

用 MATLAB 程序预测地平类药物的熔点

李志伟¹, 郭海谦¹, 吴 芬²

(1. 河北科技大学化学与制药工程学院, 河北石家庄 050018; 2. 石药集团欧意药业有限公司, 河北石家庄 050051)

摘 要:分析地平类药物分子结构和熔点之间关系, 编制相应的 MATLAB 程序, 分别建立回归方程和人工神经网络 2 种方法的模型。选取 14 个地平类药物作为回归方程和人工神经网络 2 种方法建模的样本, 再选另外 4 个地平类药物作验证, 分别进行 2 种方式的建模和测试对比。结果表明: 人工神经网络模型比回归方程方法在预测地平类药物的熔点时取得更加准确的估计。

关键词: MATLAB; 回归方程; 神经网络; 药物筛选; 数值预测; 熔点

中图分类号: R911 文献标志码: A

MATLAB applications for the forecast of dipine drugs' melting point

LI Zhiwei¹, GUO Haiqian¹, WU Fen²

(1. School of Chemical and Pharmaceutical Engineering, Hebei University of Science and Technology, Shijiazhuang Hebei 050018, China; 2. Ouyi Pharmaceutical Company Limited, CSPC, Shijiazhuang Hebei 050051, China)

Abstract: Two models of regression equation (RE) and artificial neural network (ANN) are established by analyzing the relationship between the dipine drugs' molecular structure and melting points and compiling the MATLAB program. Selecting 14 dipine drugs as the modeling samples of RE and ANN methods and other 4 dipine drugs for verification, the two models are compared. And results show that the ANN model is more accurate in forecasting the drugs' melting point than the RE method.

Key words: MATLAB; regression equation; neural network; drug screening; data forecast; melting point

药物的性质很大程度上取决于药物分子的结构, 通过分析药物分子的结构, 可以发现药物的性质和结构之间存在的某种定量关系^[1]。许多因素都影响药物的各种理化性质, 如在化合物分子中分子的电荷分布及立体因素、官能团的位置等。许多学者也用人工神经网络或回归方程等方法对各种药物进行研究^[2-14]。

1 参数及模型

根据文献[4]中的方法对地平类药物构建相似的分子参数 N 和 T , 与该药物的熔点进行关联, 构建预测的数学模型, 得到地平类药物的熔点计算值。普拉地平分子结构见图 1。

例如: 对普拉地平的分子参数 N, T 进行计算:

$$T = (\delta_o)^{-0.5} \times 6 + (\delta_{NH})^{-0.5} + (\delta_N)^{-0.5} + (\delta_C)^{-0.5} \times 9 + (\delta_{CH})^{-0.5} \times 12 + (\delta_{CH_2})^{-0.5} + (\delta_{CH_3})^{-0.5} \times 3 = 1.808 \times 6 + 0.280 + 1.191 + 0.5 \times 9 + 0.577 \times 12 + 0.707 + 1 \times 3 = 27.449,$$
$$N = 0.2 + 1/3 = 0.533.$$

收稿日期: 2012-11-05; 修回日期: 2012-11-30; 责任编辑: 张士莹

基金项目: 河北省自然科学基金资助项目(B2010000849)

作者简介: 李志伟(1977-), 男, 河北保定人, 副教授, 博士, 主要从事药物分析和分离材料方面的研究。

E-mail: lizhiwei@hebust.edu.cn

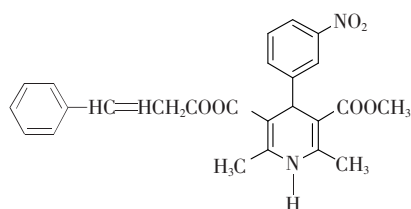


图 1 普拉地平分子结构

Fig. 1 Molecular structure of pranidipine

依照上述计算方法,将地平类药物的熔点、分子参数 N 和 T 值列入表 1。

表 1 地平类药物熔点实验值与分子参数 N, T Tab. 1 Melting point of dipine drugs and molecular parameters N, T

序号	药物名称	熔程/°C	N	T
1	硝本地平	172~174	0.700	23.204
2	尼群地平	156~159	0.533	23.911
3	氯氯地平	178~179	0.700	24.024
4	尼莫地平	124~128	0.533	28.003
5	尼卡地平	178~185	0.533	29.901
6	尼索地平	148~152	0.700	25.488
7	尼伐地平	148~150	0.533	25.472
8	非洛地平	142~145	0.400	19.569
9	阿折地平	120~126	0.533	33.380
10	拉西地平	174~175	0.700	28.581
11	马尼地平	148~150	0.533	36.175
12	乐卡地平	118~120	0.533	36.363
13	西尼地平	105~109	0.533	30.672
14	氯硝地平	75~80	0.533	37.559
15	普拉地平	143.5~144.5	0.533	27.449
16	阿雷地平	150~151	0.700	26.219
17	巴尼地平	137~139	0.533	28.063
18	伊拉地平	168~170	0.400	24.087

2 编制程序

2.1 回归方程程序

```

mmin=[172 156 178 124 178 148 148 142 120 174 148 118 105 75];
mmax=[174 159 179 128 185 152 150 145 126 175 150 120 109 80];
format bank;
mpm=(mmin+mmax)/2%程序所用熔点均为熔程平均值
n=[0.7 0.533 0.7 0.533 0.533 0.7 0.533 0.4 0.533 0.7 0.533 0.533 0.533 0.533]';
t=[23.204 23.911 24.024 28.003 29.901 25.488 25.472 19.569 33.380 28.581 36.175 36.363
30.672 37.559]';
y=[mpm]';
x=[ones(14,1) n t];%对数据拟合新的方程,同时和文献中方程作图对比;
[b,bint,r,rint,stats]=regress(y,x);
figure;

```

```

rcoplot(r,rint);
pre_mp=~6.393 1 * t+107.425 78 * n+260.080 9;
mp=b(3) * t+b(2) * n+b(1);
no=1:14;
figure;
plot(no,mmin,'ro',no,pre_mp,'* ',no,mmax,'ro');
figure;
plot(no,mmin,'ro',no,mp,'* ',no,mmax,'ro');
[t1,n1]=meshgrid(18:2:38,0.8:~0.1:0.3);
mp3=~6.393 1 * t1+107.425 78 * n1+260.080 9;
mp4=b(3) * t1+b(2) * n1+b(1);
figure;
set(gcf,'defaultlinelinerwidth',1.5);
plot3(t,n,pre_mp,'~y*'),hold on,surf(t1,n1,mp3),shading interp,hold on,...
plot3(t,n,mp,'~m'),hold on,surf(t1,n1,mp4),shading interp,...
hold on,plot3(t,n,mmin,'o'),hold on,plot3(t,n,mmax,'v'),xlabel('T'),ylabel('N'),...
zlabel('melting point'),title('Surfaces Before and after the Curve Fitting');
choose=input('Do you want to test 4 groups of data,input " 1" to go on or " 0" to stop. ');
if (choose==0),break;
else
n15=input('test 15th N: ');
t15=input('test 15th T: ');
Test_mp15=b(3) * t15+b(2) * n15+b(1)
Pre_mp15=~6.393 1 * t15+107.425 78 * n15+260.080 9
余下 3 个药物验证程序略去。

```

2.2 神经网络程序

数据输入部分略。

```

P(2,:)=(p(2,.)~min(p(2,:)))/(max(p(2,:))~min(p(2,:))); % p 为参数 T,N 组成的矩阵
T=(t~min(t))/(max(t)~min(t)); % t 为熔程平均值
rand('state',0);
%net=newff(minmax(P),[6,1],{'tansig','purelin'},'trainlm');
net=newff(minmax(P),[9,1],{'tansig','purelin'},'traingdx');
%以上两行可交替注释屏蔽训练
net=init(net)
net.trainParam.epochs = 10 000;
net.trainParam.goal = 0.01;
[net,tr]=train(net,P,T);
mmin=[172 156 178 124 178 148 148 142 120 174 148 118 105 75];
mmax=[174 159 179 128 185 152 150 145 126 175 150 120 109 80];
format bank;
mp=sim(net,P). * (max(t)~min(t))+min(t)
no=1:14;
figure;
plot(no,mmin,'ro',no,mp,'* ',no,mmax,'ro');
choose=input('Do you want to test 4 groups of data,input " 1" to go on or " 0" to stop. ');
if (choose==0),break;

```

```

else
N15=input('test 15th N: ');
t15=input('test 15th T: ');
T15=(t15~min(p(2,:)))/(max(p(2,:))~min(p(2,:)));
Test_mp=sim(net,[N15 T15]).*(max(t)~min(t))+min(t)
余下 3 个药物验证程序略去。

```

3 运行结果

3.1 回归方程程序运行结果

程序拟合所得方程为

$$\text{Test_mp} = \sim 2.85T + 128.79N + 151.85. \quad (1)$$

文献[4]中方程为

$$\text{Pre_mp} = \sim 6.39T + 107.43N + 260.08. \quad (2)$$

回归方程法熔点计算值及误差见表 2。

表 2 回归方程法熔点计算值及误差

Tab. 2 Calculation values of RE method and errors

样本序号	药物名称	方程 1 熔点/°C	Test_mp 误差/%	方程 2 熔点/°C	Pre_mp 误差/%
1	硝本地平	175.87	1.66	187.01	8.10
2	尼群地平	152.35	3.27	164.55	4.48
3	氨氯地平	173.53	2.78	181.77	1.83
4	尼莫地平	140.69	11.66	138.40	9.84
5	尼卡地平	135.28	25.47	126.27	30.43
6	尼索地平	169.36	12.91	172.41	14.94
7	尼伐地平	147.90	0.74	154.57	3.74
8	非洛地平	147.59	2.85	178.01	24.05
9	阿折地平	125.36	1.92	104.04	15.41
10	拉西地平	160.55	8.00	152.65	12.52
11	马尼地平	117.40	21.21	86.18	42.16
12	乐卡地平	116.86	1.80	84.98	28.59
13	西尼地平	133.08	24.37	121.35	13.41
14	氯硝地平	113.45	46.39	77.34	0.21
平均误差/%			11.79	14.98	

其余 4 组验证数据见表 3。

表 3 回归方程法熔点预测值及误差验证

Tab. 3 Predicted values of RE method and errors for verification

样本序号	药物名称	方程 1 熔点/°C	Test_mp 误差/%	方程 2 熔点/°C	Pre_mp 误差/%
15	普拉地平	142.27	1.20	141.94	1.43
16	阿雷地平	167.28	11.15	167.74	11.46
17	巴尼地平	140.52	1.82	138.02	0.01
18	伊拉地平	134.72	20.29	149.14	11.75

3.2 神经网络程序运行结果

神经网络法熔点计算值及误差见表 4。

表4 神经网络法熔点计算值及误差

Tab. 4 Calculation values of ANN method and errors

样本序号	药物名称	gdx9 熔点/°C	gdx9 误差/%	lm6 熔点/°C	lm6 误差/%
1	硝本地平	174.22	0.71	170.89	1.22
2	尼群地平	158.91	0.90	143.04	9.18
3	氨氯地平	180.08	0.89	172.67	3.27
4	尼莫地平	135.93	7.88	135.16	7.27
5	尼卡地平	155.23	14.47	171.75	5.37
6	尼索地平	146.11	2.59	146.30	2.47
7	尼伐地平	144.97	2.70	138.39	7.12
8	非洛地平	143.23	0.19	152.72	6.43
9	阿折地平	121.36	1.33	126.75	3.05
10	拉西地平	175.43	0.53	158.00	9.46
11	马尼地平	139.48	6.39	136.18	8.60
12	乐卡地平	130.85	9.96	127.27	6.95
13	西尼地平	128.16	19.78	113.47	6.05
14	氯硝地平	75.21	2.95	74.77	3.52
平均误差/%			5.09		5.71

其余4组验证数据见表5。

表5 神经网络法熔点预测值及误差验证

Tab. 5 Predicted values of ANN method and errors for verification

样本序号	药物名称	gdx9 熔点/°C	gdx9 误差/%	lm6 熔点/°C	lm6 误差/%
15	普拉地平	133.02	7.62	132.88	7.72
16	阿雷地平	154.11	2.40	137.41	8.70
17	巴尼地平	139.27	0.92	135.64	1.71
18	伊拉地平	155.35	8.08	152.55	9.73

4 结果与讨论

以上数据的结果表明,重新拟合的回归方程比文献[4]中方程更适合于计算地平类药物的熔点。但回归方程法预测结果与实验值比较平均误差在10%左右,而神经网络方法平均误差在5%左右,人工神经网络预测值的结果比回归方程更接近实验值。以上2个方法若进一步完善也可以对其他药物熔点进行更精准的预测,为进一步研究药物的其他理化性质及药物活性提供帮助。而其中的BP神经网络具有设计简单的优点^[15],具有较好的预测和泛化能力,为复杂、高度非线性问题的解决提供了途径。因此,BP神经网络可作为药物构效关系及计算机辅助药物设计的有效方法。

用MATLAB处理数据比较方便,而且能进行非线性拟合。作为高级计算机语言,MATLAB具有强大的计算能力,它将成为药物科研人员进行数据处理及试验设计的得力助手。

参考文献/References:

- [1] 宋倩,梅向东,宁君. 生物电子等排原理在农药与医药开发中的应用[J]. 河北工业科技, 2009, 26(2): 131-136.
SONG Qian, MEI Xiangdong, NING Jun. Application of bioisosterism in pesticide and medicine development[J]. Hebei Journal of Industrial Science and Technology, 2009, 26(2): 131-136.
- [2] HECHT D, CHEUNG M, FOGEL G B. QSAR using evolved neuralnetworks for the inhibition of mutant PfDHFR by pyrimethamine derivatives[J]. Biosystems, 2008, 92(1): 10-15.
- [3] FOMENKO I, DURST M, BALABAN D. Robust regression for high throughput drug screening[J]. Computer Methods and Programs in Biomedicine, 2006, 82: 31-37.

- [4] 周燕平, 焦健, 周先锋, 等. 碳酸酐酶 XIV 抑制剂的定量构效关系研究[J]. 华中师范大学学报(自然科学版), 2010, 44(4): 618-622.
ZHOU Yanping, JIAO Jian, ZHOU Xianfen, et al. QSAR studies of sulfonamides as the carbonic anhydrase XIV inhibitors by artificial neural network-based transformation for nonlinear partial least-square regression[J]. Journal of Huazhong Normal University (Natural Sciences), 2010, 44(4): 618-622.
- [5] LEOW J L, BARON R, CASEY P J, et al. Quantitative structure-activity relationship (QSAR) of indoloacetamides as inhibitors of human isoprenylcysteine carboxyl methyltransferase[J]. Bioorganic and Medicinal Chemistry Letters, 2007, 17(4): 1 025-1 032.
- [6] HEMMATEENEJAD B, SHAMSIPUR M. Quantitative structure electrochemistry relationship study of some organic compounds using PC-ANN and PCR[J]. Internet Electronic Journal of Molecular Design, 2004, 3(6): 316-334.
- [7] 周鲁, 夏昆华. 喹诺酮类化合物的定量构效关系[J]. 华西药学杂志, 2006, 21(5): 433-435.
ZHOU Lu, XIA Kunhua. The quantitative structure-pharmacokinetic relationship of quinolones compounds[J]. West China Journal of Pharmaceutical, 2006, 21(5): 433-435.
- [8] 陈国华, 陆瑶, 陈虹. 基于逐步回归所得变量集的遗传反向传播神经网络的 QSAR 研究[J]. 计算机与应用化学, 2010, 27(9): 1 257-1 262.
CHEN Guohua, LU Yao, CHEN Hong. The QSAR study of backpropagation neural network using genetic algorithm based on variable pool selected by stepwise linear regression[J]. Computers and Applied Chemistry, 2010, 27(9): 1 257-1 262.
- [9] 堵锡华, 陈艳, 唐自强, 等. 抗糖尿病新药熔点的定量构效关系研究[J]. 南京工业大学学报, 2006, 28(2): 98-100.
DU Xihua, CHEN Yan, TANG Ziqiang, et al. QSPR research for melting point of new antidiabetic drugs[J]. Journal of Nanjing University of Technology, 2006, 28(2): 98-100.
- [10] 高鸿斌, 张永强, 张佳. BP-神经网络算法在分布式 VOD 系统流量预测中的应用[J]. 河北科技大学学报, 2001, 22(3): 23-26.
GAO Hongbin, ZHANG Yongqiang, ZHANG Jia. Application of BP-neural network to network-flow prediction of distributed VOD system[J]. Journal of Hebei University of Science and Technology, 2001, 22(3): 23-26.
- [11] HANGAN M T, DEMUTH H B, BEALE M. Neural Network Design[M]. Beijing: China Machine Press, 2002.
- [12] KAHANER D, MOLER C B, NASH S. Numerical Methods and Software[M]. Englewood Cliffs: Prentice Hall, 1989.
- [13] SMITH B T, BOYLE J M, DONGARRA J J, et al. Matrix Eigensystem Routines-EISPACK Guide[M]. 2nd ed. New York: Springer-Verlag, 1976.
- [14] 郝平娇, 李士雨. 浅谈 MATLAB 在化工计算中的应用[J]. 计算机与应用化学, 2000, 17(4): 371-374.
HAO Pingjiao, LI Shiyu. Applications of MATLAB in chemical engineering calculating[J]. Computers and Applied Chemistry, 2000, 17(4): 371-374.
- [15] 冯剑. MATLAB 在化工计算中的应用[J]. 化学工程师, 2008, 22(5): 25-27.
FENG Jian. Applications of MATLAB in chemical engineering calculating[J]. Chemical Engineer, 2008, 22(5): 25-27.

~~~~~  
(上接第 53 页)

- [12] 廖海斌, 徐洪章. 基于鉴别主成份分析的基因表达数据特征提取[J]. 燕山大学学报, 2010, 34(5): 426-430.  
LIAO Haibin, XU Hongzhang. Gene expression data feature extraction based on discriminant principal component[J]. Journal of Yanshan University, 2010, 34(5): 426-430.
- [13] 李玉梅, 陈艳秋, 李莉. 梨品种枝条膜透性和水分状态与抗寒性的关系[J]. 北方果树, 2005(1): 3-5.  
LI Yumei, CHEN Yanqiu, LI Li. Relationship of between the membrane permeability, water status and hardiness in pear shoots[J]. Northern Fruits, 2005(1): 3-5.
- [14] 杨建民, 李艳华, 杨敏生, 等. 几个仁用杏品种抗寒性比较研究[J]. 中国农业科学, 1999, 32(1): 46-50.  
YANG Jianmin, LI Yanhua, YANG Minsheng, et al. Comparative study on cold resistance of apricot varieties[J]. Agricultural Sciences in China, 1999, 32(1): 46-50.
- [15] 马艳青, 戴雄泽. 低温胁迫对辣椒抗寒性相关生理指标的影响[J]. 湖南农业大学学报(自然科学版), 2000, 26(6): 461-462.  
MA Yanqing, DAI Xiongze. Influences of low temperature stress on pepper's physiological index related to resistance to coldness[J]. Journal of Hunan Agricultural University(Natural Sciences), 2000, 26(6): 461-462.
- [16] 王飞, 陈登文, 王卿, 等. 杏品种的需寒量与抗寒性的相关研究[J]. 中国农业科学, 2001, 34(5): 465-468.  
WANG Fei, CHEN Dengwen, WANG Qing, et al. Correlation study on chilling requirements and hardiness of apricot varieties[J]. Agricultural Sciences in China, 2001, 34(5): 465-468.
- [17] 杨建民, 周怀军, 王文凤. 果树霜冻害研究进展[J]. 河北农业大学学报, 2000, 23(3): 54-58.  
YANG Jianmin, ZHOU Huaijun, WANG Wenfeng. The advance in frost injury research of fruit trees[J]. Journal of Agricultural University of Hebei, 2000, 23(3): 54-58.