

Al-Cu-Mg-Ag 合金强度性能的支持向量回归预测

唐江凌, 蔡从中, 皇思洁, 肖婷婷

(重庆大学 应用物理系, 重庆 401331)

摘要: 为了研究不同时效工艺下 Al-Cu-Mg-Ag 合金强度性能, 根据实测数据集, 应用基于粒子群算法 (PSO) 寻优的支持向量回归 (SVR) 方法, 建立了 SVR 预测模型。模型以 Al-Cu-Mg-Ag 合金时效温度与时效时间为输入, 合金的抗拉强度、屈服强度为输出。经过与 BP 神经网络模型进行比较, 结果表明: 对于相同的训练样本和检验样本, 支持向量回归模型比 BP 神经网络模型具有更高的预测精度。

关键词: Al-Cu-Mg-Ag 合金; 强度; 支持向量回归; 粒子群优化; 回归分析

DOI: 10.3969/j.issn.1005-5053.2012.5.016

中图分类号: TG146.2

文献标识码: A

文章编号: 1005-5053(2012)05-0092-05

Al-Cu-Mg-Ag 系铝合金是在 2000 系合金的基础上通过提高合金化元素纯度和改变合金元素含量以及利用新的时效工艺发展起来的。因为具有较高的强度和热稳定性, Al-Cu-Mg-Ag 系铝合金是航空航天工业上很有发展潜力的中温高强材料, 具有广阔的应用前景。目前, 国内外新型 Al-Cu-Mg-Ag 合金已经从机理研究进入实际应用开发阶段^[1-5]。

力学性能是铝合金的一个极其重要的性能指标, 而抗拉强度、屈服强度则是铝合金材料力学性能的主要参数。因此, 在生产铝合金的工艺过程中, 通过建立加工工艺参数与合金力学强度关系的预测模型, 对于准确估计该铝合金的力学性能具有十分重要的指导意义。然而, 由于加工过程中各工艺参数与合金力学强度之间呈现出高度的非线性特性, 传统的预测模型很难精确地描述它们之间的复杂关系。因此, 一些新的建模方式, 如偏最小二乘 (PLS)、反向传播人工神经网络 (BPNN) 等方法被用来建立铝合金的力学性能预测模型^[6,7]。尽管这些方法能较好地预测铝合金的力学性能, 但它们的预测精度还不够高, 泛化能力也不够理想。所以, 进一步探索和寻求更加准确、有效的理论预测方法很

有必要。

支持向量机 (Support Vector Machines, SVM) 是 20 世纪 90 年代由 Vapnik 等^[8]提出的一种新的机器学习方法。它是在统计学习理论 (Statistical Learning Theory, SLT) 的 VC 维理论和结构风险最小化 (Structure Risk Minimization, SRM) 准则基础上, 由给定的有限样本信息, 对经验风险和泛化性能进行折中, 以得到最佳的训练和泛化性能。近几年 SVM 在其理论研究和算法实现方面都取得了突破性进展, 并逐渐成为克服“维数灾难”和“过学习”等传统困难的有力工具, 在分类和回归分析领域得到了广泛应用^[9-15]。当 SVM 用于回归预测时, 又被称为支持向量回归 (Support Vector Regression, SVR)。

本研究利用文献[7]报道的 Al-Cu-Mg-Ag 铝合金在不同的时效工艺下的强度实验数据集, 应用基于粒子群优化 (Particle Swarm Optimization, PSO) 算法进行参数寻优的 SVR 方法, 对该铝合金的强度性能进行建模和预测研究, 并与文献[7]报道的应用神经网络预测结果进行比较。

1 原理和方法

1.1 支持向量回归原理

设样本集为 $(x_1, y_1), \dots, (x_m, y_m)$ ($x_m \in \mathbf{R}^n$ 是 n 维输入变量, $y \in \mathbf{R}$ 是相应的输出变量, m 为样本个数), 寻找一个输入空间到输出空间的非线性映射 $\Phi(x)$, 通过该映射, 将样本集中的数据 x 映射到高维特征空间 F , 并在特征空间 F 中用下述线性函数进行拟合, 即:

收稿日期: 2011-12-09; 修订日期: 2012-07-13

基金项目: 教育部新世纪优秀人才支持计划资助项目 (NCET-07-0903); 教育部留学回国人员科研启动基金资助项目 (教外司留[2008]101-1); 中央高校基本科研业务费资助项目 (CDJXS11101135)

作者简介: 唐江凌 (1974—), 男, 博士研究生, 讲师, (E-mail) tangjiangling1974@hotmail.com

通讯作者: 蔡从中, 男, 教授, (E-mail) caiczhang@gmail.com。

$$f(x) = w \cdot \Phi(x) + b, \Phi: R^n \rightarrow F, w \in F. \quad (1)$$

式中 w 是回归系数向量, b 是阈值。由于 Φ 固定不变, 因此影响 w 的是经验风险的总和以及使其在高维空间平坦的 $\|w\|^2$, 即:

$$\text{minimize } R(C) = \frac{1}{2} \|w\|^2 + C \sum_{i=1}^m L_\varepsilon(f(x_i) - y_i), \quad (2)$$

$$L_\varepsilon(f(x_i) - y_i) = \begin{cases} 0, & |f(x_i) - y_i| < \varepsilon, \\ |f(x_i) - y_i| - \varepsilon, & |f(x_i) - y_i| \geq \varepsilon. \end{cases} \quad (3)$$

上两式中, $R(C)$ 为风险泛函, $L_\varepsilon(f(x_i) - y_i)$ 是损失函数, C 是惩罚因子, ε 是误差。为了控制函数的复杂性, 应使线性回归函数尽量平坦, 并考虑可能超出精度的回归误差, 引入松弛因子 ξ^* 和 ξ 以处理未满足(3)式的数据点。根据统计学习理论的结构风险化准则, SVR 通过极小化目标函数 $R(w, \xi_i, \xi_i^*)$ 来确定(1)式中的 w 和 b :

$$\text{minimize } R(w, \xi_i, \xi_i^*) = \frac{1}{2} \|w\|^2 + C \sum_{i=1}^m (\xi_i + \xi_i^*),$$

$$\text{subject to } \begin{cases} y_i - w \cdot x_i - b \leq \varepsilon + \xi_i, \\ w \cdot x_i + b - y_i \leq \varepsilon + \xi_i^*, \\ \xi_i, \xi_i^* \geq 0. \end{cases} \quad (4)$$

上式 $R(w, \xi_i, \xi_i^*)$ 中第一项是使回归函数更为平坦, 泛化能力更好, 第二项则为减少误差, 惩罚因子 C 是一个常数, 且 $C > 0$, 用来控制对超出误差 ε 的样本的惩罚程度。为求 w 和 b , 建立拉格朗日方程:

$$L(w, \xi_i, \xi_i^*) = \frac{1}{2} \|w\|^2 + C \sum_{i=1}^m (\xi_i + \xi_i^*) - \sum_{i=1}^m \alpha_i ((\varepsilon + \xi_i) + y_i + (w \cdot \Phi(x_i)) + b) - \sum_{i=1}^m \alpha_i^* ((\varepsilon + \xi_i^*) + y_i + (w \cdot \Phi(x_i)) - b) - \sum_{i=1}^m (\lambda_i \xi_i + \lambda_i^* \xi_i^*). \quad (5)$$

要使(5)式取得最小值, 对于参数 w , b , ξ_i , ξ_i^* 的偏导都应等于零。即:

$$\frac{\partial L}{\partial w} = w - \sum_{i=1}^m (\alpha_i - \alpha_i^*) \cdot \Phi(x_i) = 0,$$

$$\frac{\partial L}{\partial b} = \sum_{i=1}^m (\alpha_i - \alpha_i^*) = 0,$$

$$\frac{\partial L}{\partial \xi_i} = C - \alpha_i - \lambda_i = 0,$$

$$\frac{\partial L}{\partial \xi_i^*} = C - \alpha_i^* - \lambda_i^* = 0. \quad (6)$$

代入(5)式, 可以得到对偶优化问题:

$$Q = \sum_{i,j=1}^m (\alpha_i - \alpha_i^*) (\alpha_j - \alpha_j^*) (\Phi(x_i) \Phi(x_j)) + \sum_{i=1}^m \alpha_i (\varepsilon - y_i) + \sum_{i=1}^m \alpha_i^* (\varepsilon + y_i),$$

$$\text{subject to } \begin{cases} \sum_{i=1}^m (\alpha_i - \alpha_i^*) = 0, \\ \alpha_i, \alpha_i^* \in [0, C]. \end{cases} \quad (7)$$

由此, 支持向量机的函数回归问题就可以归结为对偶优化问题(7)。求解该对偶优化问题, 可以得到用训练样本点表示的 w :

$$w = \sum_{i=1}^m (\alpha_i - \alpha_i^*) \Phi(x_i). \quad (8)$$

式中 α_i 和 α_i^* 是最小化 $R(w)$ 的解。由此可求得线性回归函数:

$$f(x) = \sum_{i=1}^m (\alpha_i - \alpha_i^*) k(x, x_i) + b, \quad (9)$$

式中 $k(x, x_i) = \Phi(x) \cdot \Phi(x_i)$ 为核函数。选择不同形式的核函数就可以生成不同的 SVR 回归模型。常用的核函数有: 径向基函数、多项式函数、感知器函数、线性函数等。本工作采用径向基函数(10)来建立 SVR 回归模型:

$$k(x, x_i) = \exp(-\gamma \|x - x_i\|^2). \quad (10)$$

1.2 支持向量回归参数的粒子群算法寻优

由于 SVR 回归模型的泛化性能很大程度上依赖于不敏感损失函数 ε 、误差惩罚因子 C 和核函数参数 γ 等 3 个参数。因此, 对 (ε, C, γ) 参数进行寻优是十分关键的。本研究采用 PSO 算法来寻找最优组合参数 (ε, C, γ) 。

PSO 是由 Eberhart 和 Kennedy^[16] 模拟鸟群的飞行捕食行为而于 1995 年提出的一种高效多维并行寻优算法。设群体中的每个粒子由 3 维参数向量 (ε, C, γ) 组成, 第 i 个粒子在 3 维解空间的位置为 $u_i = (u_{i1}, u_{i2}, u_{i3})^T$, 其速度为 $v_i = (v_{i1}, v_{i2}, v_{i3})^T$, 当前时刻的个体极值记为 p_{ibest} , 全局极值记为 g_{best} 。在每次迭代中, 粒子跟踪个体极值、全局极值和自己前一时刻的状态来调整当前时刻的位置和速度, 迭代公式如下:

$$v_i(t+1) = \omega \cdot v_i(t) + c_1 \cdot \text{rand}() \cdot (p_{ibest} - u_i(t)) + c_2 \cdot \text{rand}() \cdot (g_{best} - u_i(t)) \quad (11)$$

$$u_i(t+1) = u_i(t) + v_i(t+1), \quad (12)$$

式中 $v(t), v(t+1), u(t)$ 和 $u(t+1)$ 分别是粒子在当前时刻、下一时刻的速度和位置; $\text{rand}()$ 是 $[0, 1]$ 之间的随机数; c_1 和 c_2 是学习因子, 通常取为 2; ω 是权重因子, 为加快收敛速度, 其值应随算法迭代的

进行而自动调节,一般定义为:

$$\omega = \omega_{\min} + (iter_{\max} - iter) \cdot (\omega_{\max} - \omega_{\min}) / iter_{\max} \quad (13)$$

式中 ω_{\max} 和 ω_{\min} 分别为最大和最小权重因子,它们的值一般分别取为 0.9 和 0.4。 $iter$ 为当前迭代次数, $iter_{\max}$ 为总的迭代次数。为了直接反映 SVR 模型的回归性能,选用均方根误差 ($RMSE$) 作为适应度函数:

$$RMSE = \sqrt{\frac{1}{m} \sum_{i=1}^m (\hat{y}_i - y_i)^2}, \quad (14)$$

式中 m 是训练样本数, y_i 和 \hat{y}_i 分别是第 i 个训练样本目标量的实测值和预测值。

2 SVR 回归模型的建立

2.1 时效参数下的 Al-Cu-Mg-Ag 合金强度性能实测数据集

文中所用 Al-Cu-Mg-Ag 合金成分如表 1。对合金的工艺处理方式见文献[7]。实验数据共 27 组,是在时效状态的每个时效时间点取两片 Al-Cu-Mg-Ag 合金试样,等时效完毕后再测量试样在室温下的抗拉强度和屈服强度,取其平均值(详见表 2)。

2.2 模型的建立

为了与文献[7]中神经网络方法的预测效果进行直接比较,本研究也采用实验数据集中的前 25 个样本进行建模,并将后两组(第 26 和 27 组)样本作为测试样本用于检验模型的预测能力。其中,时效温度、时效时间两个工艺参数作为模型的输入变量,抗拉强度和屈服强度两个力学性能量为输出变量。

2.3 模型预测性能的评价

采用平均绝对误差 (MAE)、平均绝对百分误差 ($MAPE$) 对所建模型的预测性能进行评价:

$$MAE = \frac{1}{n} \sum_{j=1}^n |\hat{y}_j - y_j|, \quad (15)$$

$$MAPE = \frac{1}{n} \sum_{j=1}^n \left| \frac{\hat{y}_j - y_j}{y_j} \right|, \quad (16)$$

在式(15)和式(16)中 n 是检验样本数, y_j 和 \hat{y}_j 分别是第 j 个检验样本的实测值和预测值。

表 1 实验合金成分

Table 1 Composition of Al-Cu-Mg-Ag alloy

Cu	Mn	Mg	Ag	Fe	Si	Al
5.8	0.2	0.3	0.7	0.02	0.02	Bal.

表 2 实验数据

Table 2 Experimental data

NO	Aging temperature / °C	Aging time / h	Tensile strength / MPa	Yield strength / MPa
1	165	1	410.725	506.35
2	165	2	492.18	526.565
3	165	4	512.38	530.825
4	165	6	548.375	555.46
5	165	8	548.63	554.605
6	165	10	556.745	561.205
7	165	14	555.755	559.05
8	165	20	546.84	552.785
9	165	50	541.475	548.715
10	165	110	534.09	540.32
11	200	1	511.07	522.864
12	200	2	505.995	520.923
13	200	4	504.875	524.733
14	200	8	471.29	501.119
15	200	10	470.64	501.015
16	200	12	458.405	491.911
17	200	14	457.035	491.884
18	200	16	450.59	485.068
19	200	18	456.08	488.773
20	200	20	446.015	481.042
21	250	0.167	474.765	500.578
22	250	0.5	440.38	474.836
23	250	1.167	417.115	462.568
24	250	2	380.58	436.693
25	250	9	309.825	389.533
26*	200	6	483.23	508.251
27*	250	0.833	429.76	467.009

*—test sample; others—training samples

3 结果分析与讨论

表 3 列出了 BP 神经网络模型和 SVR 模型的预测结果。由表 3 可见,无论是抗拉强度还是屈服强度,SVR 模型的预测结果都更接近于实测值。这表明 SVR 模型的预测精度高于 BP 神经网络模型,可以用于 Al-Cu-Mg-Ag 合金强度性能的预测。表 4 对比了 BP 神经网络模型和 SVR 模型二者预测值的平均绝对误差 MAE 和平均绝对百分误差 $MAPE$ 。表中数据显示,SVR 模型的 MAE 和 $MAPE$ 都分别小于 BP 神经网络模型。这进一步说明,SVR 模型具有更高的预测精度,更好的泛化能力,是一种更有效的预测、分析手段。

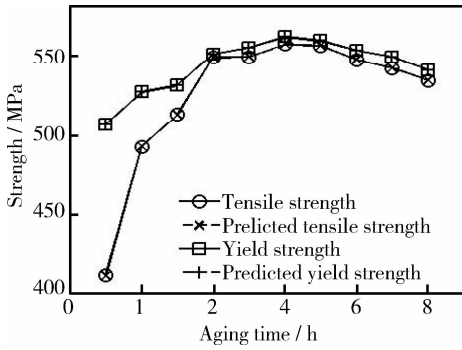


图 1 合金 165°C 时抗拉强度、屈服强度性能曲线
Fig.1 Strength properties of the alloy aged at 165°C

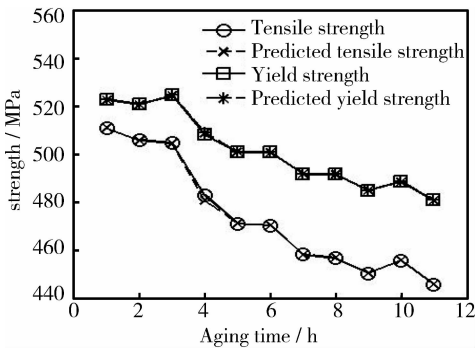


图 2 合金 200°C 时抗拉强度、屈服强度性能曲线
Fig.2 Strength properties of the alloy aged at 200°C

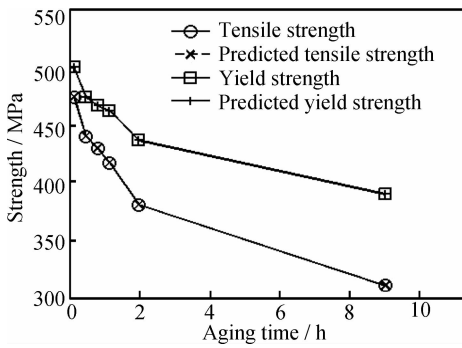


图 3 合金 250°C 时抗拉强度、屈服强度性能曲线
Fig.3 Strength properties of the alloy aged at 250°C

表 3 BP 神经网络模型和 SVR 模型的预测结果比较
Table 3 Comparison of predicted result between BP neural network and SVR

NO	Tensile strength / MPa			Yield strength / MPa		
	Exp.	BPNN Pre.	SVR Pre.	Exp.	BPNN Pre.	SVR Pre.
26*	483.23	478.44	480.99	508.25	505.92	509.33
27*	429.76	427.78	429.76	467.01	467.67	467.01

表 4 BP 神经网络模型和 SVR 模型的预测性能比较
Table 4 Comparison of prediction between BP neural network and SVR

Method	Tensile strength		Yield strength	
	MAE/MPa	MAPE/%	MAE/MPa	MAPE/%
BPNN	3.53	0.76	1.973	0.30
SVR	1.12	0.23	0.539	0.11

4 结论

(1) Al-Cu-Mg-Ag 合金力学性能与时效温度和时效时间关系密切。随着温度升高,合金抗拉强度和屈服强度达到峰值时效的时间变短;随着时间加长,合金抗拉强度和屈服强度下降。

(2) 支持向量回归方法对 Al-Cu-Mg-Ag 合金强度性能的预测非常有效,其准确程度超过 BP 神经网络,可以用于 Al-Cu-Mg-Ag 铝合金时效过程中工艺参数与强度性能的关系的预测和分析。

参考文献:

[1] HUTCHINSON C R, FAN X, PENNYCOOK S J, *et al.* On the origin of the high coarsening resistance of Omega plates in Al-Cu-Mg-Ag alloys [J]. *Acta Mater*, 2001, 49 (14): 2827 - 2841.

[2] SONG M, CHEN K H, HUANG L P. Effects of Ag addition on mechanical properties and microstructures of Al-8Cu-0.5Mg alloy [J]. *Trans Nonferrous Met Soc China*, 2006, 16 (4): 766 - 771.

[3] 肖代红, 黄伯云, 宋 旻, 等. Al-Cu-Mg-(Ag, La) 合金的显微组织与力学性能 [J]. *中国有色金属学报*, 2008, 18(4): 571 - 576.
(XIAO D H, HUANG B Y, SONG M, *et al.* Microstructures and mechanical properties of Al-Cu-Mg-(Ag, La) alloys [J]. *The Chinese Journal of Nonferrous Metals*, 2008, 18(4): 571 - 576.)

[4] 毛柏平, 雷文平, 沈健. Ag 对 Al-Cu-Mg 合金组织与性能的影响[J]. *热加工工艺*, 2009, 38(14): 48 - 50.
(MAO B P, LEI W P, SHEN J. Effects of Ag on microstructure and properties of Al-Cu-Mg Alloy [J]. *Hot Working Technology*, 2009, 38(14): 48 - 50.)

[5] 侯延辉, 刘志义, 李云涛, 等. Al-Cu-Mg-Ag 合金峰时效前微观组织演变及屈服强度数值模拟 [J]. *稀有金属材料与工程*, 2010, 39(8): 1435 - 1438.
(HOU Y H, LIU Z Y, LI Y T, *et al.* Simulation of microstructure evolution and yield strength of Al-Cu-Mg-Ag alloys before peak aging stage [J]. *Rare Metal Materials and*

- Engineering, 2010, 39(8): 1435 – 1438.)
- [6] 方善锋, 汪明朴, 王正安, 等. PLS-BPN 法用于 7005 铝合金力学性能与工艺参数的定量研究 [J]. 中国有色金属学报, 2007, 17(12): 1948 – 1954.
(FANG S F, WANG M P, WANG Z A, *et al.* Quantitative study between mechanical properties and processing parameters for 7005Al alloys predicted using PLS-BPN method [J]. The Chinese Journal of Nonferrous Metals, 2007, 17(12): 1948 – 1954.)
- [7] 侯延辉, 刘志义, 柏松, 等. Al-Cu-Mg-Ag 合金时效强化的神经网络预测模型 [J]. 材料热处理学报, 2009, 30(3): 203 – 208.
(HOU Y H, LIU Z Y, BAI S, *et al.* Prediction of aging strengthening of Al-Cu-Mg-Ag alloy based on neural network [J]. Transactions of Materials and Heat Treatment, 2009, 30(3): 203 – 208.)
- [8] VAPNIK V. The nature of statistical learning theory [M]. New York: Springer, 1995.
- [9] CAI C Z, HAN L Y, JI Z L, *et al.* SVM-Prot: web-based support vector machine software for functional classification of a protein from its primary sequence [J]. Nucleic Acids Res, 2003, 31(13): 3692 – 3697.
- [10] CAI C Z, HAN L Y, JI Z L, *et al.* Enzyme family classification by support vector machines [J]. Proteins, 2004, 55(1): 66 – 76.
- [11] LIANG G Z, LI Z L, ZHOU Y H, *et al.* A new peptide sequences representation technique and support vector machine for quantitative structure-retention modeling of peptides in HPLC [J]. Acta Phys Chim Sin, 2006, 22(9): 1052 – 1055.
- [12] DING Y M, ZHANG T T, LIU X, *et al.* Phase diagram of CsF-RbF system [J]. Acta Phys Chim Sin, 2007, 23(4): 614 – 616.
- [13] LI P, TAN N X, RAO H B, *et al.* Classification Models for HERG Potassium Channel Inhibitors Based on the Support Vector Machine Approach [J]. Acta Phys Chim Sin, 2009, 25(8): 1581 – 1586.
- [14] LU W, XUE Y. Activity Prediction of Hormone-Sensitive Lipase Inhibitors Based on Machine Learning Methods [J]. Acta Phys Chim Sin, 2010, 26(2): 471 – 477.
- [15] YANG G B, LI Z R, RAO H B, *et al.* Classification Models for Acetylcholinesterase Inhibitors Based on Machine Learning Methods [J]. Acta Phys Chim Sin, 2010, 26(12): 3351 – 3359.
- [16] KENBEDIY J, EBERHART R. Particle swarm optimization [C] // Proc IEEE Int Conf Neural Networks, 1995.

Strength Prediction for Al-Cu-Mg-Ag Alloy Based on Support Vector Regression

TANG Jiang-ling, CAI Cong-zhong, HUANG Si-jie, XIAO Ting-ting

(Department of Applied Physics, Chongqing University, Chongqing 401331, China)

Abstract: To explore the strength properties of Al-Cu-Mg-Ag alloy at different aging temperature and aging time, the support vector regression (SVR) approach combined with particle swarm optimization (PSO) algorithm was proposed to construct a SVR model based on experimental data. In the modeling process, the aging temperature and aging time were employed as input parameters, the tensile strength and yield strength acted as outputs. By comparison with BP neural network, it was found that the prediction accuracy of the established SVR model was higher than that of BPNN model by applying identical training and test samples. This investigation would provide a theoretical foundation for further study on the effect of aging condition on mechanical property, and the optimal design of the aging process for fabricating Al-Cu-Mg-Ag alloy.

Key words: Al-Cu-Mg-Ag alloy; strength property; support vector regression; particle swarm optimization; regression analysis