

正交递归方法用于研究钢中微量元素 对钢的力学性能的影响^①

李 静 章 元 朱 尔^②

(厦门大学化学系 教育部材料和生命过程分析科学开放研究实验室 厦门 361005)

摘要 通过对含有 8 种微量元素和 5 种热轧工艺条件的 CuPCrNi 钢的数据,应用化学计量学方法(包括变量筛选及 PLS 回归等方法)进行处理,选出对钢性能影响较为显著的几种变量,从而建立起预报 CuPCrNi 钢性能的准确的预报模型。

关键词 变量筛选, PLS 回归, 钢, 微量元素

中国图书分类号 TG 142

钢的主要性能指标为屈服强度(σ_s)、抗拉强度(σ_b)、延伸率(δ_5)。影响钢性能的主要因素是钢中所含的微量元素及一些热处理工艺条件。如:钢中含碳量越高,钢的硬度越高,但塑性及任性差;磷的存在可使钢的塑性及韧性明显下降,特别是低温时,影响更严重;硫是钢中的有害物质,含硫较多的钢在高温下加压时容易脆裂;锰可提高钢的淬透性和耐磨性,改善钢的抗氧化作用,提高耐腐蚀能力;镍可提高钢的强度,韧性和淬透性,含量高时,可显著改善钢的一些物理特性,提高耐腐蚀能力;硅可使钢的硬度提高,但塑性,韧性下降;铬可提高钢的淬透性行耐磨性,改善钢的抗氧化能力,提高耐腐蚀能力……等。此外,钢的热处理工艺条件,如:初轧温度,终轧温度等也对钢的性能有一定的影响。为了研究钢性能的影响因素,对钢中所含微量元素及热轧工艺条件的研究是很有必要的。

本研究对从某钢铁公司生产轧钢现场得到的有关数据进行处理,考虑到变量间的非线性关系,采用了变量扩维及压缩筛选方法(即引入原始变量的非线性项,包括变量的平方项,所有二次交叉项及变量两两相除项,再用正交递归选择法^[1]对所有的变量进行压缩筛。该法大大降低所建模型的 PRESS 值,提高预报模型的准确性)。用 PLS^[2]方法对由筛选出的变量所组成的数据进行处理,建立起准确的预报模型,并用此预报模型预报钢的性能。若能通过此预报模型确定一定的钢的性能指标所需要的元素含量和热处理条件,这对于生产实践是非常有益的,它有助于新钢种的设计。

1 数据采集

本研究工作用到的 15MnHP 钢、CuPCrNi 钢及 T521 钢的数据均来自某钢铁公司,因而

① 本文 1998-11-27 收到; 教育部留学回国人员科研启动经费(无编号)及福建省自然科学基金资助项目(B93016)

② 联系人

有一定的代表性,在此仅以 CuPCrNi 钢为例.表 1 中列出了 CuPCrNi 钢中所有的 8 种微量元素,5 种热轧工艺条件及 3 种性能指标的平均值及标准偏差,样本总数为 105 个,包括 13 个变量(元素与热轧工艺条件)和 3 个目标变量(性能指标).

表 1 CuPCrNi 钢的工艺条件与性能指标

Tab 1 Technical Conditions and Properties of CuPCrNi Steel

元素	平均值 /(%)	标准偏差 /(%)	工艺 条件	平均 值	标准 偏差	性能 指标	平均 值	标准 偏差
C	0.071	0.022	T (s)	155.56	41.57	$\sigma_s / \text{kg} \cdot \text{cm}^{-2}$	38.14	2.52
Si	0.316	0.057	CL-T/°C (c)	1082.17	25.20	$\sigma_b / \text{kg} \cdot \text{cm}^{-2}$	50.89	2.72
Mn	0.374	0.070	CZ-T/°C (c)	1044.86	24.78	$\delta 5 / \%$	33.82	2.37
P	0.0099	0.0009	ZZ-T/°C (c)	883.37	20.24			
S	0.019	0.003	JQ-T/°C (c)	677.93	33.82			
Cu	0.300	0.017						
Cr	0.439	0.074						
Ni	0.355	0.040						

注: T: 在炉时间; CL-T: 出炉温度; CZ-T: 初轧温度; ZZ-T: 终轧温度; JQ-T: 卷曲温度; σ_s : 屈服强度; σ_b 抗拉强度; $\delta 5$ 延伸率. 以下同)

2 数据处理结果与讨论

2.1 数据预处理

研究样本总数 105 个, 变量 13 个, 目标变量 3 个. 分别建立每个目标变量和变量之间的数据模型, 得到每组均含有 13 个变量, 1 个目标变量的 3 组数据. 为了使建立的模型更准确, 预报能力更强, 先对每组数据进行异常点的删除, 再将

表 2 数据

Tab 2 Data

目标变量	变量个数	建模数据个数	预报数据个数
屈服强度 (σ_s)	13	60	29
抗拉强度 (σ_b)	13	61	28
延伸率 ($\delta 5$)	13	61	29

剩下的数据分为两部分, 一部分用于建模, 另一部分用于预报检验(具体数据情况如表 2). 数据在进行分析前, 先进行标准化, 使各变量均值为 0 均方差为 1.

2.1 扩维与正交递归选择法筛变量结果

为了能建立较准确的预报模型, 对原始数据采用变量扩维及压缩筛选. 由于自变量与目标变量之间存在非线性关系, 因此在筛选的变量中包含各变量的线性与非线性因子, 非线性因子包括变量的平方项, 所有二次交叉项及变量两两相除项, 将所有的变量全部参加筛选, 即对原有变量进行扩展. 变量筛选用 PRESS 值为判据, 根据 PRESS^[3] 值为最低或接近最低, 从几百个变量里选出 6~7 个含信息量较多的变量(表 3~5).

表 3 筛选出的影响 CuPCrNi 钢的屈服强度的非线性因素

Tab 3 Selected nonlinear factors for Yield-Intensity (σ_s) of CuPCrNi Steel

参 数	Si/Cu	Mn/Si	JQ-T	Cu*Cr	Si*Cr	Mn/Cr	SiMn
平均值	1.02	1.22	677.53	0.13	0.14	0.88	0.12
平均标准偏差	0.19	0.17	38.65	0.03	0.04	0.11	0.04

表 4 筛选出的影响 CuPCrNi 钢的抗拉强度的非线性因素

Tab 4 Selected nonlinear factors for (σ_b) of CuPCrNi Steel

参数	Si*Cr	Cu/C	CZ-T* JQ-T	S/Cr	S*Cu	Si/Cu
平均值	0.15	4.72	709.733.13	0.044	0.0058	1.07
平均标准偏差	0.04	1.39	27.644.20	0.008	0.0010	0.20

表 5 筛选出的影响 CuPCrNi 钢的延伸率的非线性因素

Tab 5 Selected nonlinear factors for Extend-Rate(δ_5) of CuPCrNi Steel

参数	Si*Cr	S	CZ-T/t	JQ-T	Ni/P	P*Cr	Si/Ni
平均值	0.16	0.019	7.22	106.959.27	3.52	0.046	0.99
平均标准偏差	0.03	0.002	1.98	32.750.08	0.51	0.006	0.16

2.1 预报模型

根据 PLS 正交递归选择法筛选出的对各个目标变量影响较显著的几个变量,再用 PLS 法对选出的变量进行数据处理,并根据 PRESS 值最低,删除含噪音多的隐变量,建立对各个目标变量的预报模型,模型系数见表(6~8),其中系数 1 为数据标准化后的模型系数,系数 2 为数据未经过标准化的模型系数。

表 6 筛选出的非线性因子及模型系数— σ_s

Tab 6 Selected nonlinear factors and model coefficient

变量	模型系数 1	模型系数 2	影响顺序
Si/Cu	9.745.007	50.190.342	2
Mn/Si	2.218.719	12.758.205	5
JQ-T	-0.214.460	-0.005.549	7
Cu*Cr	7.220.819	264.172.455	3
Si*Cr	-9.848.641	-225.359.05	1
Mn/Cr	-0.623.124	-5.660.750	6
SiMn	-2.352.404	-57.188.793	4
Const	0	-17.066.235	

表 7 筛选出的非线性因子及模型系数— σ_b

Tab 7 Selected nonlinear factors and model coefficient(σ_b)

变量	模型系数 1	模型系数 2	影响顺序
Si*Cr	-1.606.002	-41.323.631	3
Cu/C	-1.054.409	-0.761.180	5
CZ-T* JQ-T	-0.328.125	-0.000.012	6
S/Cr	-2.390.469	-2.390.469	1
S*Cu	1.797.095	1.797.095	2
Si/Cu	1.475.509	1.475.509	4
Const	0	63.167.801	

表 8 筛选出的非线性因子及模型系数— δ_5 Tab 8 Selected non linear factors and model coefficient (δ_5)

变量	模型系数 1	模型系数 2	影响顺序	变量	模型系数 1	模型系数 2	影响顺序
Si Cr	- 2 922 098	- 91 814 812	3	Ni/P	1 886 752	3 673 081	4
S	- 0 039 611	- 14 320 804	7	P Cr	1 333 590	206 150 925	5
CZ-T/t	7 082 520	3 575 816	1	SiNi	1 203 538	7 699 440	6
T* JQ-T	6 213 109	0 000 019	2				

用先前预留下的样本代入各目标变量的预报模型, 可得到模型的预测值, 预测准确率为 90%, 比不经过变量扩维及筛选直接用 PLS 回归所预报的准确率高 10%。这一结果说明变量选择的准确性对模型预报的准确性有重大影响, 因为模型中自变量与目标变量可能存在非线性关系, 所以只考虑线性关系, 用 PLS 等方法处理, 得到的模型的预报准确性较低。以上分析还可得出, 对于 CuPCNi 钢的性能指标有较大影响的因素及排列顺序为

$$\sigma_s \text{ Si Cr, Si/Cu, Cu Cr Si Mn}$$

$$\sigma_b \text{ S/Cr, S Cu, Si Cr Si/Cu}$$

δ_5 初轧温度 在炉时间, 在炉时间* 卷曲温度, Si Cr, Ni/P.

参 考 文 献

- 1 朱尔一, 杨夙原, 邓志威等. 正交递归选择法及应用. 高等学校化学学报, 1993 14 1 518~ 1 521
- 2 朱尔一, 杨雍静, 庄峙厦等. 一种新的适合处理两类样本判别分析问题的 PLS 方法. 高等学校化学学报, 1997 18(2): 212~ 215
- 3 Myers R H. Classical and Modern Regression with Application. Boston, Massachusetts: Duxbury Press, 1986 105~ 108

Multi-variable Analysis Method Used for Studying Influence of Trace Elements in Steel on its Mechanical Properties

Li Jing Zhang Yuan Zhu Eryi

(Dept of Chem. The SEDC Res Lab of Anal Sci for Material and Life Chem. Xiamen Univ., Xiamen 361005)

Abstract The data of CuPCNi Steel includes 8 trace elements and 5 heat treatment parameters. Through the application of Chemometrics methods (including the variable dimension expansion and the variable selection method, PLS regression method and so on), several primary variables which influence the properties of the steel, can be selected. Thus accurate prediction models to predict the properties of the steel can be built.

Key words Variable selection, PLS regression, Trace element, Steel