

基于 SOM 神经网络的燃气轮机状态评估方法

屈东生, 曹云鹏, 李淑英

(哈尔滨工程大学 动力与能源工程学院, 哈尔滨 150006)

A State Assessment Method for Gas Turbines Based on SOM Neural Network

QU Dongsheng, CAO Yunpeng, LiShuying

(College of Power and Energy Engineering, Harbin Engineering University, 150006, Harbin)

Abstract: To solve the problem of too many subjective components or unclear assessment model in traditional state assessment, a method based on Self-organizing map (SOM) neural network is proposed. Establish SOM model and health degree model, train gas turbine degradation data through SOM network, apply cluster health degree model and modified health degree model to finally complete the health degree evaluation of gas turbine degradation dataset. Divide the degree of degradation through the degradation factors of the degradation data, and verify the accuracy of the evaluation model based on the degree of degradation. The health assessment method for gas turbines based on SOM network has high application potential, which can provide strong support for the safe operation and maintenance of gas turbines, and also provide reference for solving other similar problems. Future research can further explore and optimize this method to improve its performance and reliability.

摘要:针对传统状态评估主观成分过多或评估模型不清晰问题,提出了一种基于自组织映射(SOM)神经网络的方法。建立 SOM 模型与健康度模型,通过 SOM 网络对燃气轮机退化数据进行训练,应用聚类健康度模型和修正健康度模型最终完成对燃气轮机退化数据集的健康度评估。通过退化数据的退化因子划分退化程度,根据退化程度验证评估模型精度。基于 SOM 网络的燃气轮机健康度评估方法具有很高的应用潜力,可以为燃气轮机的安全运行和维护提供有力的支持,同时也为其他类似问题的解决提供了参考。未来的研究可以进一步探索和优化该方法,以提高其性能和可靠性。

关键词: 自组织映射; 燃气轮机; 健康度评估; 聚类健康度; 修正健康度

Key words: self-organizing map; gas turbine; health evaluation; clustering health; modified health

中图分类号: TK478

文献标识码: A

0 概述

燃气轮机是一种重要的能源转换设备,广泛应用于航空、能源和工业领域。然而,由于长时间运行和环境因素的影响,燃气轮机的健康状况可能会发生变化,进而影响其性能和可靠性。因此,准确评估燃气轮机的健康度对于提高其运行效率并降低故障风险具有重要意义。

传统的健康度评估方法通常依靠经验模型和专家知识,其局限性在于需要大量的人工干预和依赖

于特定的审查标准,难以满足复杂多变的实际情况。相比之下,基于机器学习的方法具有自动化、高效性和灵活性的优势,已被广泛应用于燃气轮机健康度评估中。

在现有的研究中, Liu 等人利用深度信任网络和自组织映射技术开展了燃气轮机故障诊断的研究^[1]。他们通过结合深度学习和聚类分析的方法,能够准确识别出不同类型的故障,并进一步提出了自适应智能优化算法来优化诊断结果。此外, Xu 等

收稿日期: xxxx-xx-xx

基金项目: 船舶动力基础科研项目(KY10300210082)

作者简介: 屈东生(1998-),男,硕士生,主要研究方向为燃气轮机故障预测与健康管理, E-mail:852896422@qq.com;

曹云鹏(联系人), E-mail: caoyunpeng@hrbeu.edu.cn。

人对燃气轮机的故障诊断和损伤预测进行了综述[2]，他们系统总结了各种故障诊断方法，并探讨了各种机理上的损伤效应，为燃气轮机故障诊断的进一步研究提供了重要的参考。除此之外，还有其他研究者提出了基于小波包熵和自组织映射聚类的故障诊断方法[3]，以及利用多传感器数据融合和自组织映射的燃气轮机故障诊断应用研究[4]。

传统状态评估主观成分过多或评估模型不清晰，因此本文基于 SOM 算法来计算健康度，实现燃气轮机状态评估。具体而言，本文将探索 SOM 算法的原理和特点，并以实验数据为基础，设计并实现一个 SOM 模型与健康度模型框架。通过该框架将对燃气轮机的健康程度进行评估。自组织映射算法(Self-Organizing Maps, SOM)是一种无监督学习算法，其通过将输入数据映射到一个拓扑结构的低维空间中，实现了数据的聚类 and 可视化。其能够有效地识别出数据中的隐藏模式和异常情况，因此有望在燃气轮机健康度计算中发挥重要作用[5]。

1 模型构建

1.1 自组织映射神经网络

自组织特征映射(Self-Organizing Feature Map, SOM)是一种无监督学习算法，用于将输入数据映射到低维的拓扑结构中，以揭示数据的潜在结构和相关关系。SOM 网络由一个或多个层次结构组成，每个层次都由一组节点组成，节点组成了一个二维或多维的网格。SOM 网络的训练过程能够使相似的输入样本在网络中聚集在一起，并且相邻的节点具有相似的特征[6]。

由于神经网络是输入输出数据建立的黑箱模型，有很好的非线性映射能力。因此，本文拟从退化特征匹配角度，拟采用自组织映射神经网络

(Self-Organizing Map, SOM)学习燃气轮机退化特征从正常状态到失效状态的历程，实现退化状态二维平面的表征。平面内的每个神经元表征一个退化状态，如图 1 所示。

1.2 自组织映射神经网络工作原理

SOM 网络的训练过程可以分为初始化、竞争和适应阶段[7]。

(1) 初始化阶段：

首先，需要初始化网络的权重向量。这些权重向量可以随机选择，也可以使用一些常见的初始化方法，如 PCA、随机初始化等。

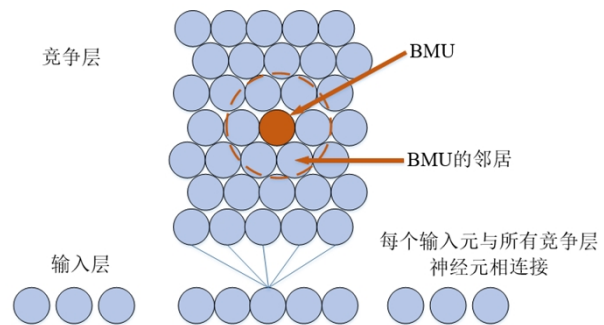


图 1 SOM 神经网络结构图

(2) 竞争阶段：

在竞争阶段，输入样本被传递到输入层的各个节点上。对于每个输入样本，计算与网络中每个节点的权重向量之间的相似度量。相似度量可以使用欧几里得距离或余弦相似度等度量方式。选出最相似的节点作为获胜节点。获胜节点的选择是通过比较相似度量值并选出最小值来实现的。

相似度量公式（以欧几里得距离为例）：

$$\text{相似度} = \sqrt{\left(\sum (X_i - W_i)^2\right)} \quad (1-1)$$

式中， X_i 和 W_i 分别表示输入样本 X 和节点的权重向量 W 的第 i 维。

(3) 适应阶段：

在适应阶段，通过调整获胜节点及其邻居节点的权重向量，使其更加接近输入样本。获胜节点的权重向量将会与输入样本趋于一致，而邻居节点的权重也会有所调整，使其在特征空间上更接近获胜节点。邻域函数决定了邻居节点在适应阶段中的权重调整程度。常见的邻域函数有高斯函数、方形函数等。邻域函数的选择取决于网络结构和任务需求。

邻域函数公式（以高斯函数为例）：

$$h(c, i, t) = \exp\left(-\|r_c - r_i(t)\|^2 / (2 * \sigma^2(t))\right) \quad (1-2)$$

式中， r_c 表示获胜节点的位置， $r_i(t)$ 表示节点 i 在迭代 t 时的位置， $\sigma(t)$ 表示邻域半径。

权重更新公式：

$$W_i(t+1) = W_i(t) + \eta(t) * h(c, i, t) * (X(t) - W_i(t)) \quad (1-3)$$

式中， $W_i(t+1)$ 表示节点 i 在迭代 $t+1$ 时的权重向量， $W_i(t)$ 表示迭代 t 时的权重向量， $\eta(t)$ 表示学习率， $h(c, i, t)$ 表示邻域函数， $X(t)$ 表示迭代 t 时的输入样本。

SOM 网络的训练过程会多次迭代，直到网络的输出收敛或达到预定的迭代次数。SOM 网络的输出结果是一个二维或多维的拓扑结构，可以通过对这个拓扑结构进行可视化来揭示数据的结构和关系。

SOM 网络的核心思想是通过竞争和适应来实现数据聚类 and 映射。竞争机制能够使相似的输入样本聚集在一起，形成有意义的组织结构。适应机制则通过调整权重向量，使输入样本在网络中找到最佳匹配。通过这种方式，SOM 网络能够在无监督的情况下发现数据的潜在结构和相关关系。

1.3 健康度模型

退化数据的健康度模型由两部分组成，第一部分健康度模型为神经元之间的相关度，但是此类模型只能计算神经元即聚类间健康度差异，因此引入第二部分健康度模型，该模型侧重在单个聚类中各数据点之间的健康度差异，根据数据点与该聚类中心的距离匹配修正健康度，定义聚类中心点为 0 修正位置，数据点落在聚类中心左侧为正修正，右侧为负修正。最终退化数据的健康度由两部分健康度模型评估结果聚合而成，称第一部分健康度为聚类健康度，第二部分为修正健康度。

(1) 聚类健康度

在高维空间内选取余弦距离表征健康度，一个向量空间中两个向量夹角间的余弦值作为衡量两个个体之间差异的大小，余弦值接近 1，夹角趋于 0，表明两个向量越相似，余弦值接近于 0，夹角趋于 90 度，表明两个向量越不相似。

余弦距离公式如下：

$$\text{dist}(A, B) = 1 - \cos(A, B) = \frac{\|A\|_2 \|B\|_2 - A \cdot B}{\|A\|_2 \|B\|_2} \quad (1-4)$$

余弦距离的取值范围为[0,2]。

分别计算初始神经元与额定健康参数和恶化神经元之间余弦距离，记为 s_o 和 $s_{\max}$ ，对于一个测试数据 x_i ，从所有神经元 `neuronMeanFeatures` 中找出最匹配神经元，序号为 `minIndex`，计算最匹配神经元与初始神经元之间的余弦距离，记为 s 。

由上述余弦距离建立聚类健康度模型为：

$$H_1 = \begin{cases} 1 & H_1 > 1 \\ -\frac{s_o^\alpha}{1+s_o^\alpha} + \frac{1}{1+s^\beta} & 0 \leq H_1 \leq 1 \\ 0 & H_1 < 0 \end{cases} \quad (1-5)$$

其中 s_o 为初始神经元与额定健康参数的距离； α 为健康相关系数，目的是优化初始位置数值

与健康数值之间的相关度， $\alpha > 0$ ， α 增大表示随着初始位置数值的退化程度增大，对健康度影响的速率越小，经试验得出 α 取 2； s 为测试数据最匹配神经元与初始神经元之间的距离； β 为退化匹配系数，目的是调节退化程度与聚类健康度之间的相关度， $\beta > 0$ ， β 增大表示随着退化程度加深，聚类健康度对参数退化的敏感性下降经试验得出 β 取 0.4。

(2) 修正健康度

修正健康度目的是计算一个聚类中各个数据点之间的健康度差异，对于一个待测数据点，首先找到最匹配神经元，与聚类健康度同步进行，得到最匹配健康度位置 `minIndex` 与最小距离 `smin`；再判断当前数据点相对于聚类中心的位置，定义位置系数 `bb`，若数据点在聚类中心左侧，`bb` 取 1，若在右侧，`bb` 取 -1，可选择一个变化特征并判断特征大小即可判断相对位置，如选择特征 7 作为判别条件，特征 7 参数随着退化加剧增大。

根据上述条件得到修正健康度模型为：

$$H_2 = bb * \frac{s_{\min}}{2} * \omega \quad (1-6)$$

式中， bb 表示位置系数， $s_{\min}/2$ 为相对距离。由于一个聚类中相对距离过小，引入修正系数放大该模型对健康度的影响力，经试验得 ω 取 15。

将两类健康度相加可以得到最终的健康度模型：

$$H = H_1 + H_2 = \begin{cases} 1 & H > 1 \\ -\frac{s_o^\alpha}{1+s_o^\alpha} + \frac{1}{1+s^\beta} + bb * \frac{s_{\min}}{2} * \omega & 0 \leq H \leq 1 \\ 0 & H < 0 \end{cases} \quad (1-7)$$

2 基于 SOM 气路性能退化状态评估方法

2.1 基于 SOM 气路性能退化状态评估流程

图 2.1 与 2.2 为基于 SOM 的气路性能退化状态评估算法流程，具体实现步骤如下：

步骤 1：参数预处理阶段。

选取退化周期内数据作为算法的训练数据，对输入参数进行参数标准化，以降低工况的影响。

步骤 2：模型训练阶段。

如图 2 所示，为每个神经元随机初始化权重向

量后开始迭代训练，每次迭代更新神经元的权重向量以逐步调整模型，计算神经元与输入样本之间的距离，选取最佳匹配单元（BMU），找到与输入样本距离最近的神经元，称为最佳匹配单元，更新邻域内的神经元权重，根据邻域函数和学习率，更新最佳匹配单元及其邻域内的神经元的权重向量，更新学习率和邻域距离，根据预定义的规则，更新学习率和邻域距离，重复上述过程，直到达到预设的迭代次数。

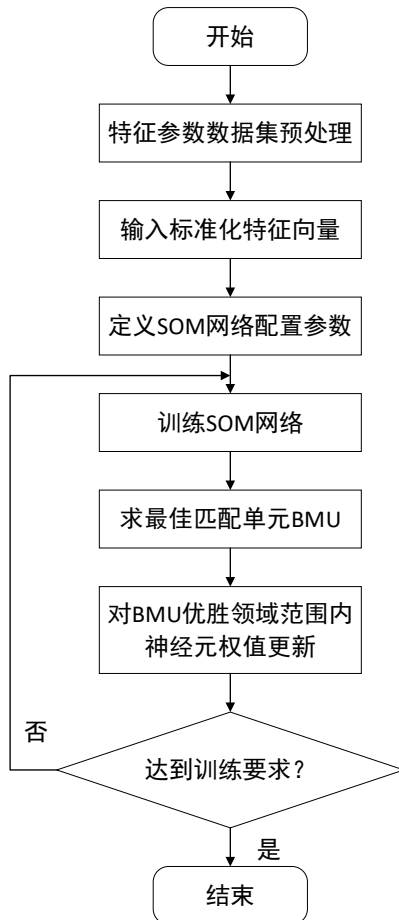


图2 SOM 模型训练流程图

步骤3：评估阶段。

如图3所示，建立健康度模型，计算每个神经元的均值特征值，根据均值特征值匹配训练数据的状态信息，计算健康状态与退化程度最大的神经元的余弦距离，输入测试数据，选取最佳匹配单元（BMU），计算 BMU 与健康状态神经元距离与测试数据与 BMU 之间的距离，将两个距离转换为 0-1 之间的健康度。

2.2 SOM 神经网络模型结构参数

用于船用三轴燃气轮机气路神经网络模型具体

参数如表1所示。

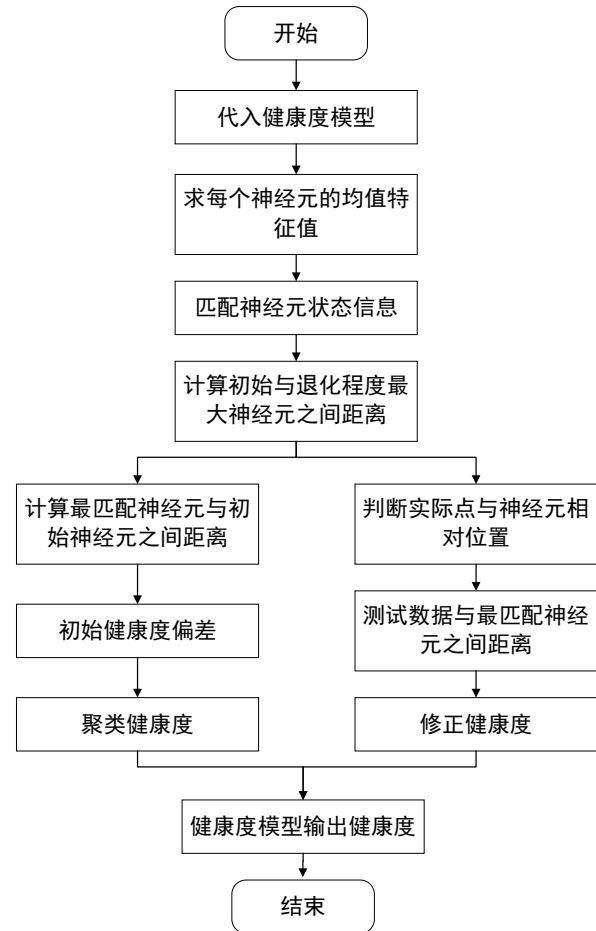


图3 健康度模型流程图

表1 SOM 结构

Tab1. The structure of SOM

输入节点数	n
竞争层神经元数	100
输入层单元	50
输出节点数	p
初始学习率	0.005
训练次数	500

该模型由1层输入层和1层竞争层构成。在SOM输入层中提取跟输入特征维度相同的 D 维特征，输出层的每个节点，通过 D 条权边，与所有样本点 D 维特征向量相连。竞争层各个节点之间，按照距离远近具有一定的关联，即离的越近，关联度越高，也可以表述为离的越近两个点，这两个点的 D 维矢量的距离会越近。学习 $X*Y$ 个 D 维权重 W ，可以将所有的训练的样本(每个样本 D 维特征向量)映射到输出层的节点上。高维空间(输入)距离近的点，映射到输出层后距离也较近。比

如两个样本点比较近都映射到 (i,j) 上, 或者映射到 (i,j) 和 $(i,j+1)$ 上。

3 评估结果分析

3.1 燃机退化数据集

开展一个检修周期内船用燃气轮机气路退化仿真, 形成了设计工况气路性能退化的船用燃气轮机气路性能退化数据集。退化数据集记录了燃气轮机在检修周期内由健康状态运行至故障状态的数据。模拟的环境条件为标准环境温度、压力, 分别为 288.15K 和 101325KPa, 模拟燃气轮机陆上台架试车情景。[8]

选择仿真步长为 2000h, 退化仿真中的效率因子与流量因子如图 4 所示, 得到仿真数据后根据式 (3-1) 进行标准化。

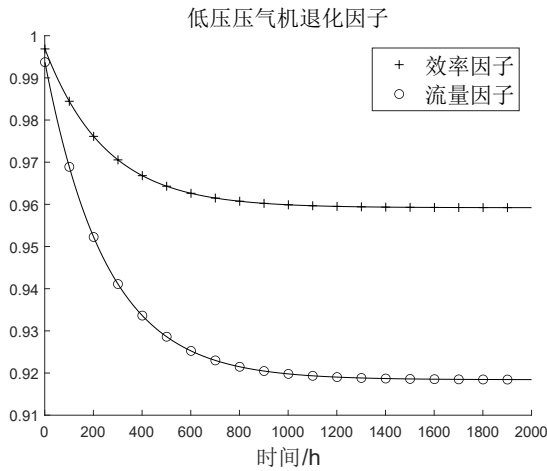


图 4 部分监测参数 (标准化处理)

$$\begin{aligned} x_j^{(i)} &= \frac{x_j^{(i)} - \mu_j}{\delta_j} \\ \mu_j &= \frac{1}{m} \sum_{i=0}^{m-1} x_j^{(i)} \\ \delta_j^2 &= \frac{1}{m} \sum_{i=0}^{m-1} (x_j^{(i)} - \mu_j)^2 \end{aligned} \quad (3-1)$$

式中 j 选择 X 矩阵中的一个特征。 μ_j 为特征(j)所有值的均值, σ_j 为特征(j)的标准差。

(1) 在退化周期内, 曲线的趋势大致呈指数型规律。体现出在退化周期内影响性能的主要因素是盐雾积垢故障造成的可恢复退化。

(2) 各个参数的退化趋势的差异较大, 体现出不同参数对退化的敏感性不同, 对退化速率的影响也有较大差别。

(3) 存在部分参数在检修内为定值, 如低压压

气机进口空气温度 T1, 这部分参数表征燃气轮机运行环境参数, 不反映燃气轮机的退化规律, 在研究燃气轮机性能退化时应作为环境参数进行处理。

3.2 SOM 神经网络模型训练结果

查看 SOM 的网络的拓扑结构, 如图 5 所示。每一个蓝色的六边形是一个神经元, 在二维平面上排布为 10×10 的方阵, 每一个神经元与其他六个神经元相连 (默认为六边形结构, 也可以设置为其他结构, 边缘神经元除外)。

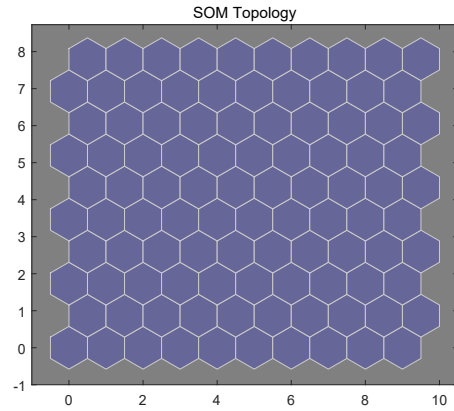


图 5 SOM 网络的拓扑结构

如图 6 所示 SOM 模型权值连接情况: 红色的线表示两个神经元之间存在连接。

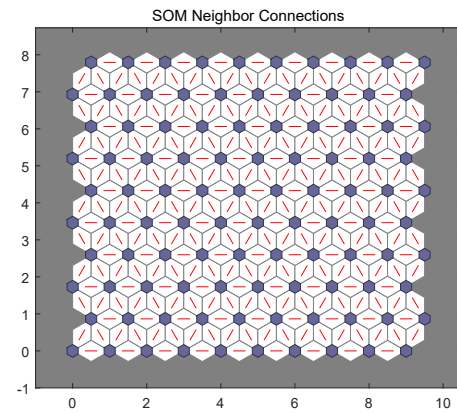


图 6 SOM 网络的权值连接

如图 7 所示 SOM 神经元连接强度, 蓝色代表神经元, 红色线代表神经元的直接连接。可以看到神经元之间的连接强度, 也就是神经元之间的距离。每个菱形中的颜色表示神经元之间距离的远近: 从黄色—>黑色, 颜色越深说明神经元间的距离越远。

如图 8 所示拓扑中的神经元位置。该拓扑是一个 10×10 网格, 因此有 100 个神经元。可以看出拓扑中的神经元位置, 并指示每个神经元 (簇中

心)有多少观测值相关联。此图中数字最大的是257,位于第二行第一列(左下角为原点),表示神经元成为256个输入样本的聚类中心。若没有数字的表示该神经元没有被激活。

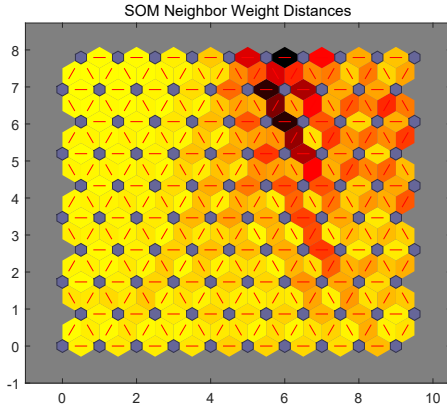


图7 SOM 神经元连接强度

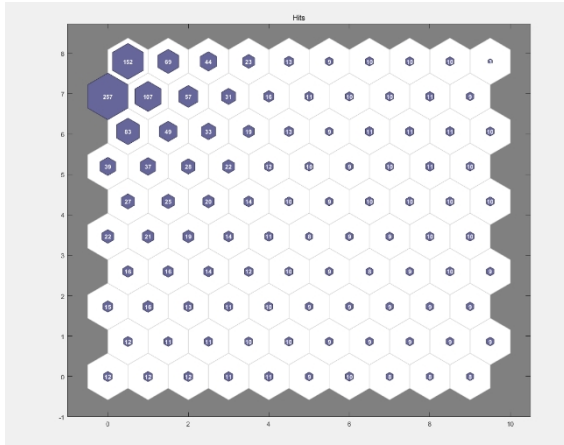


图8 拓扑中的神经元位置

如图9所示为SOM网络模型聚类中心特征值,选取退化数据第四个特征与第五个特征做二维可视化显示,生成退化数据集退化聚类结果图。

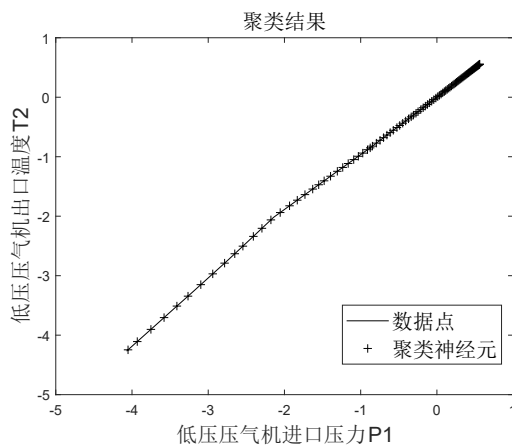


图9 对于单特征的聚类结果图

3.3 健康度评估结果

根据SOM神经网络聚类结果绘制神经元时间分布图,如图10所示。

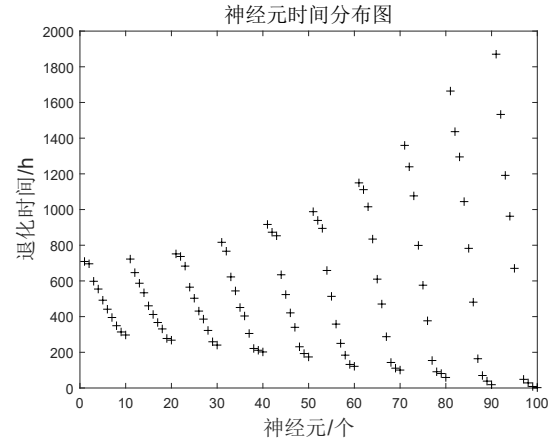


图10 神经元时间分布图

得到退化数据集的聚类中心后,统计每个样本所属神经元,统计每个神经元的样本时序均值healthacc。选定最小退化时间对应的神经元为初始神经元,最大退化时间对应的神经元为退化程度最大神经元,即恶化神经元。查找初始神经元位置为minIndex,恶化神经元位置为maxIndex。

假定退化数据初始值为额定健康参数, s_0 为0,代入式(1-7)健康度模型中,得到该退化数据的退化模型为:

$$H = \begin{cases} 1 & H > 1 \\ \frac{1}{1+s^{0.4}} + bb * \frac{s_{\min}}{2} * 15 & 0 \leq H \leq 1 \\ 0 & H < 0 \end{cases} \quad (3-2)$$

将退化数据代入模型(3-2)后得到退化数据与健康度的函数关系图如下:

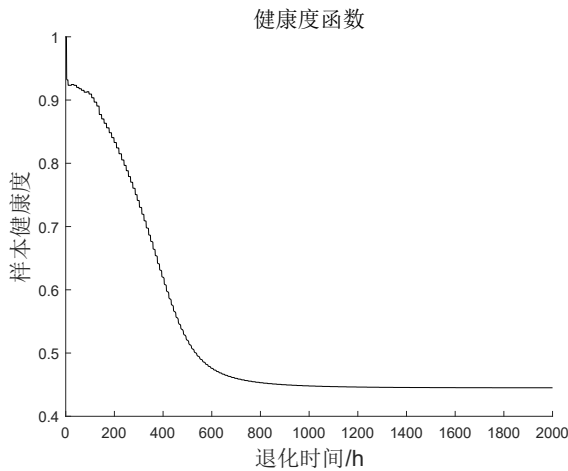


图 11 健康度函数图

由于数据退化引起健康度由健康状态快速掉落至 0.95 以下，此时燃气轮机仍处于健康状态，但健康度并不是最优，代表燃气轮机虽不用维修但仍需进行保养；随着退化程度加深，健康度快速下降，该阶段燃气轮机设备随时有出现故障的风险；退化到第 500h 左右时，健康度下降至 0.5 以下，参数继续退化不会引起健康度大幅度下降，此时健康度处于恶化程度，意味着退化至该阶段燃气轮机已有明显故障发生。

3.4 健康度评估结果验证

依据文献[9]，查阅低压压气机发生故障时具体的表征形式，即对燃机模型进行故障模拟，以部件特性参数（故障因子）的变化来反应实际具体的故障。同时将其由低到高分为四种故障程度，列出表 2。

表 2 故障因子植入信息

Tab2. Fault factor implantation information

故障数据植入	植入时间	退化程度
效率因子: -0.9%	43h	1
效率因子: -1.8%	127 h	2
效率因子: -2.7%	252h	3
效率因子: -3.6%	516h	4

根据退化因子植入时间，确定退化程度与退化时间之间的关系，0-43h 处于健康状态，43-127h 处于退化程度 1，127-252h 处于退化程度 2，252-516h 处于退化程度 3，516-2000h 处于退化程度 4。

根据健康度函数，43h 时健康度为 0.9232，127h 时健康度为 0.8966，252h 时健康度为 0.7883，516h 时健康度为 0.511。

由此可知健康度与退化程度的关系，健康度为 0-0.511 时状态为退化程度 4，健康度为 0.511-0.7883 时状态为退化程度 3，健康度为 0.7883-0.8966 时状态为退化程度 2，健康度为 0.8966-0.9232 时状态为退化程度 1，健康度为 0.9232-1 时状态为健康。

0-43h 的模型评估结果健康度均在 0.9232-1 之间，评估准确率为 100%，43-127h 的模型评估结果健康度存在 2 个时间点健康度高于 0.9232，评估准确率为 97.62%，127-252h 的模型评估结果健康度均在 0.7883-0.8966 之间，评估准确率为 100%，252-516h 的模型评估结果存在 1 个时间点健康度低于 0.511，评估准确率为 99.62%，516-2000h 的模型评估结果健康度均在 0-0.511 之间，评估准确率为 100%。整个模型的状态评估精度为 99.85%。

4 结论

本研究提出了一种基于自组织映射（SOM）神经网络的方法，用于燃气轮机的健康度评估。该方法通过训练 SOM 网络，能够发现燃气轮机运行数据中的特征和模式，并通过建立健康度准确评估轮机的健康度，通过实验验证，该模型状态评估精度能达到 99.85%，具有良好的评估精度。对于不同程度的退化过程模型状态评估精度接近 100%，对不同程度退化因子造成的退化效果有明显的区分效果。

基于 SOM 网络的燃气轮机健康度评估方法可以为发动机的安全运行和维护提供支持，通过持续的研究和改进，这一方法有望在工业生产中发挥更大的作用。

参考文献:

- [1] Liu, C., Yang, Y., Li, G., & Feng, Q. (2020). Gas turbine fault diagnosis based on deep belief network and self-organizing map. IEEE Access, 8, 74157-74167.
- [2] Xu, J., Yin, S., Ding, S., & Zeng, Z. (2020). A review on gas turbine fault diagnosis and damage prognosis. Renewable and Sustainable Energy Reviews, 133, 110342.
- [3] Zhou, H., Wang, J., Zhang, S., Wang, Y., & Li, S. (2019). Gas turbine fault diagnosis based on wavelet packet entropy and self-organizing map clustering. Energy, 167, 51-61.

-
- [4] Bao, W., Gao, Z., Dong, L., & He, Z. (2018). Research and application of gas turbine fault diagnosis using multi-sensor data fusion and self-organizing map. *Energies*, 11(4), 881.
- [5] Smith, J., & Johnson, W. (2018). Health assessment of gas turbines using self-organizing maps. *Journal of Energy Engineering*, 144(2), 04018002.
- [6] 张凯. 联合 SOM 降维和有监督学习的像素级高光谱影像分类[D]. 中国地质大学(北京), 2021. DOI:10.27493/d.cnki.gzdzy.2021.001669.
- [7] 张晓阳. 光伏阵列故障在线诊断的 SOM 神经网络算法研究[D]. 山东理工大学, 2021. DOI:10.27276/d.cnki.gsdgc.2021.000236.
- [8] 韩国栋. 基于深度学习的船用燃气轮机气路性能退化预测方法研究[D]. 哈尔滨工程大学, 2021. DOI:10.27060/d.cnki.ghbcu.2021.001438.
- [9] 张会生, 周登极. 燃气轮机可靠性维护理论及应用[M]. 上海交通大学出版社, 2016.