

基于多维时间序列的数控机床状态预测方法研究

李海, 王伟*, 黄璞, 杜丽, 张心羽

(电子科技大学 机械电子工程学院, 四川 成都 611731)

摘要:随着数控机床结构复杂化以及运行状态数据呈现多样性、时序性的特点,为了有效解决数控机床未来状态难以准确预测的难题,提出一种基于多维时间序列的数控机床状态预测方法。首先,采用OPC(OLE for process control)技术进行数控机床状态数据采集,结合Min-max标准化和自回归移动平均模型完成了数据预处理,建立了多维时间序列状态模型及度量模型,采用特征向量、特征趋势距离标示状态模型,并利用差异度进行多维时间序列状态匹配分析。其次,通过建立时间窗口滑动模型,利用时间窗口长度和滑动时长获取数控机床历史状态集合,进一步提出基于窗口滑动的多重匹配技术,利用 β -耦合相似度量标准寻找与当前状态矩阵相似度最大的历史状态集合,并根据相似性阈值得到最优滑动时长和预测时长。然后,采用密度空间聚类算法进行状态序列分析,得到了表征机床当前时刻状态的最佳历史状态矩阵,并以此状态的下一时刻作为预测状态。最后,对数控机床主轴四项参数开展了数控机床状态预测实验,通过状态序列相似性分析得到最佳预测时长为24 s,滑动单位为2 s,并利用状态序列聚类分析完成状态序列匹配。预测结果表明,基于多维时间序列的状态预测方法的最大误差、平均误差、均方误差和相对误差均低于传统的AR预测模型,验证了所提出的状态预测方法的有效性和准确性。

关键词:数控机床; 多维时间序列; 多重匹配; 状态预测

中图分类号:TH132.4

文献标志码:A

文章编号:2096-3246(2018)01-0187-09

State Prediction Method Research in NC Machine Tool Based on Multidimensional Time Series

LI Hai, WANG Wei*, HUANG Pu, DU Li, ZHANG Xinyu

(School of Mechatronics Eng., Univ. of Electronic Sci. and Techno. of China, Chengdu 611731, China)

Abstract: With the structural complexity of machine tool and the diversity and time-sequence characteristic of state data, in order to solve the problem that the future state of machine tool is difficult to accurately predict, a novel state prediction method based on multidimensional time series was proposed. Firstly, the ole for process control (OPC) technology was used to collect data of machine tool, and the Min-max normalization and autoregressive moving average model were adopted to data preprocessing. The state and evaluation models of multidimensional time series were established. Meanwhile, the feature vector and trend distance were also used to represent state model, and then the state match of multidimensional time series was completed by difference degree analysis. Secondly, through constructing the time sliding window technique, the historical state sets of machine tool were obtained by the length of time window length and sliding. On this basis, multiple matching technique based on window sliding was developed, and then the β -coupling similarity metrics was also used to find a set of historical states that were the most similar to the current state matrix. According to the similarity threshold, the optimal sliding time and prediction time were obtained. Further, the density-based spatial clustering algorithm was adopted to perform state series analysis, and the optimal historical sate matrix which can represent the current state of machine tool was obtained, and then the next state was regarded as prediction state. Finally, the state prediction experiments were carried out for four parameters of machine tool spindle. The optimal prediction time and sliding unit were 24 s and 2 s respectively, and then the state-sequence matching was completed by using state-sequence clustering analysis. The prediction results showed that the maximum error, mean error, mean

收稿日期:2017-06-07

基金项目:国家科技重大专项资助(2015ZX04001002)

作者简介:李海(1991—),男,博士生.研究方向:数控机床故障诊断与预警. E-mail: lheight@163.com

* 通信联系人 E-mail: wangwhit@163.com

网络出版时间:2018-01-17 22:58:00

网络出版地址: <http://kns.cnki.net/kcms/detail/51.1773.TB.20180117.2258.003.html>

<http://jsuese.ijournals.cn>

<http://jsuese.scu.edu.cn>

square error and relative error of the matrix and vector state prediction approach based on multidimensional time series were lower than those of traditional AR prediction model, which also verified the effectiveness and accuracy of state prediction method.

Key words: machine tools; multidimensional time series; multiple matching; state prediction

数控机床作为一种复杂的机电设备,其结构复杂,故障隐蔽性较高,难以发觉^[1]。数控机床状态预测有利于及时发现数控机床健康状况,进而对机床进行检修,有效地控制数控机床受损程度,降低生产成本。因此,对数控机床状态预测的研究具有重要的意义。

传统数控机床状态预测方法有神经网络、支持向量机、粒子群算法等。Zhang等^[2]提出串行灰色神经网络和并联灰色神经网络方法预测数控机床的热误差。Sheng等^[3]提出Back-Propagation神经网络模型预测数控机床刀具的状态监测和刀具磨损量的预测。杜柳青等^[4]通过构建一种小波神经网络模型,实现对数控机床运动精度的预测。石宁等^[5]结合遗传算法和支持向量机的优势,建立改进支持向量机预测模型对机械状态进行振动趋势预测。许志军^[6]提出采用粒子群算法优化支持向量机进行数控机床状态预测。针对传统的预测方法,神经网络存在容易陷入局部极小值、收敛速度太慢等问题,支持向量机中训练参数的选择对其预测性能有很大影响,且没有考虑机床状态时序性,限制传统预测方法的实用性和准确性。

时间序列预测是目前应用最广的状态预测方法^[7],针对数控机床状态的时序性,张海波等^[8]使用随机事件序列法建立数控机床故障频率预测模型,以保证故障频率预测值的准确度。李永祥等^[9]提出了采用时间序列分析进行机床热误差建模方法,利用实测的热误差序列进行时序分析实时预测。杜柳青等^[10]提出混沌相空间重构理论的数控机床运动精度演化分析方法。为解决时间序列的相似性匹配问题,Agrawal等^[11]提出对时间序列在频域上进行处理的方法,达到了降低数据维数的目的,并且显著提高了相似匹配的效率;Lim等^[12]研究了相似性匹配中的多指标问题,并且详细讨论了滑动时间窗口维数与滑动步长对匹配精度的影响关系。虽然时间序列分析法广泛应用于状态预测,但大多是时间序列单维匹配,且没有考虑噪音数据,影响预测准确度。

针对以上算法中存在的问题,作者提出了一种基于多维时间序列的数控机床状态预测方法。利用多维时间序列表征机床未来状态,提出基于窗口滑动的多重匹配预测方法,为消除噪音数据的影响,提出了密度空间聚类算法的状态序列分析法。通过与现有算法进行实验对比,验证了本文方法的优势,为实现数控机床状态的准确预测提供了有价值的参考。

1 多维时间序列数据的分析处理方法

1.1 基于OPC技术的机床状态数据采集

OPC协议提供了3个主要的COM对象: OPC Server、OPC Group、OPC Items。OPC Server动态地创建或释放OPC Group, OPC Group提供包容OPC Items的机制,实现对OPC项的管理, OPC Items代表与OPC服务器到数据源(主轴电流、主轴负载、回转误差等)的连接,包括值、品质、时间戳3个基本属性。

SIEMENS 840D由人机界面,数控实时操作系统(NCK),数控单元和可编程逻辑控制器(PLC)组成。OPC数据采集客户端在人机界面上运行,使用组件对象模型(COM)访问OPC服务器,OPC服务器从NCU中的动态数据交换服务NCDDE请求相应的数据。结合多点接口(MPI)的特性,NCDDE服务器与PLC之间可进行数据交换,实现OPC客户端对数控机床数据的采集。

1.2 时间序列数据预处理

1.2.1 归一化处理

为消除时间序列相同量纲下大数值的影响,需要对各个参数进行归一化处理,针对机床参数数据的随机性和不稳定性,本文采用Min-max标准化法。

1.2.2 自回归移动平均模型

自回归移动平均模型(auto-regressive and moving average model, ARMA)是研究时间序列的重要方法,模型认为对于任意时间序列 $X = \{x_1, x_2, x_3, \dots, x_t\}$ 在时刻 t 的取值 x_t 不仅与前 n 个时刻的参数值 $x_{t-1}, x_{t-2}, \dots, x_{t-n}$ 有关,而且与之前 m 个随机白噪声干扰项 $a_{t-1}, a_{t-2}, \dots, a_{t-m}$ 有关,则 x_t 可表达为 (n, m) 阶的ARMA模型^[13]:

$$x_t = \phi_1 x_{t-1} + \phi_2 x_{t-2} + \dots + \phi_n x_{t-n} + a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \dots - \theta_m a_{t-m} \quad (1)$$

式中,误差干扰项 $a_t \sim NID(0, \sigma_a^2)$, $\phi_i (0 \leq i \leq n)$ 称为自回归系数, $\theta_i (0 \leq i \leq m)$ 表示移动平均系数。当 $\theta_i = 0$ 时,称为自回归模型记为AR(n)(auto-regressive);当 $\phi_i = 0$ 时,称为移动平均模型记为MA(m)(moving average)。

考虑机床参数的实际特征,通过对序列 X 建立自回归方程(AR)。若序列 $X = \{x_1, x_2, \dots, x_n, \dots, x_{2n}\}$,则序列元素满足方程:

$$\begin{cases} x_{n+1} = \phi_1 x_n + \phi_2 x_{n-1} + \dots + \phi_{n-1} x_2 + \phi_n x_1 + a_1, \\ x_{n+2} = \phi_1 x_{n+1} + \phi_2 x_n + \dots + \phi_{n-1} x_3 + \phi_n x_2 + a_2, \\ \dots \\ x_{2n} = \phi_1 x_{2n-1} + \phi_2 x_{2n-2} + \dots + \phi_{n-1} x_{n+1} + \phi_n x_n + a_n \end{cases} \quad (2)$$

1.3 多维时间序列状态模型及度量模型

多维时间序列是指观测对象的一组指标在不同时期上获得的一系列观测值,并按时间顺序排列而构成的数字集合,如图1所示。若基于OPC状态数据采集获得的参数种类有 m 个,且以采样频率100 ms获得 n 个数值点,则机床过程数据形成 m 行 n 列的多维时

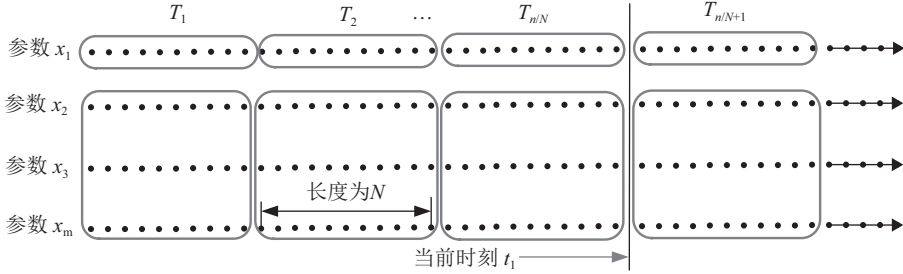


图1 多维时间序列示意图

Fig.1 Multi-dimensional time series schematic

在机床参数时间序列矩阵 $T_{m \times n}$ 中,令采样时间跨度为 N 的序列矩阵集合 $T_{m \times n}$ 为机床的状态模型,表征机床设备状态在时段 N 内的变化规律。图1中时间窗口 N 内的矩阵集合 $T = \{T_1, T_2, T_3, \dots, T_{n/N}\}$ 均为机床的历史状态,状态模型 $T_{(n/N)+1}$ 为需要预测的机床未来状态。针对机床的状态模型集合 $T_{m \times n}$,建立统一的状态度量标准和评价函数是对模型进行定量处理的基础,采用特征向量、特征趋势距离标示状态模型,利用差异度对状态模型进行对比衡量。

对于任意的状态模型 $T_k \in T_{m \times N}$,特征向量 F_k 为 (f_{kx}, f_{ky}) ,分别表示机床状态矩阵中某时刻点向量与平均值向量的最大、最小偏差值。

$$F_k(T_{m \times N}) = (f_{kx}, f_{ky}) = \left(\begin{array}{l} \text{Max}_j^N = \left| \sqrt{\sum_{i=1}^m (T_i(t_j))^2} - \frac{\sum_{j=1}^N \sqrt{\sum_{i=1}^m (T_i(t_j))^2}}{N} \right|, \\ \text{Min}_j^N = \left| \sqrt{\sum_{i=1}^m (T_i(t_j))^2} - \frac{\sum_{j=1}^N \sqrt{\sum_{i=1}^m (T_i(t_j))^2}}{N} \right| \end{array} \right) \quad (3)$$

式中, $T_i(t_j)$ 表示参数 m_i 在一个状态段 N 内时刻点 t_j 的参数值。

对于任意的状态模型 $T_k \in T_{m \times n}$,二范数 D_k 为模型的特征趋势距离,衡量机床状态模型的变化规律和平稳性特征,其表达式为:

$$D_k(T_{m \times N}) = \|T_i(t_j)\|_2 = \sqrt{\sum_{j=1}^N \sum_{i=1}^m (T_i(t_j))^2} \quad (4)$$

对于任意的状态模型 $T_k \in T_{m \times n}$,特征向量和特征

间序列矩阵 $T_{m \times n}$ 。

参数集合根据相应分类规则分为单维时间序列向量 x_1 和多维时间序列矩阵 $X = \{x_1, x_2, x_3, \dots, x_m\}$,并随着时间 t 的变化不断延伸。预测机床未来的状态需要利用机床历史的状态数据,应建立数控机床的状态模型和研究 t_1 时刻之后时间序列的预测方法。

趋势距离组成了二元组 $G_k = (F_k, D_k)$,则任意两个状态模型 T_i 和 T_j 的差异度 δ_{ij} 为:

$$\delta_{ij} = |G_i - G_j| = \sqrt{(F_i - F_j)^2 + (D_i - D_j)^2} = \sqrt{((f_{kx})_i - (f_{kx})_j)^2 + ((f_{ky})_i - (f_{ky})_j)^2 + (D_i - D_j)^2} \quad (5)$$

状态模型差异度可衡量机床两个状态之间的综合差别,差异度越大则说明状态矩阵吻合度越低,相似性最小,反之相似性最大。

2 基于多重匹配的状态预测模型研究

2.1 多重匹配预测建模理论

图1中,时间序列矩阵 $T_{n/N+1}$ 为需要预测的机床未来状态,将机床状态对应参数的序列值进行独立分析,滑动时间窗口建模过程如图2所示。

图2中,时间窗口长度为 l 的矩阵 $(X_{n+1})_{m \times l}$ 为需要预测的机床状态矩阵, $(X_n)_{m \times l}$ 为机床当前状态矩阵。从矩阵 X_n 的当前采样时刻元素 x_n 开始,以滑动单位 w ,时间窗口长度 l 进行逆向移动,可得第 k 个时间窗口代表的状态矩阵通项为:

$$X_k = [x(n-wk-l+1), x(n-wk-l+2), x(n-wk-l+3), \dots, x(n-wk)] \quad (6)$$

式中, $k=1, 2, \dots$,代表第 k 个历史状态, $w=1, 2, \dots$,为滑动时间窗口的滑动间隔长度。一般地,时间窗口长度 l 大于滑动时长 w ,且矩阵中的任意元素 $x_i (0 \leq i \leq l)$ 均为 $m \times 1$ 的列向量。结合滑动单位为 w 的时间窗口,离散时间序列被分割成了元素长度均为 l 的独立历史状态集合 $\tilde{X} = \{X_1, X_2, \dots, X_k, \dots\}$ 。

基于窗口滑动的多重匹配预测是在历史状态集合 \tilde{X} 中寻找与当前状态矩阵 X_n 相关性最大,即寻找相

似度最高的某一历史状态 \mathbf{X}_k , 且状态 \mathbf{X}_k 的跟随状态 \mathbf{X}_k^* 具有与 \mathbf{X}_{n+1} 相同的变化趋势和规律, 视为 \mathbf{X}_{n+1} 的预测值, 如图 3 所示。矩阵 \mathbf{X}_k 和 \mathbf{X}_k^* , 以及矩阵 \mathbf{X}_n 和 \mathbf{X}_{n+1} 分

别构成图 3 中时间跨度为 N 的一个完整机床状态。多重匹配即在寻找历史状态 \mathbf{X}_n 的过程中, 利用相似性分析和聚类分析分别完成初始寻找和精确匹配。

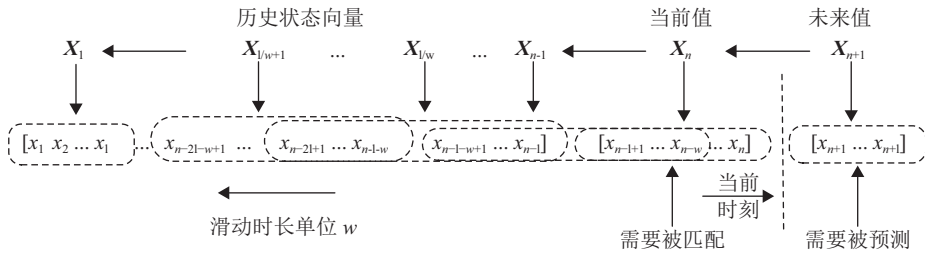


图 2 滑动时间窗口建模

Fig.2 Model of the sliding time window

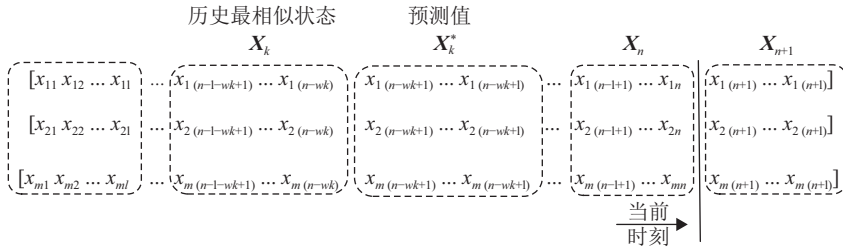


图 3 多重匹配预测

Fig.3 Multiple matching prediction

2.2 状态序列相似性分析

机床状态的相似性分析是指在机床海量历史状态数据中筛选与给定序列具有相似行为规律(变化趋势、数值大小)的状态矩阵集合, 主要包括序列分割和相似性度量。滑动时间窗口序列分割方法有效避免了参数在时间轴的扭曲, 保证了历史状态向量具有相同的长度。结合机床参数的实际特征, 提出基于矩阵相关性的 β -耦合度相似度量标准。

对于任意的机床状态矩阵 $\mathbf{X}_{m \times n} = [x_1, x_2, \dots, x_i, \dots, x_m]^T$ 和 $\mathbf{Y}_{m \times n} = [y_1, y_2, \dots, y_i, \dots, y_m]^T$, 其 β -耦合相似度定义为:

$$\cos \theta(X, Y) = \sum_{i=1}^m \left| \frac{\sum_{j=1}^n (x_{ij} - \lambda(x_i)) \times (y_{ij} - \lambda(y_i))}{\sum_{j=1}^n (x_{ij} - \lambda(x_i))^2 \times (y_{ij} - \lambda(y_i))^2} \right| / m \quad (7)$$

式中, $\lambda(x_i) = \frac{1}{n} \sum_{j=1}^n x_{ij}$, $\lambda(y_i) = \frac{1}{n} \sum_{j=1}^n y_{ij}$, $\lambda(x_i)$ 表示状态矩阵 $\mathbf{X}_{m \times n}$ 中第 i 行元素的平均值, m 为状态矩阵的行数即元素种类个数, n 为状态矩阵的列数即每个参数的长度。

β -耦合相似度可衡量机床两个状态矩阵中参数变化趋势的平均相似性和趋近性, 应满足以下性质:

- 1) 对称性且非负: $Co_{\beta}(X_1, Y_1) = Co_{\beta}(Y_1, X_1) \geq 0$;
- 2) 三角不等式: $0 \leq Co_{\beta}(X_1, Y_1) \leq Co_{\beta}(X_1, Z_1) + Co_{\beta}(Z_1, Y_1)$ 。

根据式(7)的相似性度量模型, 对于给定相似性阈值 $\varphi_d \in (0, 1)$, 当 2 个序列的相似度满足 $Co_{\beta}(X_1, Y_1) \geq \varphi_d$ 时, 称序列互为 β -耦合相似序列。相似度 Co_{β} 越大, 说

明两个状态模型中各行元素的平均相关性和相似度越大。相似度量模型不仅用于判断两个状态矩阵的相似性, 还用于确定最优滑动单位 \hat{w} 和预测时长 \hat{l} , 最优参数应满足:

$$\begin{aligned} & \text{Number}\{Co_{\beta}(T_{\text{history}}^*(\hat{w}, \hat{l}), \mathbf{X}_n) \geq \varphi_d\} \geq \\ & \text{Number}\{Co_{\beta}(T_{\text{history}}^*(w, l), \mathbf{X}_n) \geq \varphi_d\} \quad (8) \end{aligned}$$

式中, $T_{\text{history}}^*(\hat{w}, \hat{l})$ 代表以参数 \hat{w} 和 \hat{l} 构造的历史状态集合, $T_{\text{history}}^*(w, l)$ 为任意参数下的历史状态集合, Number 为历史集合中满足相似性要求的状态矩阵个数。

2.3 状态序列聚类分析

2.3.1 聚类度量函数

度量函数是聚类分析的核心, 是分析评价聚类元素共同点和差异度的关键因素。对于任意的两个状态序列 $\mathbf{X}_{m \times n} = [x_1, \dots, x_b, \dots, x_m]^T$ 和 $\mathbf{Y}_{m \times n} = [y_1, \dots, y_b, \dots, y_m]^T$ 其一般形式为:

$$d(\mathbf{X}, \mathbf{Y}) = \sqrt[\kappa]{|\mathbf{Y} - w \times \mathbf{X}|^{\kappa}} = \sqrt[\kappa]{\sum_{i=1}^m \sum_{j=1}^n |y_{ij} - wx_{ij}|^{\kappa}} \quad (\kappa \geq 1) \quad (9)$$

式中: w 为关联权重; 当 κ 取值为 1、2、 ∞ 时, 度量函数 $d(\mathbf{X}, \mathbf{Y})$ 分别为曼哈顿距离、欧几里得距离、切比雪夫距离。

2.3.2 密度空间聚类算法

密度空间聚类算法 (density-based spatial cluster-

ing of applications with noise, DBSCAN)是一种基于高密度联通区域的聚类算法,将类簇定义为高密度相连点的最大结合,可有效排除噪声数据的干扰,且能发现任意形状类簇,如图4所示。图4中 Q_1 、 Q_2 、 Q_3 为核心对象, p_1 、 p_2 、 p_3 、 p_4 、 p_5 、 p_6 为直接密度可达, q_1 、 q_2 、 q_3 、 q_4 、 q_5 、 q_6 为密度可达,则称对象 q_1 到对象 p_2 、对象 q_2 到对象 p_3 、对象 q_3 到对象 p_5 密度相连。

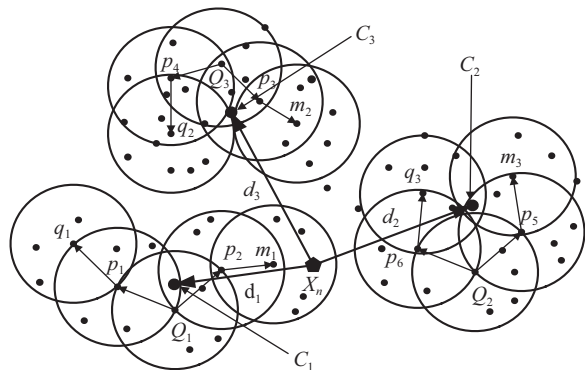


图4 密度空间聚类算法

Fig.4 Density-based spatial clustering of applications with noise

密度空间聚类算法实际上是将数据样本聚类为同簇样本之间均为密度相连点,而不同簇之间不相连,不在所有簇中的样本点作为噪声数据而丢弃。为保证各维度指标在相同的刻度级别,本文选用欧几里得距离作为聚类度量指标。

密度空间聚类算法的难点主要在于核心半径 r 、最小数目 $MinPts$ 的选取,为提高算法鲁棒性,通过下式确定参数取值:

$$\begin{aligned} r &= w \times (d_{\max} - d_{\min}), \\ MinPts &\in (0, Num/3) \end{aligned} \quad (10)$$

式中: d_{\max} 和 d_{\min} 分别为聚类元素状态中欧几里得距离的最大值和最小值; w 为缩小权重参数,取值范围为 $(0, 1)$; Num 为聚类元素的总个数。

若机床历史状态矩阵聚类中心簇个数为 k ,则聚类簇中心点 C_1, C_2, \dots, C_k 应满足:

$$D(C_k) = \text{Min} \sum_{p=1}^{M_k} d(C_k, U_p) \quad (11)$$

式中, M_k 为第 k 类中所有历史矩阵状态的样本数, U_p 为历史状态矩阵。

针对聚类元素为状态矩阵 s 时,密度空间聚类算法首先计算集合 $S = \{s_1, s_2, s_3, \dots, s_k, \dots, s_j\}$ 中任意两个状态矩阵的欧几里得距离 $d_{i,j}(s_i, s_j, i \neq j)$,根据式(10)确定核心半径 r ;然后,根据初始缩小权重 w ,最小半径数目 $MinPts$,任意选择一个未处理状态矩阵 s_k ,寻找所有与其密度相连的状态矩阵 $S^* = \{s_1^*, s_2^*, s_3^*, \dots,$

$s_k^*, \dots, s_j^*\}$,再从 S^* 中任选一个状态矩阵 s_k^* 继续寻找与其密度相连状态矩阵,直到集合 S 中的所有状态矩阵被遍历,利用式(11)得到聚类簇 $C = \{C_1, C_2, \dots, C_k\}$;最后,利用式(5)分别计算当前状态 X_n 与聚类中心 C_1, C_2, \dots, C_k 的差异度 δ_{ij} 值可完成 X_n 的精度匹配。

3 数控机床状态的预测实验

本次实验的环境为宁江机床厂的SIEMENS 840D系列机床和“S”试件,CPU为Inter Core i5-6200U 2.30 GHz, 4 G内存,测试软件为MATLAB 2014a。为验证本文算法的有效性,分别进行了基于多重匹配的向量形式、矩阵形式和AR模型的预测对比实验。

3.1 时间序列数据预处理

利用OPC数据采集客户端采集宁江机床SIEMENS 840D的运行数据,且运行数据在时间上具有一定的重复性。采集的主要参数有:主轴电机温度、主轴电机负载、主轴电流、主轴回转误差4项参数,采样周期为100 ms,机床切削S试件的累计有效加工时长(去除主轴转速为0时刻的数据)为12.8 h,产生的数据量总数为: $4 \times 10 \times 60 \times 60 \times 12.8 = 1\ 843\ 200$ 。

图5是4项参数在120 800~140 800 ms(20 s)内经过归一化处理后的图像,由图5可知,机床的回转误差、电机负载和主轴电流具有无序性和无规律性,而主轴电机温度取值比较均匀,主要在几个参数值中变化。

3.2 状态序列相似性分析

采用矩阵形式、向量形式对主轴4项参数进行相似性分析,取机床状态 $[T, T+1]$ 为机床当前状态 X_n , T 代表采样时间,并确定最优滑动单位 \hat{w} 和预测时长 \hat{l} 。令 w 以1个采样时间为公差在1~100个采样时间内依次递增,同时 l 以10个采样时间为公差在10~50个采样时间内依次递增。在给定相似度阈值区间和滑动窗口范围内,按式(7)和(8)求出最大相似集,其结果分别为图6和7(以 $T=362\ 500$,主轴回转误差为例)。

图6中,对机床主轴的4项参数以矩阵形式回滑1 000个时间窗口,相似度阈值取0.2、0.3和0.5。由图6可以看出,当预测时长为1 s时,图像峰值最高说明相似集中元素最多,结合实际情况,预测1 s的数据意义不大。因此,最终的最大相似集选用预测时长为24 s,滑动单位为2。同理,由图7可知,若主轴回转误差以单维向量的形式进行相似性分析,则应采用预测时长为24 s,滑动时长为2.3 s。

针对机床主轴的海量数据,以单维向量的形式将耗费大量计算时间,因此,对4种参数的状态预测采用矩阵形式,采用图6确定的最优预测时长和滑动

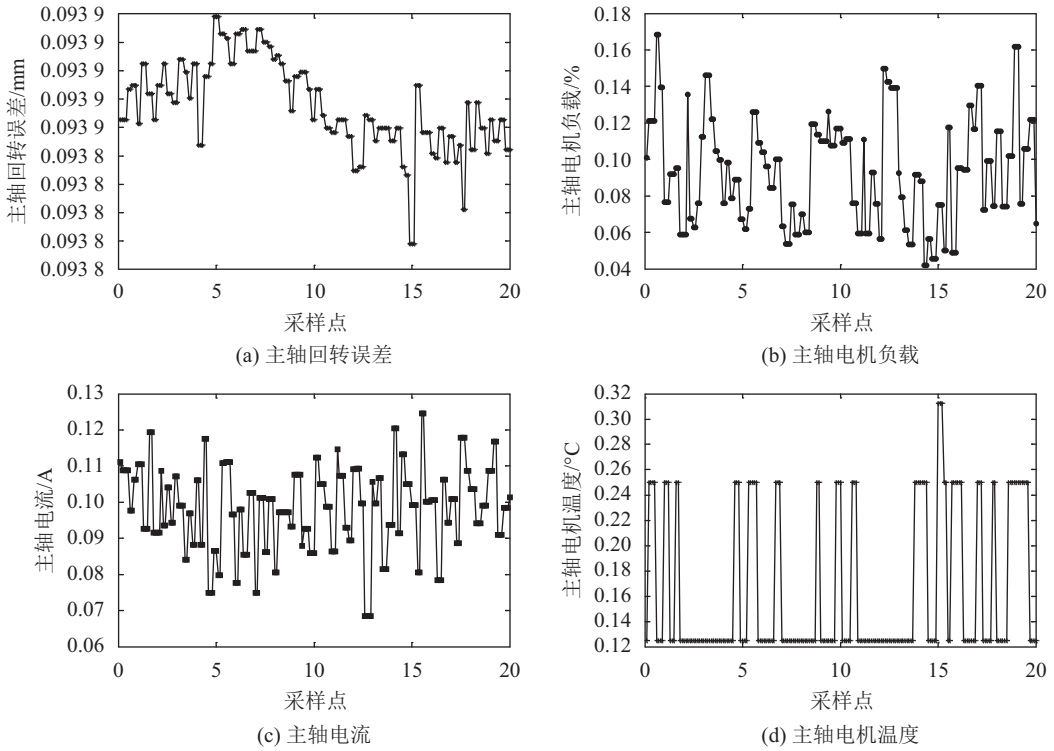


图 5 主轴4种参数归一化处理

Fig.5 Normalization of four parameters of spindle

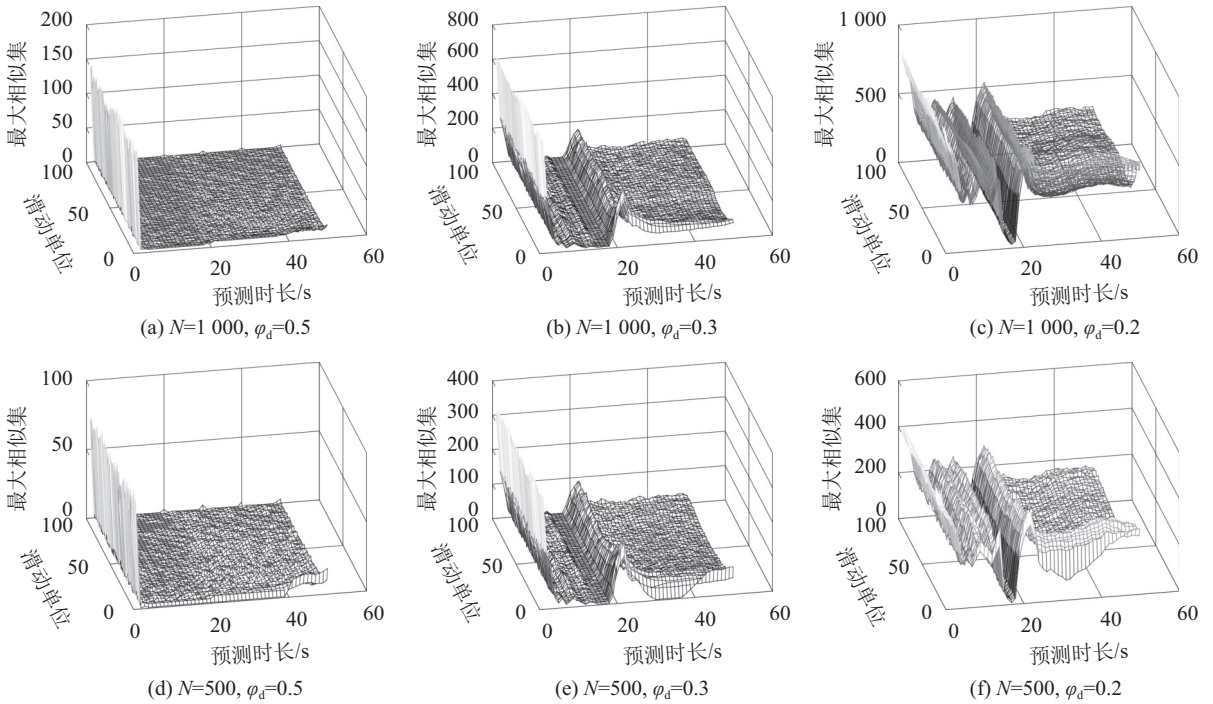


图 6 主轴回转误差以矩阵形式进行相似性匹配

Fig.6 Similarity matching with matrix form for four parameters

单位,并回滑 1 000 个时间窗口,最大相似集合的求解结果如表 1 所示。

3.3 状态序列聚类分析

根据状态序列相似性分析得到在 1 000 个历史滑

动窗口中满足 β -相似耦合度 $\varphi_d \geq 0.2$ 的 481 个历史状态矩阵 $T_{4 \times 240}$ 进行密度空间聚类分析,利用式 (10) 确定参数为 $w=0.3, MinPts=6$ 。由于聚类元素为 4 行 240 列的矩阵元素难以在平面图中展示,因此利用状

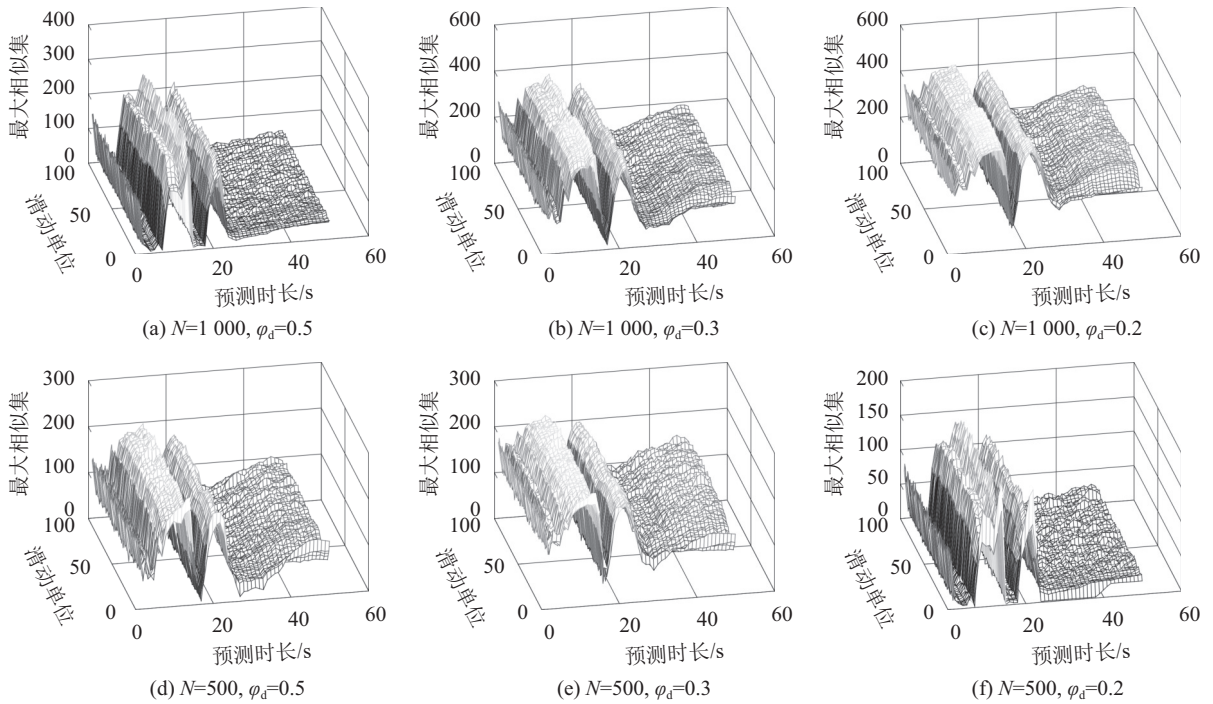


图7 主轴回转误差以向量形式进行相似性匹配

Fig.7 Similarity matching with vector form for the spindle rotating error

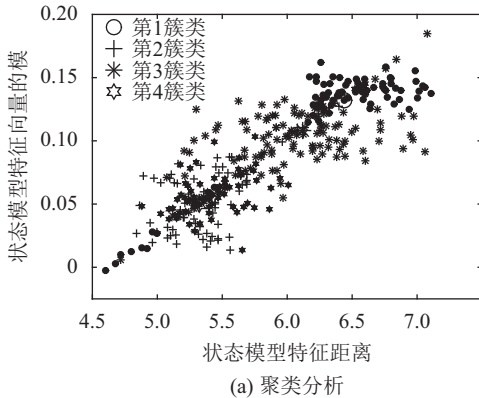
表1 回溯滑动窗口个数 $N=1\ 000$

Tab.1 Number of trace sliding window $N=1\ 000$

相似阈值 φ_d	0.2	0.3	0.5
相似集合个数	481	250	17

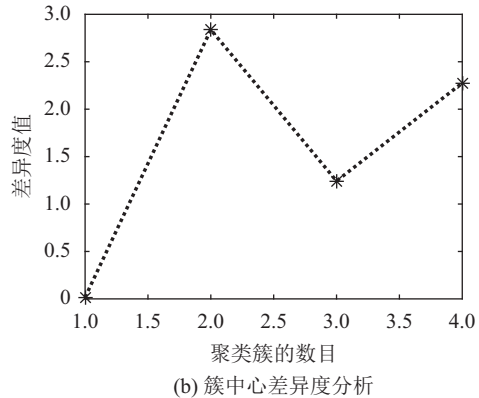
态特征值 $D(T_k)$ 和特征向量模 $\|F(T_k)\|$ 分别作为纵横坐标画出其聚类结果如图8(a)所示。对聚类后的簇中心和当前状态 X_n 进行差异度分析,其结果如图8(b)所示。

图8(a)中,聚类矩阵元素的不同类簇之间沿着



(a) 聚类分析

45°角排布,同一类簇元素大多沿着特征距离的 x 轴方向排布,这是由于聚类的指标为欧几里得距离近似于矩阵的特征距离。同时由于部分聚类元素可能具有相同的特征向量和特征距离,因此聚点结果存在叠合的现象。由图8(b)可知,4个聚类簇的中心元素,其中聚类簇1的中心状态矩阵和当前机床状态矩阵的差异度最小,因此,簇1中心点代表的状态 C_1 为多重匹配的结果。通过计算发现聚类簇 C_1 是滑动过程中的第769个窗口,利用 C_1 之后的240个采样点作为预测值。



(b) 簇中心差异度分析

图8 多重匹配结果

Fig.8 Multiple matching result

3.4 机床状态预测结果

在完成时间序列数据预处理、状态序列相似性分析、状态序列聚类分析之后,对数控机床主轴4种

参数使用基于多重匹配的矩阵形式、向量形式和传统的AR模型机床状态预测,其参数预测归一化结果如图9所示。为了便于区分,图中曲线分别做了不同

程度的平移。图9(a)中矩阵预测与向量预测分别向上平移了0.1与0.25个单位,图9(b)中向量形式预测向上平移了0.05个单位,图9(c)中向量形式预测向下平移了0.001个单位,图9(d)中矩阵预测与向量预测分别向上平移了0.15与0.3个单位。

对预测结果的评价,除差异度分析外,还可通过曲线的拟合量进行分析,采用最大误差 E_{\max} 、平均误差 E_a 、均方根误差 MSE 、相对误差率 ee 进行评价。矩阵形式、向量形式以及AR模型的预测结果及分析如表2所示。

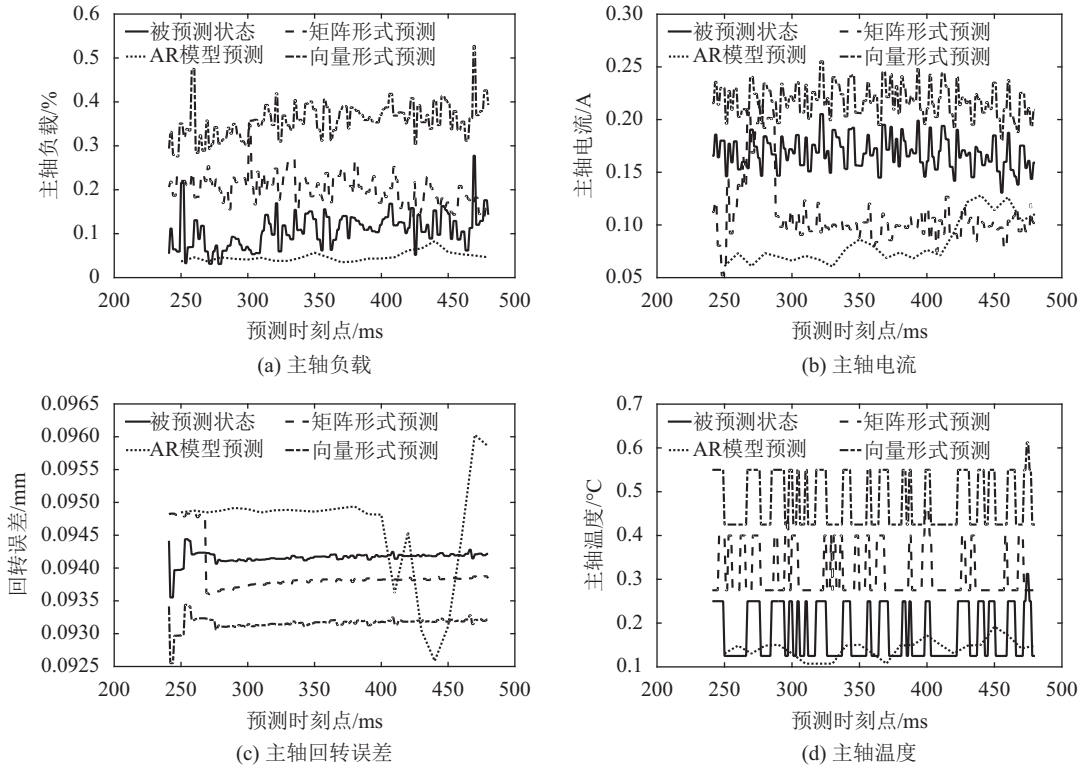


图 9 3种方法预测结果对比

Fig.9 Comparison of prediction results for three methods

表 2 3种预测方式结果

Tab.2 Prediction results for three methods

参数名	预测方法	最大误差	平均误差	均方误差	差异度	相对误差/%
回转误差	矩阵形式	1.310×10^{-3}	4.010×10^{-4}	4.310×10^{-4}	0.09	0.42
	向量形式	8.610×10^{-4}	3.910×10^{-5}	9.710×10^{-5}	0.086 9	0.04
	AR模型	9.381×10^{-2}	8.810×10^{-3}	2.710×10^{-2}	0.326 6	90.88
主轴负载	矩阵形式	0.232 7	0.031 9	0.058 8	0.09	28.55
	向量形式	0.155 5	0.010 9	0.028 1	0.012 9	10.12
	AR模型	0.257 3	0.072 8	0.083 5	0.407 6	67.60
主轴电流	矩阵形式	0.106 5	0.039 7	0.049 3	0.09	23.59
	向量形式	1.381×10^{-16}	5.710×10^{-19}	8.961×10^{-18}	0.003 9	0.01
	AR模型	0.129 6	0.082 4	0.086 5	0.452 6	48.96
电机温度	矩阵形式	0.187 5	0.057 3	0.086 0	0.09	32.26
	向量形式	2.781×10^{-15}	9.810×10^{-19}	1.281×10^{-17}	0.006 1	0.08
	AR模型	0.197 5	0.081 9	0.103 9	0.384 8	46.11

从表2的预测结果可以看出,基于多重匹配理论的矩阵形式、向量形式状态预测方法的各项误差均

比传统的AR模型低,特别是基于多重匹配的向量形式状态预测方法可以保证较高的预测置信度,其相

对误差最大值仅为10%,某些参数的相对误差低于1%,且差异度较小。相比于上述两种预测方法,基于传统的自回归移动平均模型精度较低,最大误差、平均误差、均方误差以及相对误差均较大,主轴回转误差相对误差最大值超过90%,失去了参数的真实变化规律,用于机床状态预测效果较差。因此,针对数控机床的多维时间序列,本文提出基于多重匹配的状态预测方法具有较高的预测置信度,可用于数控机床状态预测。

4 结 论

1)研究了多维时间序列数据的分析处理方法。利用OPC技术进行机床状态数据采集,对状态数据进行预处理,建立多维时间序列状态模型及其度量模型。

2)建立了基于多重匹配的机床状态预测模型。利用多重匹配预测建模理论,结合 β -耦合相似度和空间密度算法开展了状态序列相似性和聚类分析。

3)通过实验验证了方法的可行性。实验结果表明,当预测时长为24 s,滑动单位为2 s时的相似度最高,预测实验结果表明,基于多重匹配的机床状态预测方法比传统预测方法置信度更高,预测误差更低。

4)研究中忽略了多维时间序列的相似性分析时计算量的问题,下一步将针对这一情况对算法作出改进。

参考文献:

- [1] Ding Mingjun, Song Dan. Fault diagnosis expert system in CNC machine based on neural network[J]. Mechanical & Electrical Engineering Magazine, 2007, 24(5): 92-94. [丁明军, 宋丹. 基于神经网络的数控机床故障诊断专家系统[J]. 机电工程, 2007, 24(5): 92-94.]
- [2] Zhang Y, Yang J, Jiang H. Machine tool thermal error modeling and prediction by grey neural network[J]. The International Journal of Advanced Manufacturing Technology, 2012, 59(9): 1065-1072.
- [3] Sheng Z, Jin Y, Liu C, et al. Processing state monitoring of nc machine tools based on BP neural network[M]. Berlin: Springer, 2012: 508-515.
- [4] Du Liuqing, Zeng Cuilan, Yu Yongwei. Prediction of numerical control machine's motion precision based on multivariate chaotic time series[J]. Transactions of the Chinese Society for Agricultural Machinery, 2017, 48(3): 390-395. [杜柳青, 曾翠兰, 余永维. 基于多元混沌时间序列的数控机床运

动精度预测[J]. 农业机械学报, 2017, 48(3): 390-395.]

- [5] Shi Ning, Lin Yigong, Xiao Chengyong. Machine state forecast based on genetic-support vector machine[J]. Journal of Liaoning Technical University (Natural Science Edition), 2010, 29(4): 656-658. [石宁, 陆亦工, 肖成勇. 基于进化支持向量机的机械状态预测[J]. 辽宁工程技术大学学报(自然科学版), 2010, 29(4): 656-658.]
- [6] Xu Zhijun. State prediction for CNC machine based on PSO-SVM[J]. Modern Manufacturing Engineering, 2011(7): 46-49. [许志军. 基于粒子群算法优化支持向量机的数控机床状态预测[J]. 现代制造工程, 2011(7): 46-49.]
- [7] Wang Lixian. Time series prediction[D]. Tianjin University of Technology, 2012. [王丽贤. 时间序列预测技术研究[D]. 天津理工大学, 2012.]
- [8] Zhang Haibo, Huang Yangyang. The Forecasting about failure frequency of servo system in NC machine based on the random time series[J]. Journal of Northeast Dianli University (Natural Science Edition), 2014(1): 62-66. [张海波, 黄洋洋. 基于随机时间序列的数控机床伺服系统故障频率预测[J]. 东北电力大学学报(自然科学版), 2014(1): 62-66.]
- [9] Li Yongxiang, Tong Hengchao, Cao Hongtao. Application of time series analysis to thermal error modeling on NC machine tools[J]. Journal of Sichuan University (Engineering Science Edition), 2006, 38(2): 74-78. [李永祥, 童恒超, 曹洪涛, 等. 数控机床热误差的时序分析法建模及其应用[J]. 四川大学学报(工程科学版), 2006, 38(2): 74-78.]
- [10] Du Liuqing, Yin Guofu, Yu Yongwei. Evolution analysis of CNC machine tool motion precision based on chaotic phase space reconstruction[J]. Chinese Journal of Scientific Instrument, 2015, 36(8): 1810-1815. [杜柳青, 殷国富, 余永维. 基于混沌相空间重构的数控机床运动精度演化分析[J]. 仪器仪表学报, 2015, 36(8): 1810-1815.]
- [11] Agrawal R, Srikant R. Fast algorithms for mining association rules[M]. San Francisco: Morgan Kaufmann Publishers Inc, 1998: 2299-308.
- [12] Lim S H, Park H J. Using multiple indexes for efficient subsequence matching in time-series databases[J]. Information Sciences, 2007, 17(7): 5691-5706.
- [13] Sun Xiaolei, Ding Yawei, Guo Keyu, et al. Ship seawater cooling system parameter prediction based on arma model[J]. Computer Measurement and Control, 2017, 25(7): 285-289. [孙晓磊, 丁亚委, 郭克余, 等. 基于ARMA模型的船舶海水冷却系统参数预测[J]. 计算机测量与控制, 2017, 25(7): 285-289.]

(编辑 黄小川)

引用格式: Li Hai, Wang Wei, Huang pu, et al. State prediction method research in NC machine tool based on multidimensional time series[J]. Advanced Engineering Sciences, 2018, 50(1): 187-195. [李海, 王伟, 黄璞, 等. 基于多维时间序列的数控机床状态预测方法研究[J]. 工程科学与技术, 2018, 50(1): 187-195.]