

文章编号: 1000-582X(2011)03-044-06

基于参数优化的人工神经网络的 AZ31 镁合金力学性能预测模型

刘彬^a, 汤爱涛^{a, b}, 潘复生^{a, b}, 黄光杰^{a, b}, 毛建军^a

(重庆大学 a. 材料科学与工程学院; b. 国家镁合金材料工程技术研究中心, 重庆 400044)

摘要:通过力学性能试验测定了不同退火条件下 AZ31 镁合金的抗拉强度、屈服强度和延伸率, 并利用人工神经网络技术建立了对应力学性能的预测模型, 其中对模型的优化采用了一种新方法, 即参数全排列组合训练。结果表明, 基于全排列训练得到的最优参数建立的网络模型具有优良的性能, 比经传统试探法构建的模型具有更高的平均相关系数和更低的平均误差, 因此能更准确地预测 AZ31 镁合金在不同退火条件后的力学性能。

关键词:镁合金; 力学性能; 人工神经网络; 预测模型; 全排列组合训练

中图分类号: TP183

文献标志码: A

Mechanical property prediction model of AZ31 magnesium alloys based on an artificial neural network with parameter optimization

LIU Bin^a, TANG Ai-tao^{a, b}, PAN Fu-sheng^{a, b}, HUANG Guang-jie^{a, b}, MAO Jian-jun^a

(a. College of Materials Science and Engineering; b. National Engineering Research Center for Magnesium Alloys, Chongqing University, Chongqing 400044, P. R. China)

Abstract: The tensile strength, yield strength and elongation of AZ31 magnesium alloys on different annealed conditions are tested by mechanical properties experiments. A model of corresponding mechanical properties is built by applying artificial neural network, and it is optimized by a new method, namely all permutations and combinations training of parameters. The results show that the network model has an excellent performance, which is based on optimal parameters obtained from all permutations and combinations training. Compared with traditional model, whose parameters are obtained from conventional heuristic, the improved model has higher average correlation coefficient and lower average error. Therefore, it can predict the mechanical properties of AZ31 magnesium alloy on different annealed conditions more accurately.

Key words: magnesium alloys; mechanical properties; artificial neural network; prediction model; all permutations and combinations training

镁合金是当前最有发展潜力的轻合金材料, 在交通运输、航空航天、通讯电子行业等领域具有广阔的应用前景^[1-4]。然而, 镁合金的研究和生产过程一

直存在不少困难。比如大量的镁合金数据庞杂凌乱, 阻碍了数据的标准化和规范化^[5]; 大部分镁合金由于密排六方结构, 室温滑移系少, 塑性成形性较

收稿日期: 2010-09-15

基金项目: 国家 973 计划资助项目(2007CB613704); 国家杰出青年基金资助项目(50725413)

作者简介: 刘彬(1982-), 男, 重庆大学博士研究生, 主要从事镁合金和计算材料学方面的研究。

汤爱涛(联系人), 女, 重庆大学教授, 博士生导师, (E-mail) tat@cqu.edu.cn。

差^[6];特别是高性能镁合金的研究还处于探索状态,因此如何提高镁合金力学性能正是当前的研究重点。但是,力学性能受诸多因素影响,如合金成分、工艺参数、组织结构等,这些因素相互影响,迄今为止还没有传统的数学模型能够对它们之间的定量关系进行准确描述。

人工神经网络不需要任何假设或预定公式,能直接从已有的实验数据出发,经过有限次迭代计算而得到一个反映数据内在规律的数学模型,而且还可利用经训练好的网络模型进行推理预测,特别适合于研究复杂的非线性系统。但模型的预测精度受建模参数的影响极大,目前对参数的优选并没有足够完善的理论指导,大多数研究均停留在采用试探法寻找最优参数的阶段^[7],因此如何选取最优建模参数并构建最优模型对于提高人工神经网络模型的预测精度尤为重要。

AZ31是目前应用最广泛的变形镁合金^[8],国内外已对其加工过程中的组织结构、耐腐蚀性、晶粒细化等方面进行了大量研究^[9-11],不过关于人工神经网络技术对其进行性能预测的研究还未见报道。本研究基于AZ31镁合金力学性能的实验数据,采用人工神经网络技术对AZ31退火工艺参数及抗拉强度(UTS)、屈服强度(YS)和延伸率(ELO)进行建模,其中通过全排列组合训练的方式对建模参数进行了优化,并将该模型与传统方法构建的模型进行比较,最后对优化参数所建立模型的预测性能进行了研究。

1 试验材料与方法

试验所用材料为3 mm厚的AZ31镁合金板材,合金的化学成分见表1。在板材上沿着与轧制方向成0°(纵向)、45°(斜向)、90°(横向)的3个方向进行取样。退火处理使用电阻炉加热,退火温度分别设定为200、250、300、350℃,每一种温度下,均设定了4种退火时间:30、60、90、120 min。试样随炉升温,保温后出炉空冷。试样拉伸力学性能在CMT5000系列微机控制电子万能试验机上进行,按照GB/T 228—2002标准取样和实验。

表1 AZ31镁合金的化学成分

元素	Al	Zn	Mn	Si	Cu	Fe	Ni	Mg
质量分数/%	3.20	1.05	0.20	0.019	0.005	0.005	0.005	余量

2 人工神经网络模型的设计

关于人工神经网络原理、设计及应用的详细论述可见参考文献^[12-14]。神经网络的设计主要包括网络结构、传输函数、学习算法等方面。除了网络结构、建模参数对网络性能有着直接的影响外,样本数据模式及预处理方式也是影响网络性能的重要因素^[7, 14-16]。

2.1 样本数据的组织

AZ31合金经力学性能测试后获得48组实验数据,将实验数据顺次循环分为训练样本、检验样本和确认样本3部分,即训练样本用于构建网络模型,为第1,3,⋯,47组(共24组);检验样本为验证模型预测性能,对训练并无影响,取第2,6,⋯,46组(12组);确认样本为保证训练精度,防止陷入局部误差,取第4,8,⋯,48(共12组)。

不管是何种样本,何种数据模式,在参与模型的训练之前,都必须经过数据预处理。本研究将样本数据处理到[0,1]的范围之内。

2.2 网络结构及参数选择

2.2.1 网络结构

研究采用的是目前应用最广泛的基于误差反向传播算法的BP(back-propagation)网络模型。根据Kolmogorov定理^[17],对于大多数实际问题,单隐层即三层网络就已足够。因此选用单隐层网络结构:输入层、中间层(隐层)及输出层。以取样方向、退火温度及退火时间作为3个输入参数,以UTS、YS及ELO作为3个输出参数。所建立的神经网络模型的结构示意图如图1所示。

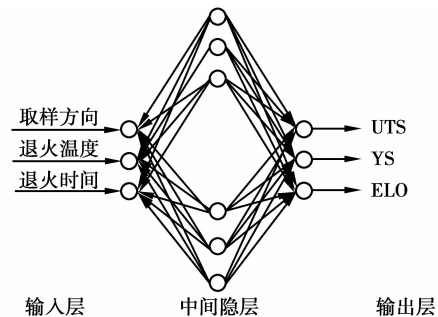


图1 BP神经网络模型结构示意图

中间隐层单元数的选择是神经网络建模的关键,根据经验公式^[18](1)预设范围为3~15个,需经后续研究才能最终确定。式中 h 为中间隐层的神经元数目, m 为输入层单元数, n 为输出层单元数。 a

为常数,一般取 1~10 之间。

$$h = \sqrt{m+n} + a. \quad (1)$$

2.2.2 传输函数的选择

常用 BP 网络的传输函数有 S 型函数(logsig)或者双曲正切函数(tansig),有时也会用到线性函数(purelin)。当网络的最后一层采用曲线函数时,输出往往被限制在一个很小的范围内,如果采用线性函数则输出可为任意值。根据本例数据非负及非线性的实际,输入层到隐层的传输函数可选取 logsig 或 tansig,隐层到输出层的传输函数直接设定为 purelin。

2.2.3 学习算法的选择

研究采用目前应用较多的基于非线性最小二乘法的 Levenberg-Marquardt(LM)算法,其根据迭代结果动态调整阻尼因子,可使每次迭代误差函数有所下降,从而使学习时间缩短,收敛速度加快,并且可以提高网络精度。

2.2.4 其他参数的选择

网络模型的正常工作还与一些参数的正确设定分不开,误差反向传播算法的主要参数为学习速率(lr)和动量项系数(mc),两者一般都设在 0~1 之间。研究中学习速率 lr 和动量项系数 mc 均取以下值:0.000 1、0.001、0.01、0.1~0.9,其中 0.1~0.9 之间以 0.1 递增。

综上设计,人工神经网络模型的建模参数选择如表 2 所示。

表 2 人工神经网络模型预设的建模参数

参数	初步设计值	数量
隐层神经元个数	3~15,间隔为 1	13
传输函数 (输入层到隐层)	logsig、tansig	2
传输函数 (隐层到输出层)	purelin	1
学习算法	trainlm	1
学习速率	0.000 1、0.001、0.01、0.1~0.9,其中 0.1~0.9 之间以 0.1 递增	12
动量项系数	0.000 1、0.001、0.01、0.1~0.9,其中 0.1~0.9 之间以 0.1 递增	12

3 人工神经网络模型的优化

3.1 建模参数的优化

在完成数据预处理、网络结构设计和建模参数选择之后就可以进行模型的训练,以期得出最优的建模参数,并进一步构建精度最高的人工神经网络模型。如表 2 所示,建模参数虽经过初选,但仍然存在 3 744 种组合(13×2×1×1×12×12)。如果采用传统的试探法寻找最优建模参数,其工作量相当大,特别是由于每一次训练都会得到一个最优模型及相应的建模参数,在有限次的比较中很难找到真正最优的参数组合,因此试探法随意性大且效率低下。

为彻底避免试探法的缺点,本研究采用了一种新的参数优化方法——建模参数全排列组合训练,即基于 MATLAB 编程,利用循环嵌套对全部预设的参数进行全排列组合的建模训练。由于部分参数组合可能引起训练震荡,造成模型精度的起伏,为进一步排除这些不稳定的建模参数,对每一种组合均进行 10 次建模训练,最后对输出的各模型平均相关系数和平均训练误差作综合比较,最终得出最优的建模参数。结果如表 3 所示。

表 3 最优建模参数

隐层神经元个数	传输函数(输入层到隐层)	传输函数(隐层到输出层)	学习算法
8	tansig	purelin	trainlm
学习速率	动量项系数	平均相关系数	平均误差 /%
默认值	默认值	0.85	3.38

学习速率和动量项系数确定为系统默认值是由于训练过程中其值的变化未对模型精度构成影响。这表明在整体数据样本较少的情况下,采用收敛速度极快的 LM 算法,网络学习时间将大大缩短,训练只需数步即达到性能误差下降的默认最小梯度。在这种情况下,学习速率和动量项系数还来不及产生影响,训练即停止。从实际来看,每次网络训练几乎都在 10 步以内完成。

根据表 3,构建 AZ31 镁合金力学性能 BP 模型的最优参数:模型结构为 3—8—3,前后传输函数分别为 tansig,purelin,算法为 trainlm,其余参数均采

用默认值。

3.2 模型的比较

研究也采用了试探法寻找最优建模参数,为尽量降低随意性,增大了每组参数的建模训练次数并以统计结果作为选择最优参数的判断标准。具体设计如下:随机设定好各建模参数,第一轮建模只改变隐层神经元个数,由此确定最优个数,在此基础上以此类推,顺次确定其余各参数。由于部分参数待选太多,只能按级差的方式来选取。最终由试探法寻找到了最优建模参数,其所建模型与全排列组合训练参数所构建的模型比较如表4所示。

表4 试探法参数所建模型与全排列组合训练参数所建模型的比较

确定参数的方法	试探法	全排列组合训练
隐层神经元个数	4	8
传递函数	logsig	tansig
学习速率	均可	均可
动量项系数	均可	均可
平均相关系数	0.84	0.85
平均误差/%	3.55	3.38

可见,2种方法确定的参数在隐层神经元个数及学习速率方面完全不同,学习速率及动量项系数方面再次证明了在LM算法的前提下,其对较小量的数据样本训练无影响。而两类模型比较,由全排列组合训练得到参数所构建的模型具有更高的平均相关系数和更低的平均误差,表明其具备更准确的预测性能。

3.3 最优模型的建立

利用全排列组合训练得到的最优参数可以构建人工神经网络模型,但是模型的构建受初始权重和阈值影响很大,而它们又是人工神经网络随机给定的,每一次建模训练都能获得一个精度较高的模型,这些模型的平均相关系数及平均误差在表4所示的对应值附近起伏,因此要得到精度越高的模型越需多次建模。经比较,目前得到的最优模型其平均相关系数已达到0.89,平均误差下降为2.91%。如图2所示,最优模型对UTS、YS、ELO的相关系数R分别达到了0.91、0.90和0.86,表明最优模型对3种力学性能都具有优良的预测性能,其中以对

UTS为最佳,YS次之,ELO最差。

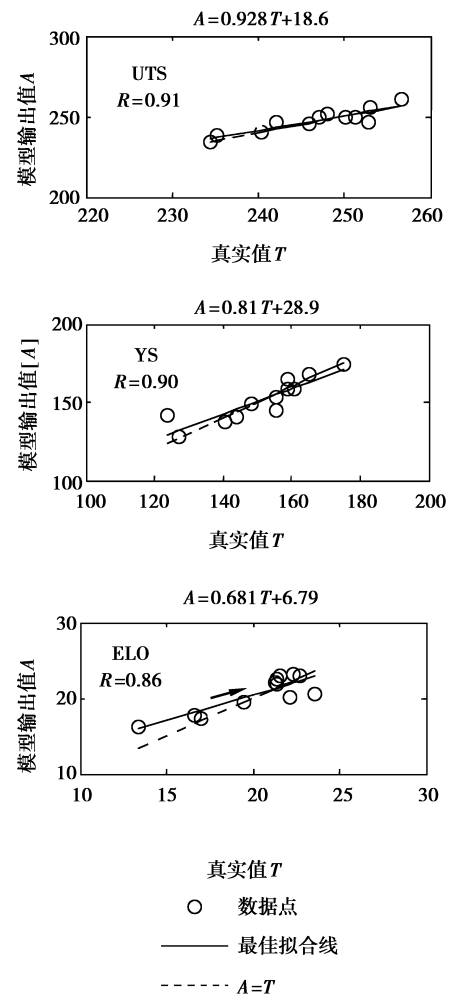


图2 最优模型对UTS、YS和ELO的拟合输出值与真实值的分析示意图

4 人工神经网络模型的验证

为验证最优神经网络模型的预测性能,对与建模训练无直接关系的检验数据组进行了预测,得到的预测值对比真实值的结果如表5所示。可见,目前得到的最优模型具有优良的预测性能,其对UTS的预测误差仅为0.98%,对YS的误差为3.03%,对ELO的误差为6.33%。同时也发现,预测误差大的数据大多出现在检验组的首尾,这是由于原始样本数据的首尾处训练、确认数据比中部的训练、确认数据少,以致相应部分拟合能力较差。这表明增加训练样本将有助于提高模型的预测和泛化能力。

表 5 实验真实值与最优模型预测值的比较

性能 级别	UTS			YS			ELO		
	真实值 /MPa	预测值 /MPa	误差 %	真实值 /MPa	预测值 /MPa	误差 /%	真实值/ /%	预测值 /%	误差 /%
1	250.1	249.4	0.28	123.6	141.1	14.16	21.4	21.9	2.34
2	256.6	260.9	1.68	159.1	164.9	3.65	22.3	23.1	3.59
3	247.0	249.7	1.09	175.1	174.2	0.51	16.6	17.8	7.23
4	248.1	251.7	1.45	148.4	149.1	0.47	21.6	23.0	6.48
5	253.0	255.2	0.87	159.1	158.6	0.31	21.4	22.5	5.14
6	251.3	249.6	0.68	165.3	167.7	1.45	19.5	19.4	0.51
7	242.2	246.1	1.61	144.0	140.6	2.36	21.3	22.1	3.76
8	252.8	246.1	2.65	155.6	144.7	7.01	23.6	20.5	13.14
9	240.3	240.2	0.04	160.9	158.2	1.68	17.0	17.3	1.76
10	235.2	238.3	1.32	127.1	128.1	0.79	22.1	20.0	9.50
11	245.9	245.8	0.04	140.4	137.3	2.21	22.7	22.9	0.88
12	234.4	234.5	0.04	155.6	152.9	1.74	13.4	16.3	21.64
平均值			0.98			3.03			6.33

5 结 论

1) 人工神经网络技术能充分描述 AZ31 镁合金的取样方向、退火温度、退火时间与抗拉强度、屈服强度、延伸率之间的映射关系,适用于研究具有非线性关系的领域。

2) 采用参数全排列组合训练的方法可以获得比传统试探法更优的建模参数,基于这些参数建立的 AZ31 镁合金力学性能的最优人工神经网络预测模型,其平均相关系数可达 0.89,平均误差下降为 2.91%。

3) 基于参数优化的 AZ31 镁合金力学性能的人工神经网络预测模型能够比较准确地预测 AZ31 镁合金在不同退火条件后的力学性能。

参考文献:

- [1] KABIRIAN F, MAHMUDI R. Impression creep behavior of a cast AZ91 magnesium alloy [J]. Metallurgical and Materials Transactions a-Physical Metallurgy and Materials Science, 2009, 40A(1): 116.
- [2] HAMDY A S. The effect of surface modification and stannate concentration on the corrosion protection performance of magnesium alloys [J]. Surface & Coatings Technology, 2008, 203(3/4): 240.
- [3] AGHION E, BRONFIN B. Magnesium alloys development towards the 21st century[J]. Materials Science Forum, 2000, 350/351:19-28.
- [4] 王渠东,丁文江. 镁合金研究开发现状与展望[J]. 世界有色金属,2004(7):8-12.

WANG QU-DONG, DING WEN-JIANG. Status and prospect of research and development for magnesium alloys[J]. World Nonferrous Metals, 2004(7): 8-12.

- [5] 袁志强,汤爱涛,刘海定,等. 基于 Web 的镁合金数据库系统的研究和设计[J]. 重庆大学学报,2007,30(2):94.
- [6] 傅定发,许芳艳,夏伟军,等. 退火工艺对轧制 AZ31 镁合金组织和性能的影响[J]. 湘潭大学自然科学学报,2005,27(4):57.
- [7] 吴良. 提高 BP 神经网络热处理工艺及材料性能预测模型的泛化能力研究[J]. 金属热处理,2003,28(5):42-45.
- [8] 肖盼,刘天模,姜丹. AZ31 镁合金的研究进展[J]. 重庆大学学报,2006,29(11):81.
- [9] ANDO D, KOIKE J, SUTOU Y. Relationship

YUAN ZHI-QIANG, TANG AI-TAO, LIU HAI-DING, et al. Research and design of magnesium alloys database systems based on web [J]. Journal of Chongqing University, 2007, 30(2): 94.

FU DING-FA, XU FANG-YAN, XIA WEI-JUN, et al. The effect of annealing on microstructure and mechanical properties of rolled AZ31 magnesium alloy[J]. Natural Science Journal of Xiangtan University, 2005, 27(4): 57.

WU LIANG. How to improve the generalized ability of ANN-based BP model used to predict technologies of heat treatment and properties of materials[J]. Heat Treatment of Metals, 2003, 28(5): 42-45.

XIAO PAN, LIU TIAN-MO, JIANG DAN. Research status and development of AZ31 magnesium alloy[J]. Journal of Chongqing University, 2006, 29(11): 81.

- between deformation twinning and surface step formation in AZ31 magnesium alloys [J]. *Acta Materialia*, 2010, 58(13): 4316-4324.
- [10] ZHANG J, YANG D H, OU X B. Microstructures and properties of aluminum film and its effect on corrosion resistance of AZ31B substrate [J]. *Transactions of the Nonferrous Metals Society of China*, 2008, 18 (z1): 312-317.
- [11] MIAO Q, HU L X, WANG X, et al. Grain growth kinetics of a fine-grained AZ31 magnesium alloy produced by hot rolling [J]. *Journal of Alloys and Compounds*, 2010, 493(1/2): 87-90.
- [12] FINE T L. Feedforward neural network methodology [M]. New York: Springer Verlag, 1999.
- [13] MARTIN T H, HOWARD B D, MARK H B. 神经网络设计[M]. 戴奎,译.北京:机械工业出版社,2002.
- [14] 张际先,宓霞. 神经网络及其在工程中的应用[M]. 北京:机械工业出版社,1996.
- [15] LIU H D, TANG A T, PAN F S, et al. A model on the correlation between composition and mechanical properties of Mg-Al-Zn alloys by using artificial neural network [J]. *Magnesium-Science, Technology and Applications*, 2005, 488/489: 793-796.
- [16] TANG A T, LIU B, PAN F S, et al. An improved neural network model for prediction of mechanical properties of magnesium alloys[J]. *Science in China Series E-Technological Sciences*, 2009, 52(1): 155-160.
- [17] CHANG C C, CHANG T Y P, XU Y G, et al. Selection of training samples for model updating using neural networks [J]. *Journal of Sound and Vibration*, 2002, 249(5): 867-883.
- [18] 飞思科技产品研发中心. 神经网络理论与 MATLAB7 实现[M]. 北京:电子工业出版社,2005.

(编辑 王维朗)

(上接第 43 页)

- [13] ANTONINO V, FRANCISCO V, STEFANO B. The impact of virtual technologies on knowledge-based processes: an empirical study [J]. *Research Policy*, 2009, 38(2): 1278-1287.
- [14] GUO T, SCHWARTZ D G, BURSTEIN F, et al. Codifying collaborative knowledge: using Wikipedia as a basis for automated ontology learning [J]. *Knowledge Management Research & Practice*, 2009, 7(1): 206-217.
- [15] 刘长勇,宁正元. 基于 XML 的学习流模型[J]. 重庆工学院学报:自然科学, 2009, 23(2): 176-180.
- LIU CHANG-YONG, NING ZHEGN-YUAN. The studying flow model based on XML [J]. *Journal of Chongqing Institute of Technology: Natural Science*, 2009, 23(2): 176-180.
- [16] LITTLE E G, ROGOVA G L. Designing ontologies for higher level fusion [J]. *Information Fusion*, 2009, 10(1): 70-82.
- [17] 朱征宇,李存青,张鹏. 基于语法模式的产品评论主题词和积极性词提取[J]. 重庆理工大学学报:自然科学, 2010, 24(5): 86-90.
- ZHU ZHENG-YU, LI CUN-QING, ZHANG PENG. Topic words and opinion words extraction from chinese productsreviews based on syntax pattern[J]. *Journal of Chongqing University of Technology: Natural Science*, 2010, 24(5): 86-90.
- [18] ANTONINO V, FRANCISCO V, STEFANO B. The impact of virtual technologies on knowledge-based processes: an empirical study [J]. *Research Policy*, 2009, 38(2): 1278-1287.

(编辑 张 苹)