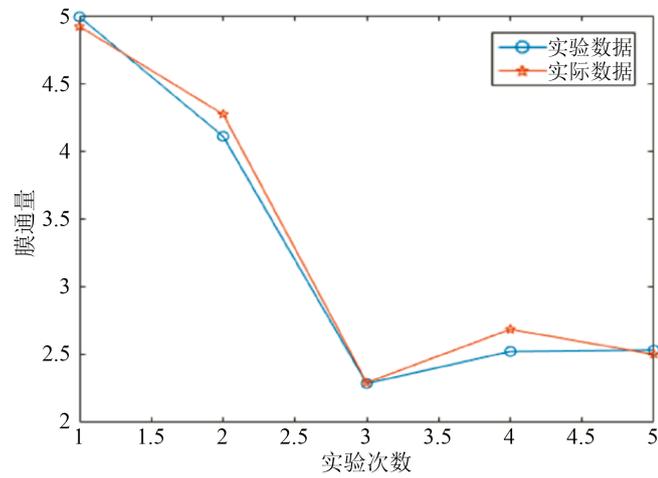


Figure 3. Relationship between actual value of membrane flux and test value

图 3. 膜通量的实际值与测试值关系图

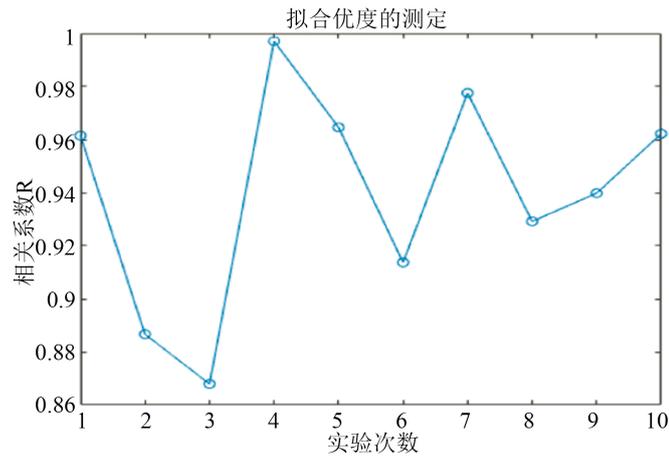


Figure 4. Fit test experiment of RNN

图 4. RNN 拟合度测定实验图

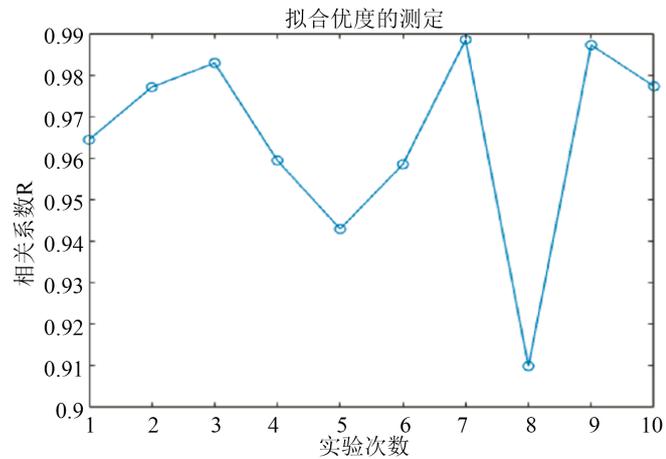


Figure 5. Fit test experiment of DE-RNN

图 5. DE-RNN 拟合度测定实验图

## 6. 结论

MBR 处理污水是一个复杂的动态过程, 具有多流程、时变、不确定等特点, 是一个运行非平稳的系统, 用数学模型难以直接建模。本文在用主成分分析法简化 MBR 系统参数的基础上, 用差分进化算法优化循环神经网络的初始权值和阈值, 建立了 DE-RNN 仿真模型用以预测膜通量。实验结果表明, DE-RNN 预测模型有着很高的精确度, 该模型是有效的、可行的。DE-RNN 模型与 RNN 模型相比拟合度更高, 说明用 DE 算法优化 RNN, 对提高精确度, 提高适应度有着积极的作用, 对模型的优化是明显有效的。本文的研究有一定的理论价值和实践意义, 对 MBR 膜通量预测可以起到积极的指导作用。

## 基金项目

国家自然科学基金项目(51378350); 国家自然科学基金青年科学基金资助项目(50808130); 国家自然科学基金青年科学基金项目(21506159)。

## 参考文献

- [1] 中华人民共和国环境保护部. 2018 中国环境状况公报[Z], 2019-05-29.
- [2] 中华人民共和国国务院. 国家环境保护“十三五”规划[Z]. 中华人民共和国国务院(国发[2016] 53 号), 2016-03-20.
- [3] Zhang, Y., Chen, Z.M., An, W., Xiao, S.M., Yuan, H.Y., Zhang, D.Q. and Yang, M. (2015) Risk Assessment of Giardia from a Full Scale MBR Sewage Treatment Plant Caused by Membrane Integrity Failure. *Journal of Environmental Sciences*, **30**, 252-258. <https://doi.org/10.1016/j.jes.2014.09.033>
- [4] Meng, F., Chae, S.R., Drews, A., et al. (2009) Recent Advances in Membrane Bioreactors (MBRs): Membrane Fouling and Membrane Material. *Water Research*, **43**, 1489-1512. <https://doi.org/10.1016/j.watres.2008.12.044>
- [5] Huang, L.Y. and Lee, D.-J. (2015) Membrane Bioreactor: A Mini Review on Recent R&D Works. *Bioresource Technology*, **194**, 383-388. <https://doi.org/10.1016/j.biortech.2015.07.013>
- [6] 韩永萍, 肖燕, 宋蕾, 王晓琳. MBR 膜污染的形成及其影响因素研究进展[J]. 膜科学与技术, 2013, 33(1): 102-110.
- [7] Zhang, M., Liao, B.Q., Zhou, X., et al. (2015) Effects of Hydrophilicity/Hydrophobicity of Membrane on Membrane Fouling in a Submerged Membrane Bioreactor. *Bioresource Technology*, **175**, 59-67. <https://doi.org/10.1016/j.biortech.2014.10.058>
- [8] Kaneko, H. and Funatsu, K. (2013) Physical and Statistical Model for Predicting a Transmembrane Pressure Jump for a Membrane Bioreactor. *Chemometrics & Intelligent Laboratory Systems*, **121**, 66-74. <https://doi.org/10.1016/j.chemolab.2012.11.013>
- [9] Goodfellow, I., Bengio, Y. and Courville, A. (2016) Deep Learning. The MIT Press, Cambridge, Massachusetts.