

溶质线性梯度保留时间的多元回归预测模型

尹小英, 姚卫峰, 胡育筑*

(中国药科大学分析化学教研室, 南京 210009)

摘要 目的:建立准确预测反相高效液相色谱线性梯度条件下溶质保留时间的数学模型。**方法:**应用最小二乘法构建了溶质保留与梯度参数的多元线性回归模型,并采用了两种不同品牌色谱柱及两种中药提取液对该模型进行了验证。**结果:**两种不同色谱柱的实验数据均能很好拟合预测模型,对粉葛提取液中14个组分和金银花提取液中13个组分保留时间预测总的平均相对偏差分别为 $(1.64 \pm 1.25)\%$ ($n=84$), $(1.28 \pm 0.95)\%$ ($n=78$),明显低于两种常用梯度保留公式的预测偏差。**结论:**所建立的回归模型计算简单且对复杂体系中弱保留及强保留组分均能准确地预测。

关键词 梯度洗脱;保留预测;多元线性回归模型

中图分类号 Q141 文献标识码 A 文章编号 1000-5048(2007)03-0236-07

Predictive model of the retention time for solutes based on multiple regression under linear gradient condition

YIN Xiao-ying, YAO Wei-feng, HU Yu-zhu

Department of Analytical Chemistry, China Pharmaceutical University, Nanjing 210009, China

Abstract Aim:To build a model predicting accurately the retention time of components under linear gradient elution in RP-HPLC. **Methods:**The best fitting equation between the retention time and the three parameters was explored by a least squares regression. The regression model is validated by two columns from different manufacturers and two herbal medicine extracts. **Results:**The results on the two different columns showed solid goodness-of-fit measures and the average relative deviation of the prediction is $(1.64 \pm 1.25)\%$ ($n=84$) and $(1.28 \pm 0.95)\%$ ($n=78$) respectively for 14 components of *Pueraria thomsonii* extracts and 13 components of *Flos Lonicerae* extracts using the regression model. The relative deviations are obviously lower than from two known equations. **Conclusion:**The proposed model is simple and especially suited to the retention prediction of complex mixtures.

Key words gradient elution; retention prediction; multiple regression model

梯度分离条件的优化一直是液相色谱领域研究的热点和难点,其中溶质保留时间的准确预测是优化的关键。对于线性梯度保留时间的预测,国内外的色谱工作者做了大量研究工作,卢佩章等^[1]根据色谱保留理论提出了线性梯度保留时间预测公式:

$$t_R = t_0 - \frac{\ln(1 - c \cdot \Delta c \cdot t_0 \cdot \exp(a + c \cdot c_0))}{c \cdot \Delta c} \quad (1)$$

式中 c_0 为起始浓度, Δc 为浓度变化率, t_0 是色谱柱的死时间, a, c 为保留值系数。

国外最常用的有 Snyder 等^[2]提出的 LSS(linear

solvent strength)模型:

$$t_R = (t_0/b) \log(2.3k_0 b + 1) + t_0 + t_d \quad (2)$$

式中, $b = t_0 \Delta \varphi S / t_G$, $\Delta \varphi$ 是有机相溶剂体积分数的改变量, t_d 是系统滞留时间, t_G 是洗脱时间, S, k_0 是由式(3)定义的溶质组分的保留常数。

$$\ln k = \ln k_0 - S \varphi \quad (3)$$

其中 k_0 是组分在假设水相中的容量因子, φ 是有机溶剂的体积分数, S 是 $\ln k$ 和 φ 线性关系的斜率。这里实际假设了 $\ln k$ 和 φ 仅存在一次线性关系。

式(1)是假设无系统延迟条件下推导的,在实

际梯度操作过程中,系统滞留难以扣除,故会对弱保留组分的保留产生影响,导致弱保留组分的保留值预测偏差大;这一问题在式(2)中同样存在,故上述两式适合于强保留组分的保留预测。本文建立了一种不需考虑系统延迟的影响,对弱保留及强保留组分均能准确预测的模型,即通过数个初始实验构建的描述溶质保留与线性梯度参数关系的多元回归模型。本文采用了两种不同品牌色谱柱对粉葛与金银花提取液中各组分线性梯度保留时间进行了预测,并与常用公式(1)和(2)进行了比较,数据结果表明本模型的平均相对偏差略低于采用公式(1)预测的偏差,明显低于公式(2)的预测偏差,验证了回归模型的有效性。

1 建立多元回归的预测模型

在线性梯度洗脱系统中,影响色谱分离的因素很多,多数参数可以参照文献或通过初始实验确定,而少数参数如起始浓度(c_0),洗脱时间(t_G)及最终浓度(c_i)等决定了单线性梯度的洗脱条件,是影响色谱分离的主要因素。故选择以上3个参数作为多元回归模型的变量。

溶质保留与 c_0 , t_G , c_i 之间可能是线性关系,也可能是曲线关系,为了构建两者之间最佳的拟合模型,可采用最小二乘法先从三元一次项开始拟合,然后再进行曲线拟合。三元线性回归模型可以描述为:

$$\log \frac{t_{R(i)}}{t_{R(s)}} = \alpha_0 + \alpha_1 c_0 + \alpha_2 t_G' + \alpha_3 c_i \quad (4)$$

其中 $t_{R(i)}$ 是第 i 组分的保留时间; $t_{R(s)}$ 是参照峰的保留时间; $\alpha_0 \sim \alpha_3$ 是回归系数; $t_G' = t_G/100$, 采用 t_G' 是为了保持 c_0 , t_G , c_i 的回归系数 $\alpha_1, \alpha_2, \alpha_3$ 的数量级一致,使得3个参数由相等的变化量所引起因变量的改变在同一数量级,便于比较。然后进行三元二次项模型拟合,其可以描述为:

$$\log \frac{t_{R(i)}}{t_{R(s)}} = \beta_0 + \beta_1 c_0 + \beta_2 t_G' + \beta_3 c_i + \beta_4 c_0 t_G' + \beta_5 c_0 c_i + \beta_6 t_G' c_i + \beta_7 c_0^2 + \beta_8 t_G'^2 + \beta_9 c_i^2 \quad (5)$$

其中 $\beta_0 \sim \beta_9$ 是偏回归系数,为了拟合公式(5),至少需要9次实验,本文实验只设计7组初始实验,故需采用逐步回归分析对上式(5)进行变量筛选,选择对因变量贡献最大的自变量进入回归模型。逐步回归操作可以通过 SPSS 11.5 软件完成,选择 F 检

验的概率作为变量引入模型的判据,应用系统默认值 0.05,得到的模型为三元一次模型,与公式(4)一致。至于三元三次项或更高次项拟合可以忽略,因为变量太多已失去统计意义。故确定最优的拟合模型为公式(4)所表达的线性回归模型。

2 实验部分

2.1 仪器、试剂和材料

岛津 LC-10ADvp 双泵, SPD-M10AvpDAD 检测器, Class-vp 工作站; 甲醇、乙腈(色谱纯, 汉邦科技有限公司); 粉葛药材, 产地安徽亳州, 经中国药科大学刘丽芳副教授鉴定为 *Pueraria thomsonii*; 金银花药材, 产地山东平邑, 经中国药科大学刘丽芳副教授鉴定为 *Lonicera japonica* Thunb; 其他试剂均为市售分析纯。

2.2 药材前处理

粉葛提取液: 将粉葛药材研成细粉, 称取 0.8 g 加入 30% 甲醇 50 mL, 超声提取 30 min 后, 冷却至室温, 取上清液过 0.45 μm 微孔滤膜, 收集滤液作为供试品溶液。

金银花提取液: 将金银花药材干燥后粉碎, 称取适量加入 20% 甲醇至终浓度为 5 mg/mL, 超声 40 min, 冷却, 取上清液过 0.45 μm 微孔滤膜, 收集滤液作为供试品溶液。

2.3 色谱条件

色谱柱 1: Hypersil C₁₈ (250 mm \times 4.6 mm), 色谱柱 2: Kromasil C₁₈ (250 mm \times 4.6 mm), 流速: 1 mL/min; 粉葛提取液: 以甲醇与 0.1% 醋酸梯度洗脱, 检测波长为 250 nm; 金银花提取液: 以乙腈与 0.4% 磷酸梯度洗脱, 检测波长为 237 nm; 死时间采用亚硝酸钠测定, 色谱柱 1、色谱柱 2 分别为 2.133 和 2.688 min; 在系统不接色谱柱的条件下, 以甲醇为流动相, 0.3% 的丙酮为示踪剂进行实验^[3], 测得系统滞留时间为 2.740 min。

2.4 实验设计

为了验证式(4)多元回归模型预测溶质保留的有效性, 采用粉葛和金银花提取液、两种不同品牌的 ODS 色谱柱从构建模型的统计评价和模型的稳健性两方面进行了考察。其中多元回归模型的构建采用均匀设计方法来安排初始实验。设计 7 组初始实验, 各参数设定 7 水平, 选取均匀表 $U_7(7^3)$ ^[4] 安排实验。通过运行 1 次预实验将有机溶

剂从 0~100% 变化将全部组分洗脱出来, 粉葛提取液 c_0 , t_G , c_i 的取值范围为: 色谱柱 1: c_0 : 16%~22%、 t_G : 34 min~58 min、 c_i : 48%~72%, 色谱柱 2: c_0 : 16%~28%、 t_G : 34 min~58 min、 c_i : 52%~76%; 金银花提取液 c_0 , t_G , c_i 的取值范围为: 色谱柱 1: 16%~18%、 t_G : 34 min~58 min、 c_i : 56%~80%, 色谱柱 2: c_0 : 10%~22%、 t_G : 34 min~58 min、 c_i : 36%~72%; 在上述范围内将 c_0 、 t_G 、 c_i 分别均匀分成 7 个水平, 按表 $U_7(7^3)$ 进行初始实验。

为了考察回归模型的稳健性, 对不同提取液和不同色谱柱在其优化范围内分别另选取 3 组梯度条件进行了实验。粉葛提取液: 色谱柱 1 的 3 组梯度条件 c_0 , t_G 和 c_i 分别为: 17%、42 min、64%; 19%、50 min、58% 和 22%、46 min、56%, 色谱柱 2 的 3 组梯度

条件 c_0 , t_G 和 c_i 分别为: 17%、44 min、70%; 23%、40 min、66%、25% 和 45 min、65%; 金银花提取液: 色谱柱 1 的梯度条件 c_0 , t_G 和 c_i 分别为: 13%、44 min、70%; 12%、50 min、64% 和 15%、54 min、56%, 色谱柱 2 的梯度条件 c_0 , t_G 和 c_i 分别为: 18%、52 min、65%; 21%、58 min、72% 和 18%、52 min、53%。

3 结果与讨论

3.1 粉葛提取液中 14 个组分梯度洗脱条件下保留时间的预测

3.1.1 多元回归模型的构建及统计评价 根据 7 组初始实验数据, 按式(4)采用最小二乘法对各组分进行线性回归分析, 得到不同色谱柱各溶质的保留预测模型系数及统计结果, 见表 1。

Tab.1 Multiple regression equation coefficients and goodness of fit assessment from the initial data sets of *Pueraria thomsonii* extracts ($n = 7$, 95% confidence interval)

| Column | Peak No. | Coefficients | | | | Statistical parameters | | | |
|--------|----------|--------------|------------|------------|------------|------------------------|---------|-------|---------|
| | | α_0 | α_1 | α_2 | α_3 | R | SE | F | P |
| 1 | 1 | -1.028 7 | 1.391 3 | -0.307 6 | 0.396 7 | 0.992 5 | 0.009 7 | 65.93 | 0.003 1 |
| | 2 | -0.599 3 | 0.072 3 | -0.343 7 | 0.350 0 | 0.987 3 | 0.010 6 | 38.54 | 0.006 8 |
| | 3 | -0.693 3 | 1.796 7 | -0.339 9 | 0.391 0 | 0.993 3 | 0.010 2 | 73.50 | 0.002 6 |
| | 4 | -0.255 3 | 0.225 5 | -0.244 1 | 0.243 3 | 0.991 1 | 0.006 4 | 55.35 | 0.004 0 |
| | 5 | -0.001 7 | -0.830 8 | -0.078 3 | 0.116 4 | 0.990 7 | 0.004 2 | 52.83 | 0.004 3 |
| | 6* | 1.728 3 | -2.187 9 | 0.307 0 | -0.376 8 | 0.986 4 | 0.015 3 | 36.02 | 0.007 5 |
| | 7 | 0.034 2 | 0.339 5 | 0.087 7 | -0.117 2 | 0.989 9 | 0.003 2 | 48.59 | 0.004 8 |
| | 8 | 0.018 9 | 0.528 2 | 0.070 6 | -0.095 7 | 0.987 4 | 0.003 6 | 38.86 | 0.006 7 |
| | 9 | 0.027 7 | 0.596 2 | 0.081 3 | -0.114 8 | 0.990 1 | 0.003 6 | 49.63 | 0.004 7 |
| | 10 | 0.051 5 | 0.789 6 | 0.139 9 | -0.197 2 | 0.988 8 | 0.006 0 | 43.81 | 0.005 6 |
| | 11 | 0.052 7 | 1.018 6 | 0.143 6 | -0.200 0 | 0.988 7 | 0.006 7 | 43.63 | 0.005 7 |
| | 12 | 0.051 8 | 1.042 9 | 0.142 0 | -0.197 3 | 0.989 2 | 0.006 6 | 45.42 | 0.005 3 |
| | 13 | 0.122 5 | 1.659 9 | 0.237 1 | -0.334 3 | 0.986 2 | 0.012 2 | 35.61 | 0.007 6 |
| | 14 | 0.066 5 | 1.715 2 | 0.303 5 | -0.226 8 | 0.990 7 | 0.004 2 | 52.83 | 0.004 3 |
| 2 | 1 | -0.650 0 | 0.615 8 | -0.346 7 | 0.303 9 | 0.960 1 | 0.018 4 | 11.78 | 0.036 3 |
| | 2 | -0.078 9 | -1.213 3 | -0.244 0 | 0.225 8 | 0.976 5 | 0.017 9 | 20.52 | 0.016 8 |
| | 3 | -0.009 8 | -1.166 7 | -0.141 1 | 0.171 8 | 0.996 9 | 0.005 8 | 162.8 | 0.000 8 |
| | 4 | -0.133 5 | -0.182 9 | -0.138 2 | 0.148 9 | 0.916 8 | 0.010 1 | 5.268 | 0.090 3 |
| | 5 | -0.274 4 | 0.833 7 | -0.164 0 | 0.151 7 | 0.962 4 | 0.011 9 | 8.370 | 0.108 5 |
| | 6* | 1.726 0 | -2.482 7 | 0.333 1 | -0.348 3 | 0.999 6 | 0.004 8 | 117.8 | 0.000 0 |
| | 7 | 0.050 9 | 0.114 3 | 0.088 6 | -0.095 5 | 0.952 9 | 0.004 7 | 9.871 | 0.000 0 |
| | 8 | 0.022 4 | 0.510 1 | 0.065 7 | -0.094 8 | 0.999 5 | 0.001 1 | 921.9 | 0.000 0 |
| | 9 | 0.034 5 | 0.693 1 | 0.084 2 | -0.135 6 | 0.999 7 | 0.001 1 | 1 702 | 0.000 0 |
| | 10 | 0.066 0 | 0.721 9 | 0.137 5 | -0.199 9 | 0.998 4 | 0.002 8 | 303.9 | 0.000 3 |
| | 11 | 0.060 1 | 1.091 2 | 0.135 6 | -0.216 7 | 0.999 7 | 0.001 8 | 1 573 | 0.000 0 |
| | 12 | 0.061 4 | 1.142 5 | 0.134 6 | -0.220 5 | 0.999 6 | 0.002 0 | 1 335 | 0.000 0 |
| | 13 | 0.141 6 | 1.942 7 | 0.233 7 | -0.388 6 | 0.999 6 | 0.003 8 | 1 113 | 0.000 0 |
| | 14 | 0.165 9 | 2.178 6 | 0.277 1 | -0.443 4 | 0.999 3 | 0.006 0 | 470.3 | 0.002 1 |

* The 6th peak was identified as puerarin by comparison with standard puerarin and taken as the reference peak

表 1 分别列出了色谱柱 1, 2 的多元预测模型

的回归系数和统计结果。从表 1 中的两组回归系

数均显示: $|\alpha_1|$ 远远大于 $|\alpha_2|$ 和 $|\alpha_3|$, 这表明参数 c_0 对回归模型的贡献最大, 因此 c_0 是 3 个参数中最主要的影响因素。统计结果显示, 两个色谱柱 14 个峰的复相关系数 R 均接近 1, 标准误差 SE 很小, F 统计量绝对大多数大于临界值 $F_{0.05} = 9.28$, 且 14 个峰多数具有高度显著水平的 P 值 ($P < 0.01$), 这些结果表明 $\log t_{R(i)} / t_{R(s)}$ 与 3 个梯度参数之间呈显著的线性相关, 模型拟合良好。但在色谱柱 2 中第 4、5 组分的预测模型显示无统计意义, 可能是因为有些组分在不同洗脱条件下易出现峰交叉或峰重叠从而导致拟合质量下降。统计结果表明, 两不同品牌色谱柱均能很好地拟合预测模型。

3.1.2 多元回归模型稳健性的考察和比较 对于不同色谱柱分别选取了 3 组梯度条件对粉葛提取液进行实验。预测的准确性采用预测值与实测值的相对偏差来评价。表 2 比较了 3 种保留预测模

型的预测结果, 因色谱柱 1、2 对于 14 个组分中同一组分的预测偏差相近, 故将每个组分的 6 组数据合并以一组评价值来表示。

从表 2 可以看出, 14 个组分的回归模型的平均相对偏差为 $(1.64 \pm 1.25)\%$, 略低于公式 (1) 的预测结果 $(2.14 \pm 2.04)\%$, 远远低于公式 (2) 的结果 $(10.93 \pm 8.70)\%$ 。公式 (1)(2) 的预测结果呈现一定的规律性: 梯度洗脱初期, 弱保留溶质预测的相对偏差较大, 随着组分保留的增大, 预测偏差逐步减小, 结果表明系统滞留对弱保留组分的影响大, 导致其预测偏差大, 随着保留组分的极性的减弱, 这种影响逐步减小, 预测值趋于接近实测值, 故式 (1) 和 (2) 适合强保留组分的预测。而本文建立的回归模型则不受其影响, 对弱保留及强保留组分均能很好的预测。

Tab.2 Comparison of the retention time and the relative deviation between the predicted and experimental data from three gradient conditions of *Pueraria thomsonii* extracts

| Peak No. | Retention time(min) | | | | Average relative deviation (% , $n=6$) | | |
|------------------------|---------------------|-------|-------|-------|---|------------------|-----------------|
| | Exp. | Eq.1 | Eq.2 | Eq.4 | Eq.1 | Eq.2 | Eq.4 |
| 1 | 3.41 | 3.45 | 3.44* | 3.43 | 1.15 | 7.40 | 1.35 |
| 2 | 4.29 | 4.35 | 4.33* | 4.28 | 2.59 | 13.91 | 0.95 |
| 3 | 8.61 | 8.87 | 8.76* | 8.75 | 2.83 | 14.41 | 2.60 |
| 4 | 9.85 | 10.26 | 11.33 | 9.90 | 2.85 | 13.05 | 0.97 |
| 5 | 10.10 | 11.14 | 13.19 | 10.52 | 3.80 | 12.84 | 1.83 |
| 6 | 14.53 | 15.60 | 17.91 | 15.04 | 4.14 | 14.66 | 2.46 |
| 7 | 17.54 | 18.61 | 21.36 | 18.23 | 2.60 | 13.28 | 2.56 |
| 8 | 18.90 | 19.95 | 22.58 | 19.55 | 2.28 | 11.70 | 2.02 |
| 9 | 19.71 | 20.75 | 23.44 | 20.38 | 2.06 | 11.16 | 1.84 |
| 10 | 21.87 | 22.90 | 25.90 | 22.72 | 1.75 | 11.07 | 1.95 |
| 11 | 24.73 | 25.74 | 28.73 | 25.59 | 1.41 | 9.67 | 1.49 |
| 12 | 25.04 | 26.04 | 29.02 | 25.90 | 1.37 | 9.42 | 1.44 |
| 13 | 37.56 | 38.37 | 41.79 | 38.61 | 0.66 | 6.63 | 1.02 |
| 14 | 43.82 | 44.23 | 47.97 | 43.02 | 0.54 | 5.60 | 1.16 |
| $\bar{x} \pm s (n=84)$ | | | | | 2.14 ± 2.04 | 10.93 ± 8.70 | 1.64 ± 1.25 |

* The retention time of solutes is calculated by eq. (3) because the retention of solutes is too weak to be calculated accurately by eq. (2). Exp. indicates the actual retention time (with experiment condition: c_0 , t_G and c_i is 22%, 46min and 56%, respectively.)

3.2 金银花提取液中 13 个组分梯度洗脱条件下保留时间的预测

为了进一步验证提出的回归模型预测的有效性, 本研究另选用金银花提取液为例, 采用前述方法进行实验。

3.2.1 多元回归模型的构建及统计评价 按公式 (4), 根据 7 组初始实验数据采用最小二乘法进行线性回归拟合, 得到各模型系数及统计结果, 见表 3。

表 3 列出了色谱柱 1、2 模型拟合的结果。表 3

中的回归系数显示出: c_0 项的系数 $|\alpha_1|$ 远远大于其他两项的系数 $|\alpha_2|$ 和 $|\alpha_3|$, 表明参数对回归模型的贡献最大, 是 3 参数中最主要的影响因素。统计结果表明 $\log t_{R(i)} / t_{R(s)}$, 两个色谱柱 13 个模型的复相关系数 R 均接近 1, 标准误差 SE 亦很小, F 值均远大于临界值, 且 13 个模型的 P 值均具有高度显著水平 ($P < 0.01$), 这些结果进一步表明与 3 个梯度参数之间呈显著的线性关系, 模型拟合亦良好, 且两种不同品牌色谱柱所得结果十分接近。

Tab.3 Multiple regression equation coefficients and goodness of fit assessment from the initial data sets for the samples of *Flos Lonicerae* ($n = 7$, 95% confidence interval)

| Column | Peak No. | Coefficients | | | | Statistical parameters | | | |
|--------|----------|--------------|------------|------------|------------|------------------------|---------|-------|---------|
| | | α_0 | α_1 | α_2 | α_3 | R | SE | F | P |
| 1 | 1 | -0.365 1 | -2.402 6 | -0.597 2 | 0.462 8 | 0.993 5 | 0.018 3 | 76.42 | 0.002 5 |
| | 2 | -0.298 5 | -1.798 2 | -0.514 3 | 0.443 0 | 0.998 5 | 0.007 1 | 343.3 | 0.000 3 |
| | 3 | -0.181 0 | -1.105 4 | -0.330 4 | 0.301 8 | 0.996 4 | 0.007 1 | 137.8 | 0.001 0 |
| | 4 | -0.160 4 | -1.024 8 | -0.311 1 | 0.280 5 | 0.996 3 | 0.006 7 | 135.9 | 0.001 1 |
| | 5 | -0.147 5 | -0.909 1 | -0.282 2 | 0.258 2 | 0.996 1 | 0.006 2 | 127.2 | 0.001 2 |
| | 6 | -0.119 4 | -0.742 8 | -0.258 3 | 0.227 9 | 0.996 5 | 0.004 9 | 142.9 | 0.001 0 |
| | 7 | -0.067 2 | -0.425 6 | -0.167 6 | 0.145 7 | 0.996 2 | 0.003 1 | 132.3 | 0.001 1 |
| | 8 | -0.054 8 | -0.350 8 | -0.148 7 | 0.129 4 | 0.996 5 | 0.002 6 | 142.3 | 0.001 0 |
| | 9 | -0.021 3 | -0.130 1 | -0.038 6 | 0.031 9 | 0.991 4 | 0.001 3 | 57.68 | 0.003 8 |
| | 10* | 1.557 6 | -0.727 8 | 0.626 8 | -0.455 0 | 0.997 8 | 0.007 5 | 227.5 | 0.000 5 |
| | 11 | 0.018 6 | 0.071 3 | -0.009 3 | 0.007 4 | 0.992 9 | 0.000 6 | 70.15 | 0.002 8 |
| | 12 | 0.043 6 | 0.133 8 | 0.014 5 | -0.027 3 | 0.976 5 | 0.002 1 | 20.57 | 0.016 7 |
| | 13 | 0.038 9 | 0.168 5 | 0.038 4 | -0.040 1 | 0.990 3 | 0.001 7 | 50.92 | 0.004 5 |
| 2 | 1 | -0.456 9 | -1.798 9 | -0.563 6 | 0.574 7 | 0.964 5 | 0.037 0 | 22.22 | 0.002 6 |
| | 2 | -0.288 0 | -1.880 5 | -0.546 9 | 0.587 5 | 0.991 9 | 0.019 0 | 101.9 | 0.000 1 |
| | 3 | -0.142 2 | -1.556 5 | -0.463 4 | 0.490 8 | 0.995 9 | 0.012 9 | 200.1 | 0.000 0 |
| | 4 | -0.097 2 | -1.234 9 | -0.384 4 | 0.404 8 | 0.995 0 | 0.013 1 | 166.3 | 0.000 0 |
| | 5 | -0.087 7 | -1.096 7 | -0.345 8 | 0.367 0 | 0.994 4 | 0.013 5 | 147.5 | 0.000 0 |
| | 6 | -0.073 1 | -0.792 0 | -0.293 5 | 0.310 4 | 0.994 4 | 0.012 5 | 147.5 | 0.000 0 |
| | 7 | -0.036 4 | -0.314 8 | -0.141 6 | 0.154 7 | 0.991 3 | 0.015 3 | 95.03 | 0.000 1 |
| | 8 | -0.031 3 | -0.207 4 | -0.116 7 | 0.129 3 | 0.991 4 | 0.015 2 | 95.13 | 0.000 1 |
| | 9 | -0.001 4 | -0.096 9 | -0.020 6 | 0.020 1 | 0.986 6 | 0.021 1 | 60.80 | 0.000 2 |
| | 10* | 1.725 6 | -1.265 4 | 0.553 4 | -0.626 4 | 0.986 9 | 0.020 9 | 62.56 | 0.000 2 |
| | 11 | 0.017 4 | 0.061 3 | -0.008 3 | 0.007 8 | 0.999 0 | 0.004 5 | 514.9 | 0.000 1 |
| | 12 | 0.005 0 | 0.306 4 | -0.026 9 | 0.041 3 | 0.997 6 | 0.007 4 | 207.5 | 0.000 6 |
| | 13 | 0.015 3 | 0.412 1 | -0.004 6 | 0.017 3 | 0.997 1 | 0.008 0 | 170.9 | 0.000 8 |

* The 10th peak was taken as the reference peak

3.2.2 多元回归模型稳健性的考察和比较 为了进一步验证模型的稳健性,对于不同色谱柱同样分

别另选了 3 组梯度条件对金银花提取液进行实验。3 种模型的预测结果见表 4。

Tab.4 Comparison of the retention time and the relative deviation between the predicted and experimental data from three gradient conditions of *Flos Lonicerae* extract

| Peak No. | Retention time(min) | | | | Average relative deviation (% , $n = 6$) | | |
|--------------------------|---------------------|-------|-------|-------|---|---------------|-------------|
| | Exp. | Eq.1 | Eq.2 | Eq.4 | Eq.1 | Eq.2 | Eq.4 |
| 1 | 5.47 | 5.94 | 6.99 | 5.54 | 2.73 | 22.30 | 1.61 |
| 2 | 8.62 | 9.52 | 10.15 | 8.59 | 3.47 | 20.03 | 1.28 |
| 3 | 15.22 | 16.25 | 16.38 | 14.99 | 2.51 | 16.93 | 1.14 |
| 4 | 16.31 | 17.40 | 17.48 | 16.11 | 2.56 | 15.35 | 1.11 |
| 5 | 17.69 | 18.62 | 18.71 | 17.39 | 2.34 | 14.69 | 1.30 |
| 6 | 19.77 | 20.61 | 20.71 | 19.47 | 2.02 | 12.97 | 1.13 |
| 7 | 25.11 | 25.44 | 25.77 | 24.66 | 1.51 | 11.06 | 1.01 |
| 8 | 26.61 | 26.81 | 27.19 | 26.11 | 1.41 | 10.62 | 1.39 |
| 9 | 31.33 | 31.14 | 31.80 | 30.78 | 1.20 | 10.33 | 1.72 |
| 10* | 34.71 | 34.35 | 35.06 | 34.04 | 1.28 | 9.91 | 2.20 |
| 11 | 37.13 | 36.66 | 37.36 | 36.35 | 1.17 | 6.26 | 0.90 |
| 12 | 39.48 | 39.04 | 39.48 | 38.75 | 0.68 | 5.12 | 0.90 |
| 13 | 40.19 | 39.58 | 40.34 | 39.3 | 0.74 | 4.87 | 0.93 |
| $\bar{x} \pm s (n = 78)$ | | | | | 1.82 ± 1.94 | 12.34 ± 12.43 | 1.28 ± 0.95 |

Chromatographic condition: c_0 , t_G and c_i is 10%, 58 min and 46%, respectively by column 1

表4结果表明,13个组分的回归模型的平均相对偏差为 $(1.28 \pm 0.95)\%$ ($n=78$),同样低于公式(1) $(1.82 \pm 1.94)\%$ ($n=78$)和(2) $(12.34 \pm 12.43)\%$ ($n=78$)的预测结果。采用公式(1)和(2)的预测结果同样显示出弱保留组分预测偏差较大,但随着组分保留的增大,预测更趋于准确。公式(4)则不受系统滞留的影响,能准确地预测强保留及弱保留两种组分的保留时间。

表2、4均显示公式(2)预测的平均相对偏差较大,最大平均相对偏差可达到22.3%,这种偏差可能源于两方面,一是实际梯度系统中,溶质的 $\ln k$ 和洗脱浓度 φ 之间的非线性关系;二是实际梯度系统的非理想化^[2]。非理想化主要体现在3方面的影响:系统滞留、仪器误差及柱内因素^[5]。因此,应用公式(2)准确预测溶质保留还受到许多条件的限制。

3.3 粉葛及金银花提取液的最佳优化条件

分别以“3.1.1”项和“3.2.1”项下各组分的多元回归模型作为保留时间预测模型,校正标准分离度积(r^*)^[6]选为优化指标,网格搜索作为寻优算法, q_0 、 t_G 和 c_i 的搜索步长分别为0.5%、1 min、1%,在3参数的范围内搜索 r^* 最大值时所对应的参数值,即最优分离条件。这样可以得到粉葛提取液的最佳分离条件:色谱柱1: q_0 、 t_G 和 c_i 分别为:22%、46 min、68%见图1(a),色谱柱2: q_0 、 t_G 和 c_i 分别为:25%、48 min、62%见图1(b);金银花提取液的最佳分离条件:色谱柱1: q_0 、 t_G 和 c_i 分别为:16%、56 min、58%,见图2(b),色谱柱2: q_0 、 t_G 和 c_i 分别为:18%、55 min、66%见图2(a)。从图1可以看出14个组分均能在适当的分析时间内得到基线分离,且色谱柱2的分离优于色谱柱1,从难分离物质对11号峰与12号峰的分

图2(c)为色谱柱2初始条件的色谱图,1号峰前有两个色谱峰,实为多个色谱峰的重叠峰,经优化后分离为5个峰见图2(a),图2(a)为优化后的色谱图,表明经优化后可以分离更多的组分且分析时间为缩短,因1号峰前5个峰靠近溶剂峰数据不易准确测定,故未进行分析。图2(b)为色谱柱1最优分离色谱图,同样可以看出色谱柱2的分离效果优于色谱柱1。

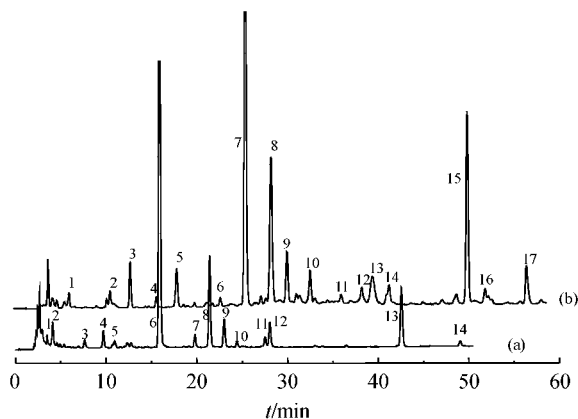


Fig.1 Optimal chromatograms of *Pueraria thomsonii* extracts under the gradient condition by two columns: q_0 , t_G and c_i is 22%, 46 min and 68%, respectively in column 1 (a); q_0 , t_G and c_i is 25%, 48 min and 62%, respectively in column 2 (b)

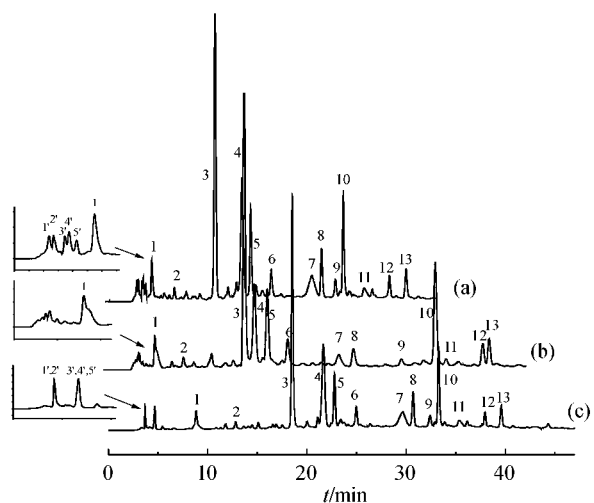


Fig.2 Optimal chromatograms of *Flos Lonicerae* extract under the gradient condition by two columns: q_0 , t_G and c_i is 16%, 56 min and 58%, respectively in column 1 (b); q_0 , t_G and c_i is 18%, 55 min and 66%, respectively in column 2 (a); and initial experimental condition: q_0 , t_G and c_i is 10%, 50 min and 54%, respectively in column 2 (c)

4 结论

本文建立的基于多元回归分析的预测模型能准确地预测线性梯度条件下溶质的保留时间,预测准确度优于两个常用预测公式,且不需考虑系统或柱内因素的影响、计算简单,适合应用于复杂体系分离条件优化中诸多组分保留值的准确预测。

参考文献

- [1] 卢佩章(Lu PZ), 张玉奎(Zhang YK), 梁鑫淼(Liang XM). 高效液相色谱法及其专家系统[M]. 沈阳: 辽宁科技出版社, 1994: 322 - 327.
- [2] Snyder LR, Dolan JW, Gant JR. Gradient elution in high-performance liquid chromatography: I. Theoretical basis for reversed-phase systems[J]. *J Chromatogr*, 1979, **165**(1): 3 - 30.
- [3] 李瑞江(Li RJ), 张维冰(Zhang WB), 倪坚毅(Ni JY), 等. 液相色谱梯度淋洗过程中溶剂的分布规律[J]. 高等学校化学学报(*Chem J Chin Univ*), 1998, **19**(10): 1 556 - 1 559.
- [4] 张学中(Zhang XZ). 关于均匀设计表的应用[C]. 均匀设计论文集(第一集), 1995: 116 - 120.
- [5] Quarry MA, Grob RL, Snyder LR. Measurement and use of retention data from high-performance gradient elution: Correction for "non-ideal" processes originating within the column[J]. *J Chromatogr A*, 1984, **285**(1): 19 - 51.
- [6] Vanbel PF, Tilquin BL, Schoenmakers PJ. Criteria for optimizing the separation of target analytes in complex chromatograms[J]. *Chemom Intell Lab Syst*, 1996, **35**(1): 67 - 86.

近年来, 研究人员开发了各种呼吸系统药物新制剂、新剂型, 如口腔速崩片、缓(控)释制剂、口腔吸入剂和鼻腔吸入剂等。预计在未来几年中, 呼吸系统药物市场将持续稳步增长, 至 2010 年全球哮喘/慢性阻塞性肺疾病(COPD)治疗药物的销售额可能达到 230 亿美元, 其中肾上腺皮质激素类药物/长效支气管扩张药吸入制剂有望名列前茅。研究开发的新剂型如下:

口腔速崩片

- 盐酸氨溴素口腔速崩片
- 复方柠檬酸苯海拉明 - 盐酸伪麻黄碱口腔速崩片
- 复方氯苯那敏 - 右美沙酚 - 伪麻黄碱口腔速崩片
- 复方感冒口腔速崩片

缓(控)释制剂

- 复方地洛他定 - 硫酸伪麻黄碱 12 小时缓释片
- 复方地洛他定 - 硫酸伪麻黄碱 24 小时控释片
- 复方愈创甘油醚 - 溴氢酸右美沙酚控释片
- 复方氯雷他定 - 硫酸伪麻黄碱 24 小时控释片
- 复方盐酸非索非那定 - 盐酸伪麻黄碱 24 小时控释片

口腔吸入剂

- 布地奈德新氢氟烷抛射剂揿压式定量吸入剂
- 糠酸莫米松粉雾剂
- 环索奈德吸入剂
- 丙酸氟替卡松氢氟烷吸入气雾剂
- 溴化异丙托溴铵氢氟烷气雾剂
- 丙卡特罗粉雾剂
- 酒石酸 arformotero 吸入液;
- 复方丙酸氟替卡松 - 沙美特罗粉雾剂
- 复方福莫特罗 - 二丙酸倍氯米松氢氟烷吸入剂

鼻腔吸入剂

- 曲安缩松鼻用氢氟烷气雾剂
- 小儿用丙酸氟替卡松鼻腔喷雾剂
- 环索奈德鼻腔喷雾剂

(摘自世界临床药物)