

(12) 按照专利合作条约所公布的国际申请

(19) 世界知识产权组织
国际局

(43) 国际公布日
2021 年 6 月 17 日 (17.06.2021)



(10) 国际公布号
WO 2021/114320 A1

(51) 国际专利分类号:
G01N 33/18 (2006.01)

(21) 国际申请号: PCT/CN2019/125888

(22) 国际申请日: 2019 年 12 月 17 日 (17.12.2019)

(25) 申请语言: 中文

(26) 公布语言: 中文

(30) 优先权:
201911298706.X 2019 年 12 月 14 日 (14.12.2019) CN

(71) 申请人: 北京工业大学(BEIJING UNIVERSITY OF TECHNOLOGY) [CN/CN]; 中国北京市朝阳区平乐园100号, Beijing 100124 (CN)。

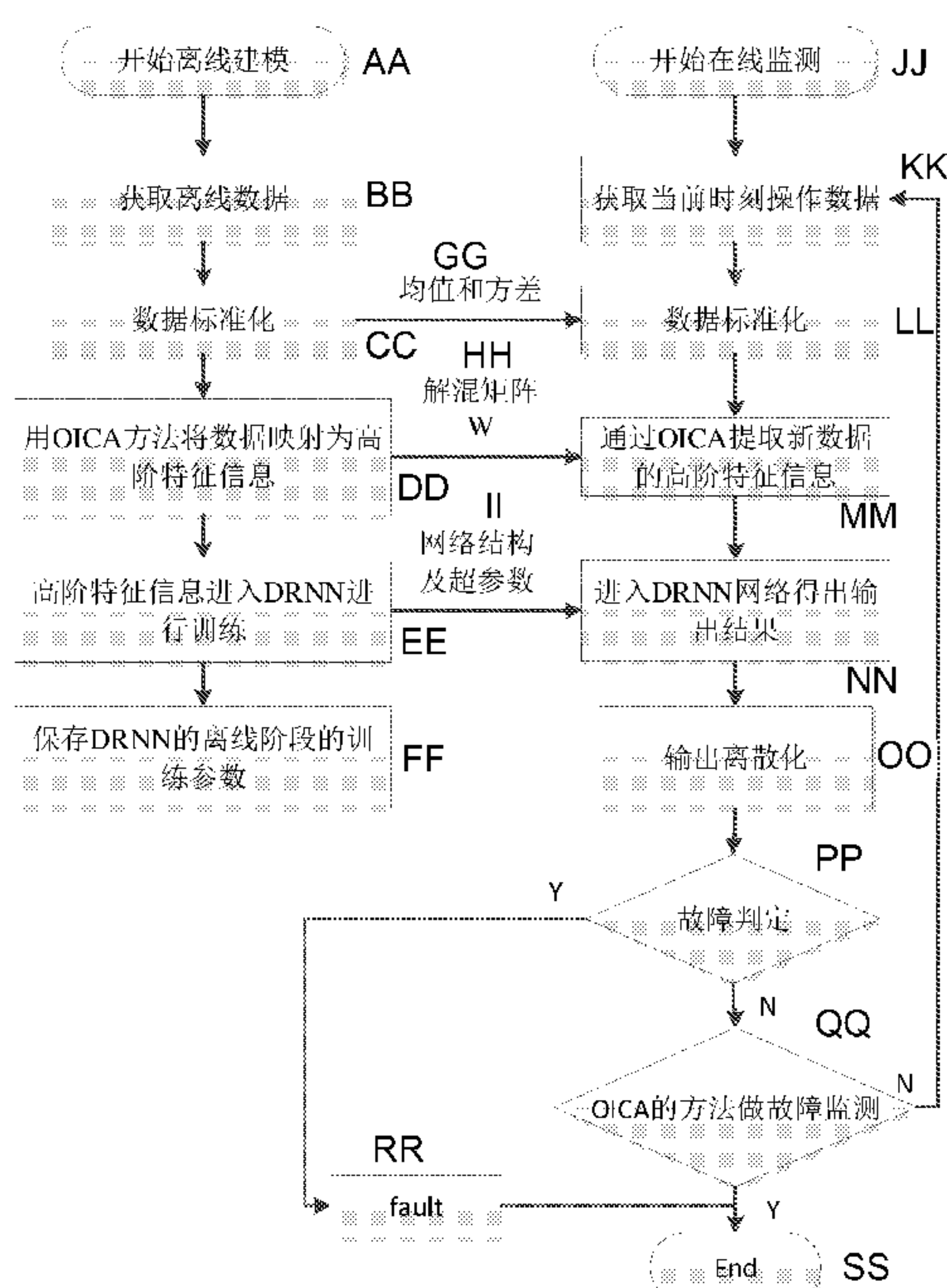
(72) 发明人: 常鹏(CHANG, Peng); 中国北京市朝阳区平乐园100号, Beijing 100124 (CN)。 李泽宇(LI,

Zeyu); 中国北京市朝阳区平乐园100号, Beijing 100124 (CN)。 王凯(WANG, Kai); 中国北京市朝阳区平乐园100号, Beijing 100124 (CN)。 丁春豪(DING, Chunhao); 中国北京市朝阳区平乐园100号, Beijing 100124 (CN)。 金辰(JIN, Chen); 中国北京市朝阳区平乐园100号, Beijing 100124 (CN)。 张祥宇(ZHANG, Xiangyu); 中国北京市朝阳区平乐园100号, Beijing 100124 (CN)。 卢瑞炜(LU, Ruiwei); 中国北京市朝阳区平乐园100号, Beijing 100124 (CN)。 王普(WANG, Pu); 中国北京市朝阳区平乐园100号, Beijing 100124 (CN)。

(74) 代理人: 北京思海天达知识产权代理有限公司(BEIJING SIHAI TIANDA INTELLECTUAL PROPERTY AGENCY LTD.); 中国北京市朝阳区百子湾西里403号楼1011, Beijing 100124 (CN)。

(54) Title: WASTEWATER TREATMENT PROCESS FAULT MONITORING METHOD USING OICA-RNN FUSION MODEL

(54) 发明名称: 一种OICA和RNN融合模型的污水处理过程故障监测方法



AA START OFFLINE MODELING
BB OBTAIN OFFLINE DATA
CC DATA STANDARDIZATION
DD USE OICA METHOD TO MAP DATA AS HIGH-ORDER FEATURE INFORMATION
EE HIGH-ORDER FEATURE INFORMATION ENTERS DRNN AND TRAINING IS CARRIED OUT
FF SAVE DRNN OFFLINE-PHASE TRAINING PARAMETERS
GG MEAN VALUE AND VARIANCE
HH DEMIXING MATRIX W
II NETWORK STRUCTURE AND HYPER-PARAMETERS
JJ START ONLINE MONITORING
KK OBTAIN CURRENT OPERATION DATA
LL DATA STANDARDIZATION
MM EXTRACT HIGH-ORDER FEATURE INFORMATION OF NEW DATA BY MEANS OF OICA
NN ENTER DRNN AND OBTAIN OUTPUT RESULTS
OO OUTPUT DISCRETIZATION
PP FAULT DETERMINATION
QQ CARRY OUT FAULT MONITORING USING OICA METHOD
RR FAULT
SS END
Y YES
N NO

图 1

(57) Abstract: A smart fault monitoring method based on a high-order information-enhanced recurrent neural network, said method being used to carry out real-time monitoring of wastewater treatment process faults, and comprising two phases: offline training and online soft measurement. The offline phase first uses OICA to extract raw data into high-dimensional high-order information features, used for the effective processing of non-Gaussian properties of the data and determining the correlation between variables. Then training of the extracted features is carried out by means of a DRNN. In the online phase, the data are mapped directly into new high-order feature

(81) 指定国(除另有指明, 要求每一种可提供的国家保护): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW。

(84) 指定国(除另有指明, 要求每一种可提供的地区保护): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), 欧亚 (AM, AZ, BY, KG, KZ, RU, TJ, TM), 欧洲 (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG)。

本国际公布:
— 包括国际检索报告(条约第21条(3))。

components and are classified and discriminated by the DRNN that was trained offline. If the results are fault-free, then a monitoring model composed only of OICA is used to carry out unsupervised monitoring. If a fault has still not been detected at this point, the process is determined to be fault-free, and if a fault occurs, then the process is determined to be faulty and the fault information is entered into the network training data and training is carried out, thereby continuously improving the monitoring precision of the DRNN.

(57) 摘要: 一种基于高阶信息增强的循环神经网络的智能故障监测方法, 用于对污水处理过程故障进行实时监测, 包括离线训练和在线软测量两个阶段。离线阶段首先采用OICA把原始数据提取成为高维的高阶信息特征, 用于有效的处理数据的非高斯性, 并解决变量间的相关性。之后将提取的特征通过DRNN进行训练。在线阶段中, 将数据直接映射成新的高阶特征分量并且通过离线训练好的DRNN网络进行分类判别。若结果无故障, 再进入由单纯的OICA构成的监测模型中进行无监督监测, 此时还没有监测出故障则判定为过程无故障, 若发生故障则判定过程故障, 并把故障信息加入网络的训练数据中进行训练, 从而不断提高DRNN的监测精度。

一种 OICA 和 RNN 融合模型的污水处理过程故障监测方法

技术领域

本发明涉及基于深度学习的故障监测技术领域，特别是涉及一种针对复杂工业过程的故障监测技术。本发明的基于深度学习的方法即是在典型复杂工业过程——污水处理过程故障监测方面的具体应用。

背景技术

污水处理过程是一个外界干扰强烈、时变性强、耦合性强、非线性的复杂动态生物化学过程，因此控制系统的可靠性和稳定性显得尤其重要。但对于过程中出现的很多异常变化(故障)，控制器往往无能为力。由于污水处理系统的工作连续性和不可替代性，一旦发生故障，就会造成严重影响。由于污水处理过程的处理过程机理特性复杂和外界环境干扰严重等特点，导致污水处理过程的数据带有明显的非线性、非高斯性和时间相关性等特点。传统方法在对于污水处理过程故障监测上效果欠佳。

近年来基于数据驱动的方法得到了广泛的发展，基于数据驱动的方法不需要研究污水处理过程复杂的机理知识，只是通过过程变量的变化就可以实时的得出监控结果，得到了广泛的应用。在传统的基于数据驱动的方法上，以 KPCA(Kernel Principal Component Analysis, KPCA)和 KPLS(Kernel Partial Least Squares, KPLS) 等的多元统计学方法为主，这些方法可以提取过程的潜在特征变量，从而捕获过程变化的信息，从而反映故障的发生。基于 KPCA, KPLS 等的方法可以有效的处理数据的非线性，但是以上方法均需要假设过程数据服从高斯分布，在实际的工业过程的数据由于复杂环境的干扰大多是不服从高斯分布的，在实际应用中有很多的限制。为了处理数据的非高斯问题，独立成分分析(independent component analysis, ICA)被提出并广泛的应用于数据的非高斯性特征的提取上。ICA 可以有效的利用数据的非高斯性提取特征。但 ICA 在求解的过程中要大量的迭代并且得到的解有高度的不确定性，导致 ICA 在应用时困难。现在缺乏一种有效的数据处理手段对于污水处理过程进行监测。近年来神经网络的方法也广泛的应用于污水的过程监测上，如 BP 神经网络，RBF 神经网络等。相对于多元统计方法，神经网络的非线性处理能力更强，但是在应用于污水监测过程中没有考虑

数据的非高斯性和时间相关性。并且神经网络的方法为有监督的监测，数据的标签会对于污水处理的过程监测产生一定的限制。

发明内容

为了克服上述所提的两个技术要素的不足。本文建立一种基于高阶信息增强的循环神经网络的智能故障监测方法。在特征提取阶段，本文选择运用 OICA (Overcomplete Independent Component Analysis) 的方法把原始数据提取成为高阶信息特征，OICA 算法由麻省理工学院的 Anastasia 等人提出，该算法无需假设数据服从高斯分布，计算复杂度低，而且不受混合矩阵形式的限制。之后将 OICA 提取的特征数据进入到多层的循环神经网络 DRNN (Deep Recurrent Neural Network) 进行逐层的训练。循环神经网络可以学习数据中具有多个抽象层次的时间序列信息，对于数据的特征变化更加敏感，更易监测到故障。在通过 DRNN 进行监测的同时，提取的高阶统计信息直接建立监测模型进行监测，OICA 直接建立监测的方法是无监督的监测方法，这样做的目的是为了监测到在已有的标签信息中没有的故障类型，可以在提高监测准确率的基础上扩充已有的故障数据数据库，使得监测结果随着时间的增加监测能力逐渐增高。

本发明采用了如下的技术方案及实现步骤：

A. 离线建模阶段：

- 1) 对采集污水处理过程正常工况下的历史数据，所述的历史数据X由离线测试得到的污水处理过程正常操作状态的数据构成，数据包含N个采样时刻，每个采样时刻采集J个过程变量形成数据矩阵 $X = [x_1, x_2, \dots, x_N]^T \in \mathbb{R}^{N \times J}$ 。其中，对于每个采样时刻 $x_i = (x_{i,1}, x_{i,2}, \dots, x_{i,j})$ ， $x_{i,j}$ 表示第i个采样时刻的第j个变量的测量值；
- 2) 然后对历史数据X进行标准化，其中第i个采样时刻的第j个变量的标准化公式如下：

$$\overline{x_{i,j}} = \frac{x_{i,j} - \text{Mean}(j)}{\text{Std}(j)}$$

其中， $i = 1, 2, \dots, N$ ， $j = 1, 2, \dots, J$ ；将步骤 2 标准化后的数据重新构造成二维矩阵，如下式所示：

$$\bar{X} = \begin{bmatrix} \overline{x_{1,1}} & \cdots & \overline{x_{1,j}} \\ \vdots & \ddots & \vdots \\ \overline{x_{l,1}} & \cdots & \overline{x_{l,j}} \end{bmatrix}$$

- 3) 利用上述所提到的 OICA 算法将 $\bar{X}$ 映射为高阶特征矩阵 S ，映射的高阶特征可以有效的反映数据的非高斯特征，能够提供更多的故障信息。具体的步骤如下，通过 OICA 计算出解混矩阵 W ，之后利用 W 将原数据 $\bar{X}$ 映射成为高阶特征矩阵 S 。通过 W 得到 $\bar{X}$ 的高阶特征矩阵 S 的公式如下：

$$S = W^T \bar{X}^T$$

进一步的，根据 S 得到残差 E ，求得残差的公式如下所示：

$$E = \bar{X} - WS$$

- 4) 分别根据 S 和 E 计算独立成分空间的统计量 I^2 和残差空间的统计量 SPE ，如下式所示：

$$I^2 = S^T S$$

$$SPE = E^T E$$

利用核密度估计算法求得上述 I^2 和 SPE 统计量在预设置的置信限时的估计值 I_{limit}^2 和 SPE_{limit} ，并将其作为后续运用 OICA 进行故障监测的控制限。

- 5) 之后对于历史数据 X 设立标签 Y 。根据每一时刻 X 对应的故障类型，污水处理过程为正常时设为 1，过程为故障设为 0。
- 6) 将步骤 3 得到的高阶特征矩阵 S 和步骤 5 得到的标签数据 Y 进入深度循环神经网络 DRNN 中进行有监督训练。深度循环神经网络的输入为 OICA 得到的高阶特征信息 S ，网络输入对应的标签数据为步骤 5 得到的故障分类标签为得到的标签 Y 。经过训练后保存 DRNN 经过监督训练过后网络中神经元的参数和结构。

B. 在线监测阶段：

- 1) 在线监测时新数据的预处理方式如离线的步骤 2，得到处理过后的新数据 X_{new}
- 2) 将新数据 X_{new} 通过离线阶段得到的解混矩阵 W 得到新的高阶特征信息特征数据 S_{new}

$$S_{new} = W^T \bar{X}_{new}^T$$

- 3) 将 S_{new} 作为网络的输入进入在离线阶段训练好的网络参数的 DRNN 深度

循环神经网络中进行运算，数据经过 DRNN 神经元的运算会得到一个输出 y ， y 为我们得到当前评判是否故障的指标数据。当 y 大于 0.5 则表示当前故障，当 y 小于 0.5 表示通过 DRNN 得到的监测结果为当前时刻无故障。

- 4) 基于 DRNN 的方法可以很好的对于故障进行有监督分类，但是当发生 DRNN 网络的训练库中没有的故障时，上述的方法的监测性能可能会降低。进一步的，本发明算法提出一种基于 OICA 的无监督算法对于上述故障进行监测，从而对于 DRNN 的监测结果进行校准。当 DRNN 得到的监测结果为正常时，进行二次监测，具体步骤如下，首先通过高阶统计信息 S_{new} 得到新数据 X_{new} 的残差 E_{new} ，如下式所示：

$$E_{new} = \bar{X}_{new} - WS_{new}$$

其中 W 为步骤 4) 中所确定的解混矩阵；

- 5) 计算当前采样时刻 k 的监控统计量 I_k^2 和 SPE_k ，如下式所示：

$$I_k^2 = S_{new}' S_{new}$$

$$SPE_k = E_{new}' E_{new}$$

- 6) 将上述步骤得到的监控统计量 I_k^2 和 SPE_k 与步骤 6) 得到的控制限 I_{limit}^2 和 SPE_{limit} 进行比较，若上述两个指标中其中任意一个指标超限就认为发生故障并报警；否则即认为是正常；
- 7) 将故障数据按照离线步骤 5 设立故障标签并加入 DRNN 的训练数据库中进行训练，不断的迭代训练使 DRNN 网络可以一直学习到新的故障信息。

有益效果

与现有技术相比，基于高阶信息增强的循环神经网络的智能故障监测方法能够处理数据的非高斯性，提高了对于原始数据的特征提取能力，并且融合循环神经网络结构可以提取不同层次的污水数据的时序性信息，在污水的监测上可以有效提高监测的精确度。并且通过同时进行监测的 OICA 无监督模型的监测校准，可以不断的提高故障的有监督训练数据，提升整体监测模型的监测精度。

附图说明

图 1 为本发明算法的整体流程图；

图 2 为晴天下对于污水污泥膨胀故障的监测图；

- 图 3 为晴天下对污水毒性冲击故障的监测图；
- 图 4 为雨天下对于污水污泥膨胀故障的监测图；
- 图 5 为雨天下对污水毒性冲击故障的监测图；
- 图 6 本方法依托的硬件系统逻辑框图；
- 图 7 为本发明方法提出的网络结构示意图。

具体实施方式

为了解决上述问题，提出一种 OICA 和 RNN 融合模型的污水处理过程故障监测方法，该方法基于一种在线监测仪器设备。整个设备包含输入模块，信息处理模块，控制台模块，输出结果可视化模块。将所提出的方法导入信息处理模块，然后用实际工业所保留的过程数据建立网络监测模型，所建立的模型保存下来，用于在线故障监测。在实际的工业过程在线监测时，首先将工厂数据传感器采集到的实时过程变量连接到输入模块，作为监测设备的输入信息，然后通过控制台选择之前所训练好的模型进行监测，并将监测结果通过可视化模块实时的显示出来，以便现场工作人员能够根据可视化监测结果及时做出相应的措施，以减少由于过程故障带来的经济损失。

污水处理过程极其复杂，不仅包含了各种物理、化学，而且也包括了生化反应，此外，各种不确定性因素充斥其中，如进水流量、水质和负荷变化等，这给污水处理监控模型的建立带来了巨大的挑战。本发明采用国际水协会（IWA）研发的“仿真基准模型”（Benchmark Simulation Model 1）作为实际污水处理过程进行实时仿真。该模型由五个反应池（5999 m³）和一个二沉池（6000 m³）组成，此外还有三个曝气池。曝气池有 10 层，深 4 米，占地 1500 m²，反应过程有内回流和外回流。平均污水处理流量为 20 000 m³/d，化学需氧量为 300 mg/l。污水模型的出水水质指标如表 1 所示。在模型故障设置上本发明基于 BSM1 模型模拟两种故障，污泥膨胀故障和毒性冲击故障

表 1 污水出水指标

变量	单位
出水流量	m ³ . d
出水 SI 浓度	g COD. m ⁻³

出水 SS 浓度	$\text{g COD} \cdot \text{m}^{-3}$
出水 XI 浓度	$\text{g COD} \cdot \text{m}^{-3}$
出水 XS 浓度	$\text{g COD} \cdot \text{m}^{-3}$
出水 XBH 浓度	$\text{g COD} \cdot \text{m}^{-3}$
出水 XBA 浓度	$\text{g COD} \cdot \text{m}^{-3}$
出水 XP 浓度	$\text{g COD} \cdot \text{m}^{-3}$
出水 SO 浓度	$\text{g } (-\text{COD}) \cdot \text{m}^{-3}$
出水 SNO 浓度	$\text{g N} \cdot \text{m}^{-3}$
出水 SNH 浓度	$\text{g N} \cdot \text{m}^{-3}$
出水 SND 浓度	$\text{g N} \cdot \text{m}^{-3}$
出水 XND 浓度	$\text{g N} \cdot \text{m}^{-3}$
出水 SALK 浓度	$\text{mol HCO}_3^- \cdot \text{m}^{-3}$
出水 TSS 浓度	$\text{g SS} \cdot \text{m}^{-3}$
出水 Kjeldahl N 浓度	$\text{g N} \cdot \text{m}^{-3}$

本发明在上述 BSM1 仿真平台的应用过程具体陈述如下：

A. 离线建模阶段：

步骤 1：本发明模拟污水处理过程的污泥膨胀故障和毒性冲击故障对于算法进行验证。BSM1 模型采集到正常天气和暴雨 14 天的数据，采样间隔为 15min，每个天气的总采样点为 1344 个。实验采用同类型下的多批不同故障度的污泥膨胀数据和正常数据进行离线训练，并再训练一组新的单批污泥故障数据作为测试，模拟的毒性冲击故障的训练和测试数据与污泥膨胀故障相同。

步骤 2：对采集污水处理过程正常工况下的离线数据进行处理，其包含多批数据采集到的 N 个采样时刻，和采集 16 个过程变量形成数据矩阵 $\mathbf{X} = [x_1, x_2, \cdots x_N]^T \in \mathbb{R}^{N \times 16}$ 。其中，对于每个采样时刻 $x_i = (x_{i,1}, x_{i,2}, \cdots, x_{i,j})$ ， $x_{i,j}$ 表示第 i 个采样时刻的第 j 个变量的测量值；

步骤 3：然后对历史数据 $\mathbf{X}$ 进行标准化，其中第 i 个采样时刻的第 j 个变量的标准化公式如下：

$$\overline{x_{i,j}} = \frac{x_{i,j} - \text{Mean}(j)}{\text{Std}(j)}$$

其中, $i = 1, 2, \dots, N$, $j = 1, 2, \dots, J$; 将步骤 2 标准化后的数据重新构造成二维矩阵, 如下式所示:

$$\bar{X} = \begin{bmatrix} \overline{x_{1,1}} & \cdots & \overline{x_{1,J}} \\ \vdots & \ddots & \vdots \\ \overline{x_{i,1}} & \cdots & \overline{x_{i,J}} \end{bmatrix}$$

步骤 4: 利用上述所提到的 OICA 算法将 $\bar{X}$ 映射为高阶特征矩阵 S , 映射的高阶特征可以有效的反映数据的非高斯特征, 能够提供更多的故障信息。具体的步骤如下, 通过 OICA 计算出解混矩阵 W , 之后利用 W 将原数据 $\bar{X}$ 映射成为高阶特征矩阵 S 。通过 W 得到 $\bar{X}$ 的高阶特征矩阵 S 的公式如下:

$$S = W^T \bar{X}^T$$

进一步的, 根据 S 得到残差 E , 求得残差的公式如下所示:

$$E = \bar{X} - WS$$

步骤 5: 分别根据 S 和 E 计算独立成分空间的统计量 I^2 和残差空间的统计量 SPE , 如下式所示:

$$I^2 = S^T S$$

$$SPE = E^T E$$

利用核密度估计算法求得上述 I^2 和 SPE 统计量在预设置的置信限时的估计值 I_{limit}^2 和 SPE_{limit} , 并将其作为后续运用 OICA 进行故障监测的控制限。

步骤 6: 之后对于历史数据 X 设立标签 Y 。根据每一时刻 X 对应的故障类型, 污水处理过程为正常时设为 1, 过程为故障设为 0。

步骤 7: 将步骤 3 得到的高阶特征矩阵 S 和步骤 5 得到的标签数据 Y 进入深度循环神经网络 DRNN 中进行有监督训练。深度循环神经网络的输入为 OICA 得到的高阶特征信息 S , 网络输入对应的标签数据为步骤 5 得到的故障分类标签为得到的标签 Y 。经过训练后保存 DRNN 经过监督训练过后网络中神经元的超参数和结构。DRNN 的具体神经网络结构及参数如下表所示。

表 1 DRNN 的网络结构及超参数

超参数	参数值
迭代次数	100
隐含层层数	3
隐含层中每层的神经元个数	30-20-10

学习率

0.01

B. 在线监测阶段:

步骤 8: 在线监测时新数据的预处理方式如离线的步骤 3, 得到处理过后的新数据 X_{new}

步骤 9: 将新数据 X_{new} 通过离线阶段得到的解混矩阵 W 得到新的高阶特征信息特征数据 S_{new}

$$S_{new} = W^T \bar{X}_{new}^T$$

步骤 10: 将 S_{new} 作为网络的输入进入在离线阶段训练好网络参数的 DRNN 深度循环神经网络中进行运算, 数据经过 DRNN 神经元的运算会得到一个输出 y , y 为我们得到当前评判是否故障的指标数据。当 y 大于 0.5 则表示当前故障, 当 y 小于 0.5 表示通过 DRNN 得到的监测结果为当前时刻无故障。

步骤 11: 基于 DRNN 的方法可以很好的对于故障进行有监督分类, 但是当发生 DRNN 网络的训练库中没有的故障时, 上述的方法的监测性能可能会降低。进一步的, 本发明算法提出一种基于 OICA 的无监督算法对于上述故障进行监测, 从而对于 DRNN 的监测结果进行校准。当 DRNN 预测为正常时, 做二次监测, 监测步骤如下, 首先通过高阶统计信息 S_{new} 得到新数据 X_{new} 的残差 E_{new} , 如下式所示:

$$E_{new} = \bar{X}_{new} - W S_{new}$$

其中 W 为步骤 4) 中所确定的解混矩阵;

步骤 12: 计算当前采样时刻 k 的监控统计量 I_k^2 和 SPE_k , 如下式所示:

$$I_k^2 = S_{new}' S_{new}$$

$$SPE_k = E_{new}' E_{new}$$

步骤 13: 将上述步骤得到的监控统计量 I_k^2 和 SPE_k 与步骤 6) 得到的控制限 I_{limit}^2 和 SPE_{limit} 进行比较, 若上述两个指标中其中任意一个指标超限就认为发生故障并报警; 否则即认为是正常;

步骤 15: 将故障数据按照离线步骤 5 设立故障标签并加入 DRNN 的训练数据库中进行训练, 不断的迭代训练使 DRNN 网络可以一直学习到新的故障信息。

以上即为本发明再 BSM1 污水仿真平台上的污水处理过程故障监测的具体应用步骤，为了验证本方法的有效性，本发明在污水的晴天和雨水下分别设置污泥膨胀和毒性冲击两种故障，检验本发明在不同天气下的监测准确性。图 2-图 5 分别是晴天和雨天下污泥膨胀的监测图，图中离散化的分类值中 1 代表故障发生。表 1 为故障的报警时间及误报率和漏报率。从图 2-5 和表 1 中可以看出，本发明方法对于污泥的故障发生可以有效的监测，同时具有较低的漏报率和误报率。并且在雨天较复杂的环境下也有很好的监测性能，表明本发明的鲁棒性较强。

表 2 不同条件下本发明的监测性能

故障类型	故障时间	报警时间	误警数	漏警数
晴天污泥膨胀故障	672-864	672	0	1
晴天毒性冲击故障	672-864	672	3	1
雨天污泥膨胀故障	672-864	672	1	2
雨天毒性冲击故障	672-864	672	0	1

1. 一种 OICA 和 RNN 融合模型的污水处理过程故障监测方法，包括“离线建模”和“在线监测”两个阶段，具体步骤如下：

A. 离线建模阶段：

- 1) 采集污水处理过程的历史数据，所述的历史数据X由离线测试得到的污水处理过程正常的数据构成，数据包含N个采样时刻，每个采样时刻采集J个过程变量形成数据矩阵 $X = [x_1, x_2, \dots, x_N]^T \in \mathbb{R}^{N \times J}$ ，其中， $x_i = (x_{i,1}, x_{i,2}, \dots, x_{i,J})$ ， $x_{i,j}$ 表示第i个采样时刻的第j个变量的测量值；
- 2) 然后对历史数据X进行标准化，其中第i个采样时刻的第j个变量的标准化公式如下：

$$\bar{x}_{i,j} = \frac{x_{i,j} - \text{Mean}(j)}{\text{Std}(j)}$$

其中， $i = 1, 2, \dots, N, j = 1, 2, \dots, J$ ；将步骤 2 标准化后的数据重新构造造成二维矩阵，如下式所示：

$$\bar{X} = \begin{bmatrix} \bar{x}_{1,1} & \cdots & \bar{x}_{1,J} \\ \vdots & \ddots & \vdots \\ \bar{x}_{i,1} & \cdots & \bar{x}_{i,J} \end{bmatrix}$$

- 3) 利用 OICA 算法将 $\bar{X}$ 映射为高阶特征矩阵S，具体的步骤如下，通过 OICA 计算出解混矩阵W，之后利用W将原数据 $\bar{X}$ 映射成为高阶特征矩阵S，通过W得到 $\bar{X}$ 的高阶特征矩阵S的公式如下：

$$S = W^T \bar{X}^T$$

进一步的，根据S得到残差E，求得残差的公式如下所示：

$$E = \bar{X} - WS$$

- 4) 分别根据S和E计算独立成分空间的统计量 I^2 和残差空间的统计量SPE，如下式所示：

$$I^2 = S^T S$$

$$\text{SPE} = E^T E$$

利用核密度估计算法求得上述 I^2 和SPE统计量在预设置的置信限时的估计值 I_{limit}^2 和 SPE_{limit} ，并将其作为后续运用 OICA 进行故障监测的控制

限；

- 5) 之后对于历史数据X设立标签Y，即正常、故障两种。
- 6) 将步骤 3 得到的高阶特征矩阵S和步骤 5 得到的标签数据Y输入深度循环神经网络 DRNN 中进行有监督训练；经过训练后保存 DRNN 经过监督训练过后网络中神经元的参数和结构。

B. 在线监测阶段：

- 1) 在线监测时新数据的预处理方式如离线的步骤 2，得到处理过后的新数据 X_{new} ；
- 2) 将新数据 X_{new} 通过离线阶段得到的解混矩阵 W 得到新的高阶特征信息特征数据 S_{new}

$$S_{new} = W^T \bar{X}_{new}^T$$

- 3) 将 S_{new} 输入离线阶段训练好的 DRNN 深度循环神经网络中当输出的故障指标数据大于 0.5，则表示当前故障，当输出的故障指标数据小于 0.5 则表示当前正常；
- 4) 当 DRNN 深度循环神经网络预测结果为正常时，需要进行二次监测：首先计算数据 X_{new} 的残差 E_{new} ，如下式所示：

$$E_{new} = \bar{X}_{new} - W S_{new}$$

其中W为离线阶段得到的解混矩阵；

- 5) 计算当前采样时刻k的监控统计量 I_k^2 和 SPE_k ，如下式所示：

$$I_k^2 = S_{new}' S_{new}$$

$$SPE_k = E_{new}' E_{new}$$

- 6) 将上述步骤得到的监控统计量 I_k^2 和 SPE_k 与离线监测阶段步骤 6) 得到的控制限 I_{limit}^2 和 SPE_{limit} 进行比较，若上述两个指标中其中任意一个指标超限就认为发生故障并报警；否则即认为是正常；
- 7) 将故障数据按照离线步骤 5 所述增加故障标签，并加入 DRNN 的训练数据库，利用更新后的训练数据再次训练 DRNN 网络，用于不断学习新的故障信息，从而更加准确的进行监测。
- 8) 根据权利要求 1 所述的故障监测方法，其特征在于：DRNN 深度循环神

经网络的损失函数为交叉熵损失函数。

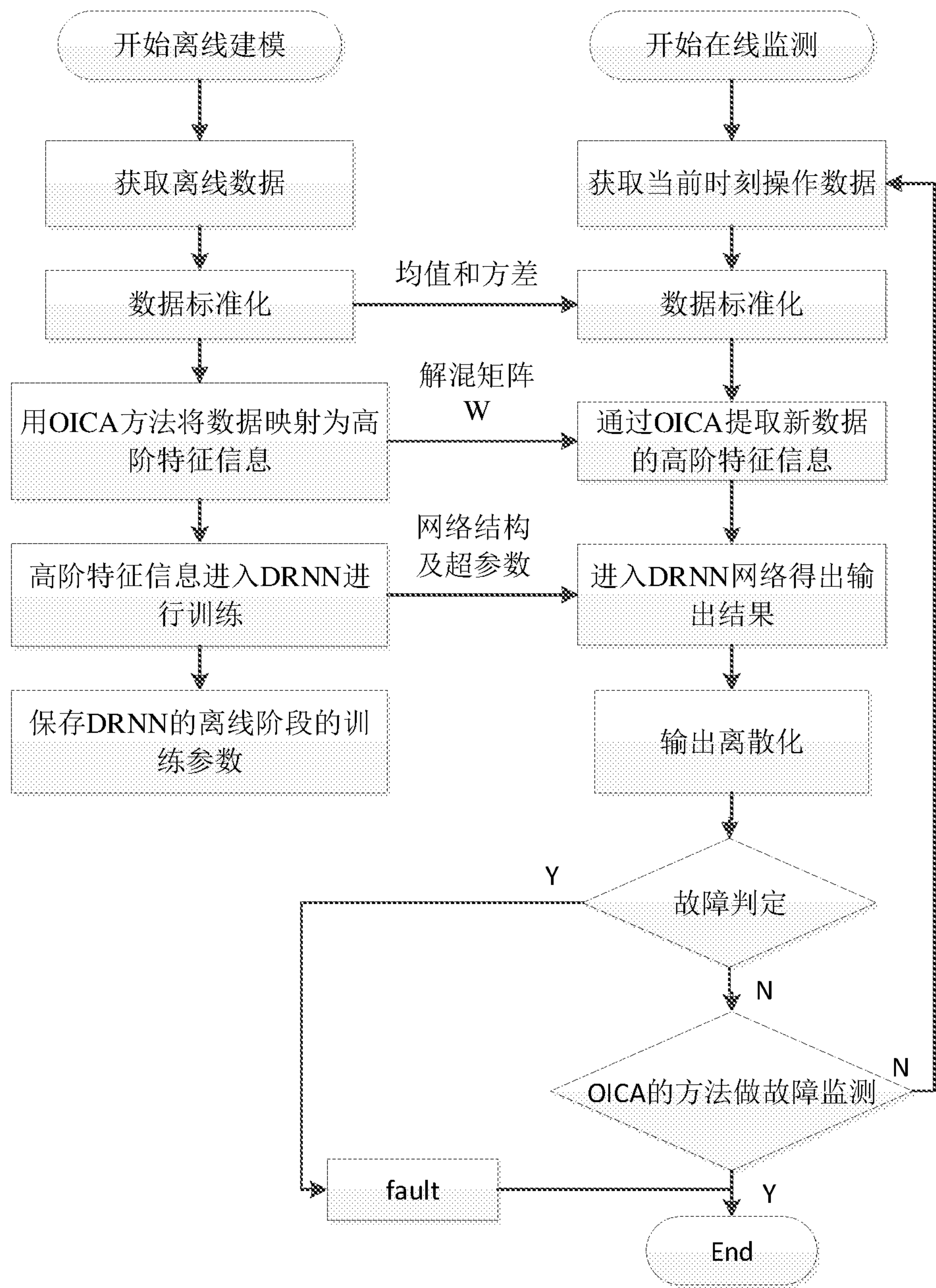


图 1

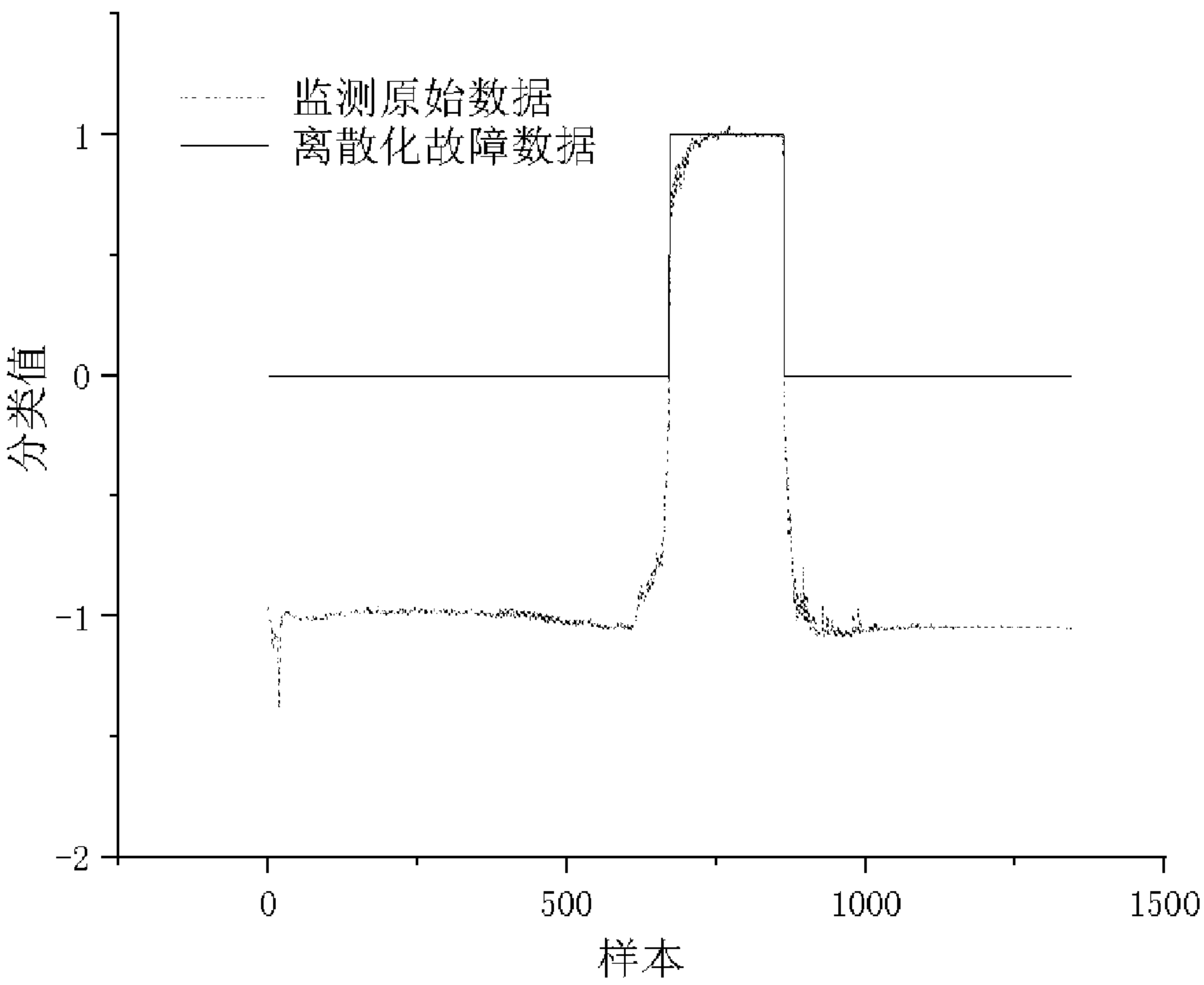


图 2

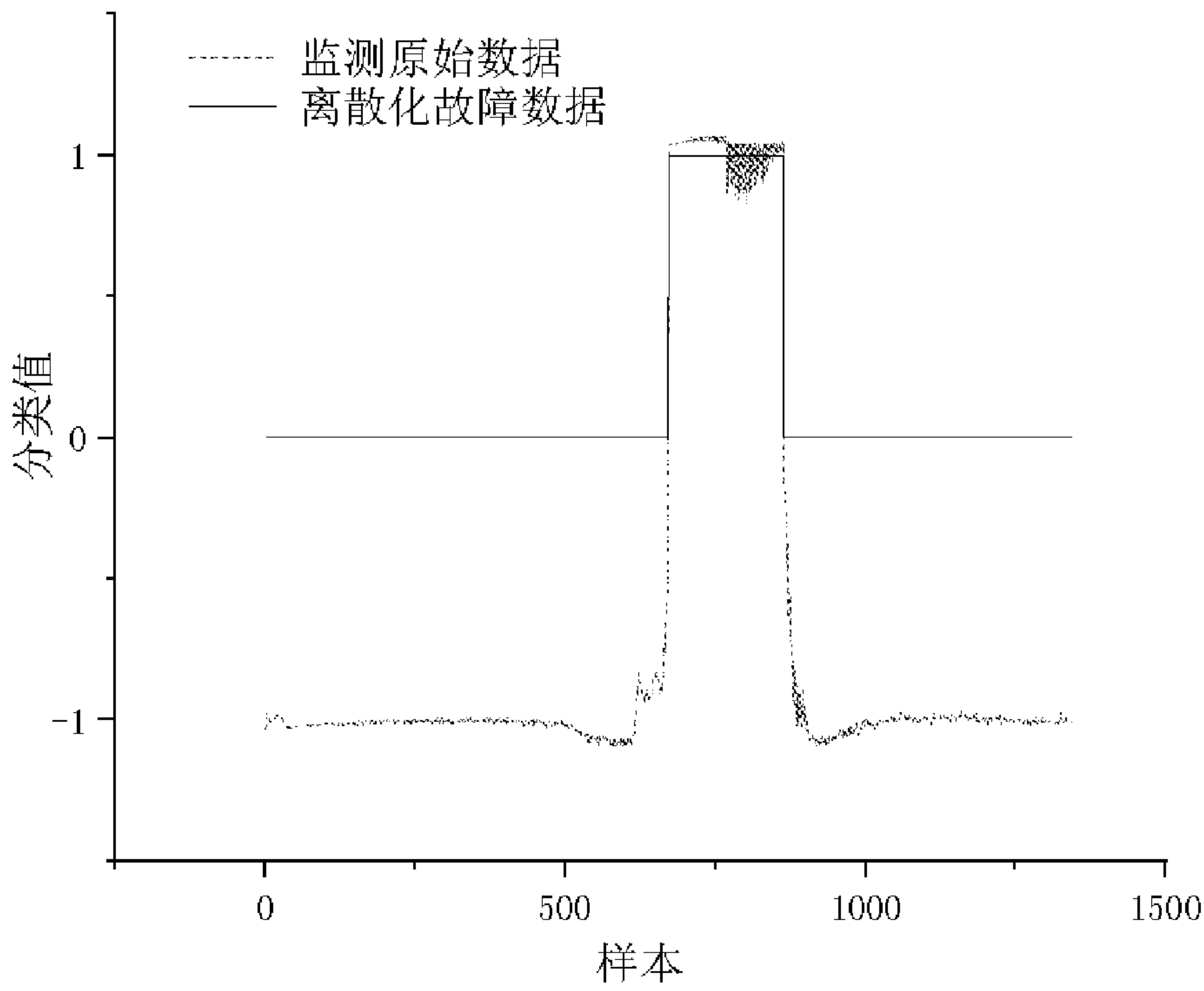


图 3

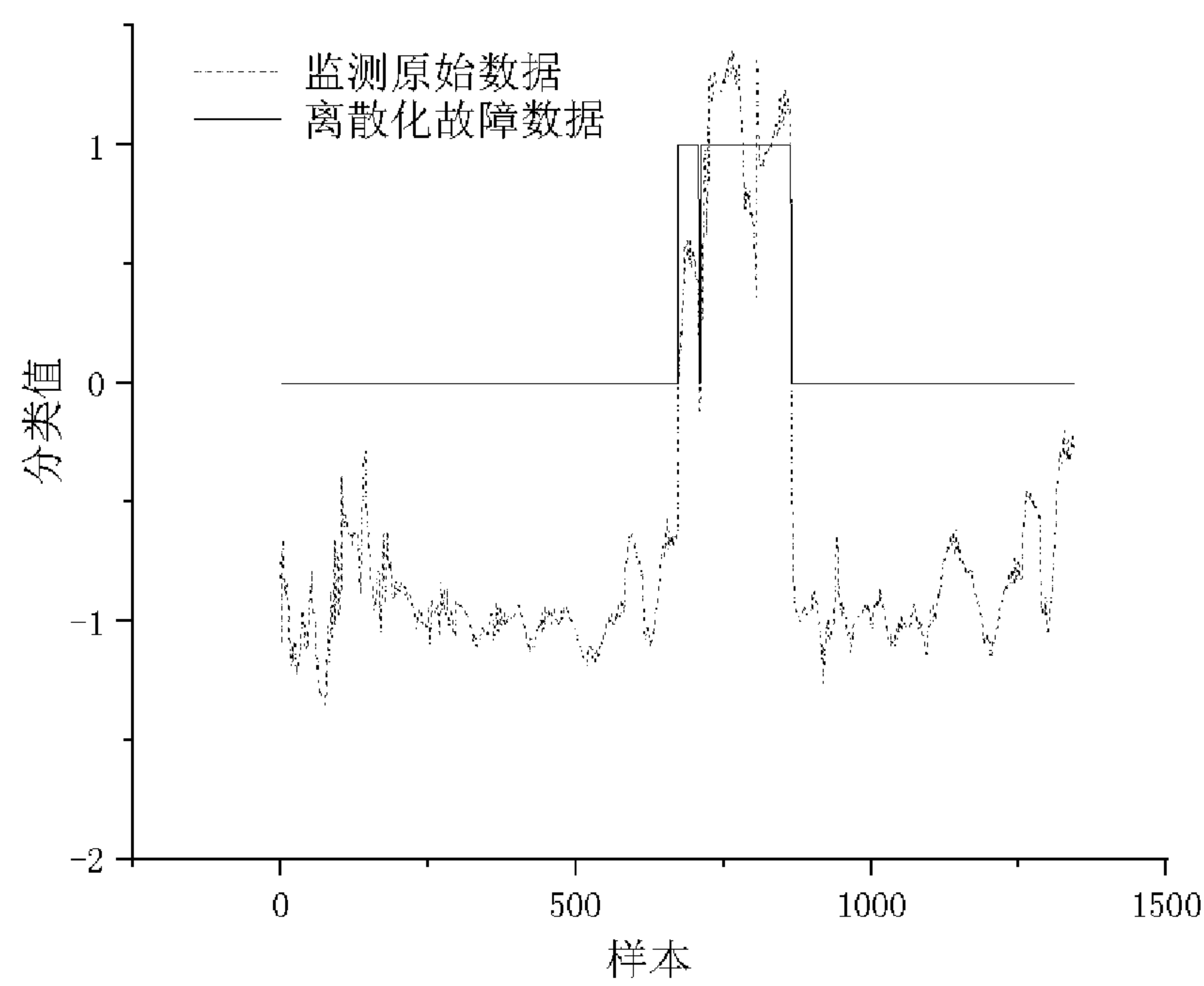


图 4

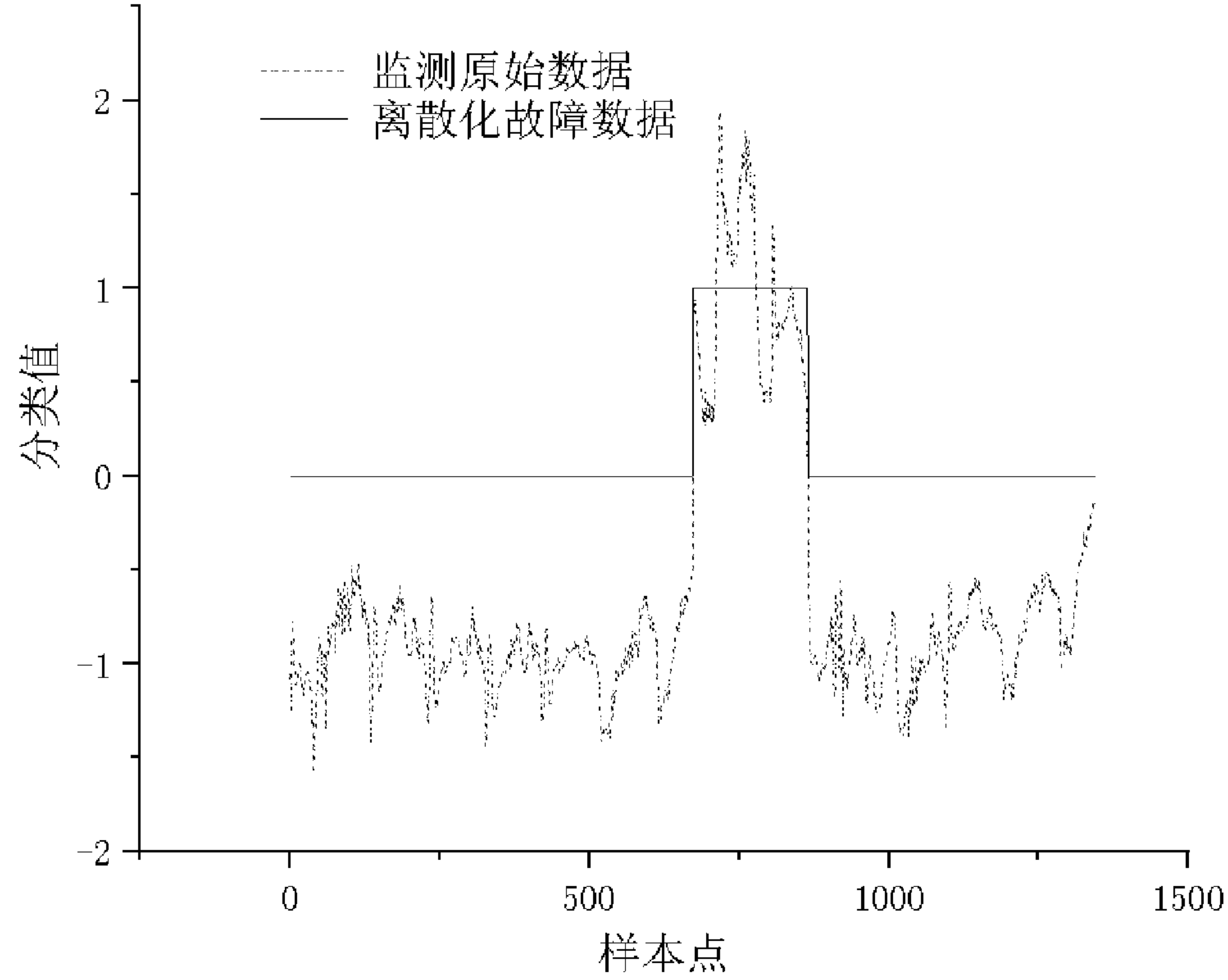


图 5

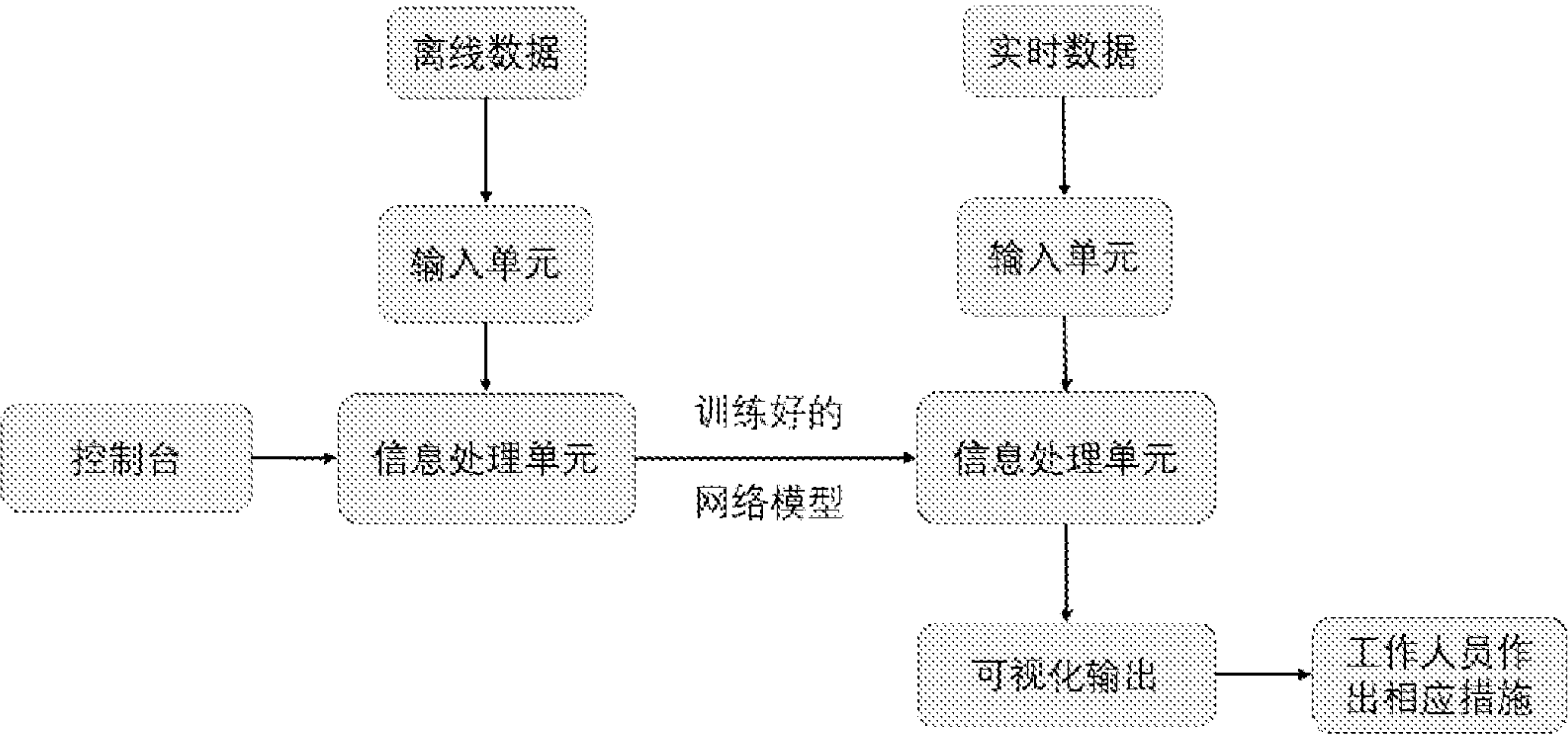


图 6

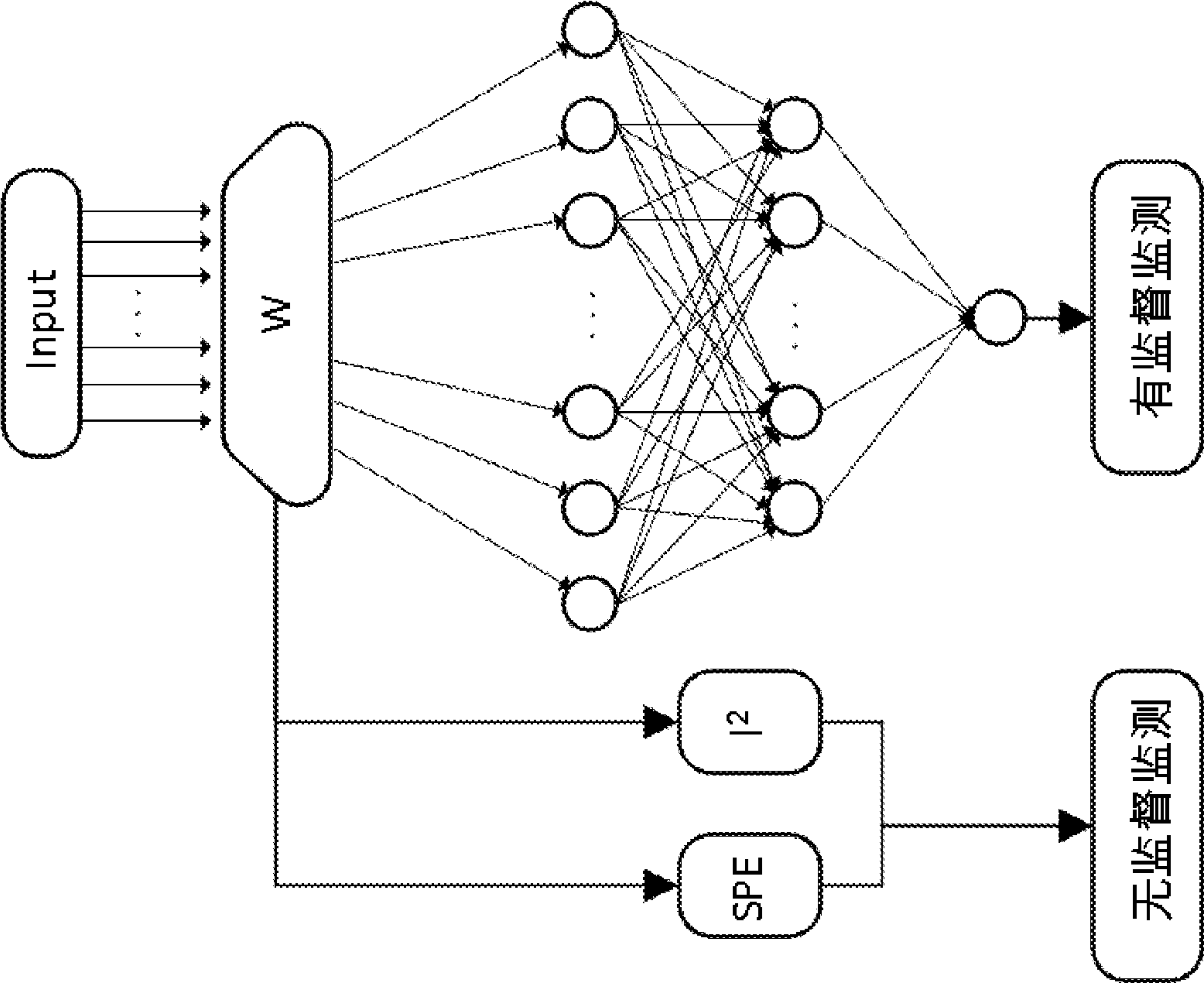


图 7

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2019/125888

A. CLASSIFICATION OF SUBJECT MATTER

G01N 33/18(2006.01)i

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

G01N 33/-

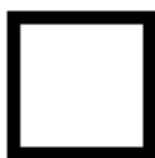
Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

CNPAT, CNKI, WPI, EPODOC: 高阶信息增强, 循环神经网络, 故障, 监测, 污水, 模型, 离线, 在线, 训练, 特征, over complete independent component analysis, recurrent neural network fusion model, wastewater, treatment, failure, monitor, training, offline, online

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	CN 105740619 A (SOUTH CHINA UNIVERSITY OF TECHNOLOGY) 06 July 2016 (2016-07-06) description, paragraphs 1-83, and figures 1-4	1,
A	CN 110088619 A (BL TECHNOLOGIES INC.) 02 August 2019 (2019-08-02) entire document	1
A	CN 107741738 A (CHONGQING HUALV ENVIRONMENTAL PROTECTION TECHNOLOGY DEVELOPMENT CO., LTD.) 27 February 2018 (2018-02-27) entire document	1
A	CN 106056127 A (JIANGNAN UNIVERSITY) 26 October 2016 (2016-10-26) entire document	1
A	CN 110119579 A (BEIJING UNIVERSITY OF TECHNOLOGY) 13 August 2019 (2019-08-13) entire document	1,
A	JP 3301428 B2 (OGAWA ENVIRONMENTAL RES. INST. INC.) 15 July 2002 (2002-07-15) entire document	1



Further documents are listed in the continuation of Box C.



See patent family annex.

* Special categories of cited documents:	“T” later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
“A” document defining the general state of the art which is not considered to be of particular relevance	“X” document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
“E” earlier application or patent but published on or after the international filing date	“Y” document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
“L” document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	“&” document member of the same patent family
“O” document referring to an oral disclosure, use, exhibition or other means	
“P” document published prior to the international filing date but later than the priority date claimed	

Date of the actual completion of the international search

25 August 2020

Date of mailing of the international search report

16 September 2020

Name and mailing address of the ISA/CN

China National Intellectual Property Administration (ISA/
CN)
No. 6, Xitucheng Road, Jimenqiao, Haidian District, Beijing
100088
China

Facsimile No. (86-10)62019451

Authorized officer

Telephone No.

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.
PCT/CN2019/125888

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)			Publication date (day/month/year)
CN	105740619	A	06 July 2016	CN	105740619	B	12 June 2018
CN	110088619	A	02 August 2019	EP	3552013	A1	16 October 2019
				CA	3049807	A1	18 April 2019
				WO	2019071384	A1	18 April 2019
				BR	112019017301	A2	22 April 2020
CN	107741738	A	27 February 2018	None			
CN	106056127	A	26 October 2016	None			
CN	110119579	A	13 August 2019	None			
JP	3301428	B2	15 July 2002	JP	2001255319	A	21 September 2001

国际检索报告		国际申请号 PCT/CN2019/125888
A. 主题的分类 G01N 33/18(2006.01)i 按照国际专利分类(IPC)或者同时按照国家分类和IPC两种分类		
B. 检索领域 检索的最低限度文献(标明分类系统和分类号) G01N 33/- 包含在检索领域中的除最低限度文献以外的检索文献 在国际检索时查阅的电子数据库(数据库的名称, 和使用的检索词(如使用)) CNPAT, CNKI, WPI, EPDOC: 高阶信息增强, 循环神经网络, 故障, 监测, 污水, 模型, 离线, 在线, 训练, 特征, over complete independent component analysis, recurrent neural network fusion model, wastewater, treatment, failure, monitor, training, offline, online		
C. 相关文件		
类 型*	引用文件, 必要时, 指明相关段落	相关的权利要求
A	CN 105740619 A (华南理工大学) 2016年 7月 6日 (2016 - 07 - 06) 说明书第1-83段, 图1-4	1
A	CN 110088619 A (BL科技有限责任公司) 2019年 8月 2日 (2019 - 08 - 02) 全文	1
A	CN 107741738 A (重庆华绿环保科技发展有限公司) 2018年 2月 27日 (2018 - 02 - 27) 全文	1
A	CN 106056127 A (江南大学) 2016年 10月 26日 (2016 - 10 - 26) 全文	1
A	CN 110119579 A (北京工业大学) 2019年 8月 13日 (2019 - 08 - 13) 全文	1
A	JP 3301428 B2 (OGAWA ENVIRONMENTAL RES. INST. INC.) 2002年 7月 15日 (2002 - 07 - 15) 全文	1
☐ 其余文件在C栏的续页中列出。 ☒ 见同族专利附件。		
* 引用文件的具体类型: “A” 认为不特别相关的表示了现有技术一般状态的文件 “E” 在国际申请日的当天或之后公布的在先申请或专利 “L” 可能对优先权要求构成怀疑的文件, 或为确定另一篇引用文件的公布日而引用的或者因其他特殊理由而引用的文件(如具体说明的) “O” 涉及口头公开、使用、展览或其他方式公开的文件 “P” 公布日先于国际申请日但迟于所要求的优先权日的文件 “T” 在申请日或优先权日之后公布, 与申请不相抵触, 但为了解发明之理论或原理的在后文件 “X” 特别相关的文件, 单独考虑该文件, 认定要求保护的发明不是新颖的或不具有创造性 “Y” 特别相关的文件, 当该文件与另一篇或者多篇该类文件结合并且这种结合对于本领域技术人员为显而易见时, 要求保护的发明不具有创造性 “&” 同族专利的文件		
国际检索实际完成的日期 2020年 8月 25日		国际检索报告邮寄日期 2020年 9月 16日
ISA/CN的名称和邮寄地址 中国国家知识产权局(ISA/CN) 中国北京市海淀区蓟门桥西土城路6号 100088 传真号 (86-10)62019451		授权官员 赵云峰 电话号码 86-(10)-53962539

国际申请号
PCT/CN2019/125888

检索报告引用的专利文件			公布日 (年/月/日)	同族专利			公布日 (年/月/日)
CN	105740619	A	2016年 7月 6日	CN	105740619	B	2018年 6月 12日
CN	110088619	A	2019年 8月 2日	EP	3552013	A1	2019年 10月 16日
				CA	3049807	A1	2019年 4月 18日
				W0	2019071384	A1	2019年 4月 18日
				BR	112019017301	A2	2020年 4月 22日
CN	107741738	A	2018年 2月 27日	无			
CN	106056127	A	2016年 10月 26日	无			
CN	110119579	A	2019年 8月 13日	无			
JP	3301428	B2	2002年 7月 15日	JP	2001255319	A	2001年 9月 21日