

卷烟辅材参数与有害成分释放量的多元模型的建立、传递和验证

聂聪¹ 谢复炜¹ 赵乐¹ 彭斌¹ 杨俊² 张悠金² 邵利民^{* 2}

¹(中国烟草总公司郑州烟草研究院, 郑州 450001) ²(中国科学技术大学化学与材料科学学院, 合肥 230026)

摘要 研究了卷烟辅材参数与有害成分释放量的关系。这些有害成分包括焦油、烟碱、CO、HCN、苯并[a]芘(B[a]P)、4-(甲基亚硝胺基)-1-(3-吡啶基)-1-丁酮(NNK)、巴豆醛、苯酚和 NH₃。使用多元校正分析, 模型传递等化学计量学方法, 对实验数据进行分析, 建立了多元线性和多项式模型, 通过独立验证样品和 *t* 检验确立了每种有害成分的最优模型; 分析了各项辅材参数对有害成分释放量的影响权重。

关键词 多元校正; 模型传递; 烟气; 有害成分

1 引言

卷烟烟气非常复杂, 含有许多对人体有害的成分。了解这些有害成分及其在主流烟气中的释放量是控烟运动的重要内容。20 世纪 60 年代以来, 卷烟烟气中的有害成分被逐渐列出, 最著名的是 Hoffmann 名单^[1]。1998 年, 加拿大政府根据此名单, 通过立法要求卷烟生产商定期检测卷烟主流烟气中 13 类 46 种有害成分的含量并公布其结果。

降低有害成分释放量的有效手段之一是调控卷烟辅材参数^[2~8]。为了达到此目的, 需要明确这些参数与主流烟气中有害成分释放量的关系, 即建立多变量数学模型, 通过模型明确有害成分释放量的主要影响因素, 从而为低危害卷烟辅材参数的优化选择提供支持。谢剑平等^[9]通过研究确定了苯并[a]芘(B[a]P), 苯酚, CO, 4-(甲基亚硝胺基)-1-(3-吡啶基)-1-丁酮(NNK), NH₃, HCN 和巴豆醛作为卷烟主流烟气的代表性有害成分, 发展了一种危害性评价指数, 用于对低危害卷烟产品进行量化风险度评价。

本研究运用化学计量学法考察了卷烟辅材参数与烟气有害成分释放量的定量关系, 建立了多元线性模型以及多项式模型; 通过模型传递及统计检验手段, 对建立的模型进行验证, 获得了关于每种成分的最佳定量模型。研究了各参数对有害成分释放量的影响权重。本研究涉及的辅材参数包括卷烟纸克重、卷烟纸透气度、成型纸透气度、接装纸透气度和滤棒吸阻, 这种情况等价于五变量数学模型; 还可以使用嘴嘴通风度替代成型纸透气度和接装纸透气度, 即研究四变量模型。

2 实验及数据处理

本研究需要的辅材以及相应参数如下。卷烟纸自然透气度(由同一种纸浆加工, 助燃剂相同): 20, 40, 50, 60 和 80 CU。卷烟纸克重: 26, 28, 30, 32 和 34 g·m²。接装纸透气度: 0, 100, 400, 800 和 1200 CU。(预打孔方式: 固定中心打孔位置 19 mm, 激光打孔, 长度固定 32 mm)。成型纸透气度: 3300, 4500, 6000, 10000 和 15000 CU。吸阻(144 mm 滤棒, 丝束 3.0 Y/35000): 3600, 4000, 4400, 4800 和 5200 Pa。叶组配方: 5 个单等级烟叶混配, 不加入香精香料, 只添加 8% 保润剂。

使用上述辅材, 通过实验设计, 制备了 90 个样品。卷烟烟气中的焦油、烟碱、CO、HCN、B[a]P、NNK、巴豆醛、苯酚采用行业标准方法进行测定; NH₃ 采用加拿大卫生部推荐分析方法进行测定^[10]。这些测量数据和辅材参数构成了校正集(Calibration set), 用于建立多变量模型。

用于模型验证的样品分别由 6 家卷烟生产企业制备。使用上述叶组, 任意变化一个或多个辅材参数, 制备验证样品。各工厂至少制备 3 个独立样品, 共制备 23 个验证样。对于每个验证样, 均采用与校正

集样品同样的分析方法测定其有害成分释放量。这些数据与验证样品辅材参数构成了模型的验证集 (Validation set), 数据计算采用 MATLAB 7.1.0 (The MathWorks Inc., Natick MA), 在 Windows XP 操作系统。

3 结果与讨论

3.1 模型的建立

烟气成分释放量和辅材参数之间并没有确定的理论关系, 所以在建立模型时需要试验不同的函数形式。研究了 4 种情况, 分别是五参数线性和多项式模型, 线性和四参数多项式模型。建模前, 自变量均通过减去均值的中心化处理。模型质量通过残差曲线和校正均方根误差 (Root mean square error of calibration, RMSEC) 进行评价。下面以焦油释放量为例, 说明模型建立的情况。

首先研究了焦油释放量的五参数线性模型。根据多元拟合, 获得了以下多元线性函数:

$$y = -1.26 \times 10^{-3} x_1 + 3.54 \times 10^{-2} x_2 - 1.27 \times 10^{-4} x_3 - 5.74 \times 10^{-3} x_4 - 1.25 \times 10^{-3} x_5 + 1.02 \quad (1)$$

其中, 自变量 x_1, x_2, x_3, x_4 和 x_5 分别表示卷烟纸克重、卷烟纸透气度、成型纸透气度、接装纸透气度和滤棒吸阻 5 种参数, 因变量 y 为焦油释放量。该模型的 RMSEC 为 0.613。通过该模型预测了校正集中 90 个样品的焦油释放量, 并计算了与实测值的差, 结果见图 1a。焦油释放量的预测值与实测值吻合得较好; 残差很小, 并且基本在零值附近波动。此外, 该线性模型(式 1)显示出焦油释放量与 x_2, x_3, x_4 和 x_5 参数呈现负相关。此结论与实际情况相符, 因为增大参数 x_2, x_3 和 x_4 的值会增加空气对烟气的稀释; 而增大参数 x_5 的值会降低烟气的吸入量, 结果降低了烟气中的焦油释放量。所以这一线性模型能够有效解释实验数据。又研究了四参数线性模型, 得到了类似的结论, 并且 RMSEC 小于五参数线性模型。

多项式模型的建立相对复杂一些, 因为需要确定所有变量的多项式次数。通过穷举法获得了变量次数各不相同的多个候选多项式模型, 计算各模型的 RMSEC 值, 取值最小者被确定为所需的多项式模型。下面是焦油释放量的四参数多项式模型:

$$y = -9.00 \times 10^{-3} x_1 - 3.03 \times 10^{-4} x_2^2 - 3.40 \times 10^{-2} x_2 - 1.36 \times 10^{-3} x_5 - 3.19 \times 10^{-4} x_6^2 - 0.139 x_6 + 1.02 \quad (2)$$

其中, x_1, x_2, x_5 和 x_6 分别为卷烟纸克重、卷烟纸透气度、滤棒吸阻和滤嘴通风度 4 种参数的数值, 该模型的 RMSEC 为 0.447。通过该模型预测了校正集中 90 个样品的焦油释放量, 并计算了与实测值的差, 结果见图 1b。此多项式模型对校正集样品焦油释放量的预测也很好, 能够有效解释实验数据。从图 1a 和图 1b 的对比以及相应的 RMSEC 数值可以看出, 对于焦油释放量而言, 多项式模型的误差更小。

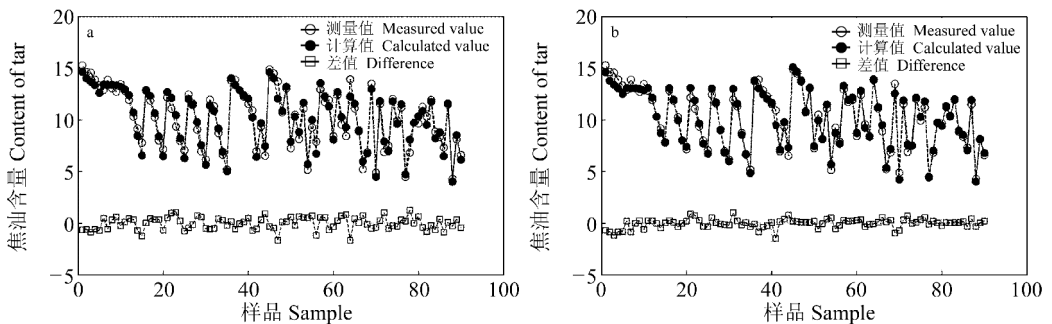


图 1 焦油释放量的建模情况: (a) 为五参数线性模型, (b) 为四参数多项式模型

Fig. 1 Measured and predicted amount of tar in the mainstream smoke. The prediction was done with (a) 5-variable linear and (b) 4-variable polynomial models, respectively

类似对焦油释放量的建模, 对其它 8 种烟气成分进行了研究, 各自都建立了 4 种模型。模型的 RMSEC 和相对误差列于表 1。数据分析结果表明, 除 NNK 外, 其它烟气成分都获得了较好的线性和多项式模型, 模型对烟气成分释放量的预测值与实测值吻合得很好, 而且能够有效表现出释放量在不同样品中的变化情况。对于 NNK, 线性模型或者多项式模型均无法有效地解释实验数据, 如图 2 所示。原因是烤烟型卷烟烟气中 NNK 的含量极低, 接近方法的检出限, 因此实测值误差较大, 导致模型无效。

表 1 数据表明: 多项式模型优于线性模型, 四参数模型优于五参数模型。但此结论仅是基于校正

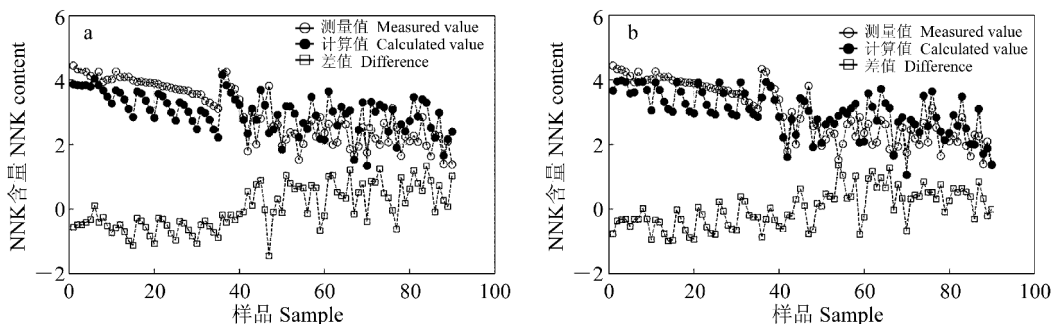


图 2 NNK 释放量的建模情况;图 2a 为五参数线性模型,图 2b 为四参数多项式模型

Fig. 2 Measured and predicted amount of 4-(methylnitrosamino)-1-(3-pyridyl)-butanone (NNK) in the mainstream smoke. The prediction was done with (a) 5-variable linear and (b) 4-variable polynomial models, respectively

表 1 不同模型* 的 RMSEC(a)和相对误差绝对值的平均(b)

Table 1 (a) Root mean square error of calibration (RMSEC) and (b) the mean average percentage errors of the predictions with different models

烟气成分 Component	5L 模型(a/b) Model 5L	5P 模型(a/b) Model 5P	4L 模型(a/b) Model 4L	4P 模型(a/b) Model 4P
烟碱 Nicotine	0.0455/4.78	0.0423/4.44	0.0405/4.19	0.0318/3.14
焦油 Tar	0.613/5.61	0.528/4.59	0.463/3.52	0.447/3.41
苯并[a]芘 Benzo[a]pyrene	0.598/4.21	0.490/3.52	0.579/4.04	0.427/2.88
苯酚 Phenol	0.994/7.39	0.972/7.09	0.988/7.23	0.828/6.02
CO	0.871/6.92	0.768/6.16	0.607/4.75	0.573/4.45
NH ₃	0.589/7.86	0.571/7.52	0.588/7.79	0.497/6.52
HCN	12.8/10.3	12.0/9.43	10.2/7.88	9.99/7.62
巴豆醛 Crotonaldehyde	1.96/11.0	1.830/9.75	1.92/10.4	1.73/8.59

* :L 和 P 分别表示线性 and 多项式模型;数字 5 和 4 分别表示参数个数 (Letters L and P indicate linear and polynomial models, whereas digits 5 and 4 indicate the number of variables, respectively)。

集样品得出的,可能存在过拟合现象。为了提高模型的可靠性,必须通过独立的验证集样品进一步验证。

3.2 模型验证

3.2.1 传递模型 出于实际可操作性目的,验证样品的制备以及烟气成分释放量的测量在不同的卷烟加工厂完成。验证样品与校正样品在原料和生产工艺方面并不完全相同;另外烟气成分释放量的测量仪器也存在差异(分析方法则完全一致)。在这种情况下,在模型应用之前,需要进行模型传递。模型传递在光谱分析中的应用较为广泛,文献[11,12]给出了比较全面和系统的综述。

本研究中的模型传递采用 SBC 算法(Slope and bias correction)^[13]。SBC 算法具有性能良好,计算简单的优点,在模型传递中得到了广泛应用^[12]。SBC 算法的基本公式如下:

$$y_{\text{validation}} = s \cdot f_{\text{calibration}}(x_1, x_2, x_3, x_4, x_5) + b \tag{3}$$

其中, $f_{\text{calibration}}$ 为通过校正集确立的模型(以五参数模型为例); $f_{\text{validation}}$ 是被传递的,适用于验证集样品的模型; s 为斜率, b 为截距,代表了模型从校正集传递到验证集的系统误差。将同一来源的同批次样品中第一个样品作为基准,所以式(3)中的斜率 $s=1$ 。根据该基准样品的相应辅材参数和烟气成分释放量的实测值,求解出 b ,以完成模型传递。23 个验证样品来自 6 个不同产地,因此实际的独立验证样有 17 个。

3.2.2 验证模型 通过式(3)得到了适用于验证集样品的模型,预测了除 NNK 之外 8 种烟气成分的释放量,并与实测值进行比较,使用验证均方根误差(Root mean square error of validation, RMSEV)评价验证结果。下面以烟碱为例,说明模型验证过程。

对于烟碱的释放量,五参数线性 and 多项式模型,四参数线性 and 多项式模型的 RMSEV 分别是 0.0588, 0.0613, 0.0249 和 0.0303。可以看出,四参数模型的误差小于相应的五参数模型。进一步通过 t 检验确定四参数线性 and 四参数多项式这两种模型对烟碱释放量的预测值是否存在显著性差异,计算 t 值为 4.83,大于临界值 $t_{0.05, 16} = 2.1$ 。因此四参数线性模型的预测结果在统计意义上显著优于四参数多项式模型,被

确定为烟碱释放量的最终模型。

对其它烟气成分释放量的 4 种模型进行了验证,通过 RMSEV 和 t 检验确定了最终模型,结果见表 2。通过验证集数据确定的最佳模型和建模时 RMSEC 最小的模型并不一致(建模时 4 参数多项式模型是最好的)。结果表明,如果没有已知的确定性模型,建模阶段应该给出多个候选模型。候选模型必须通过独立的验证样进行验证,以确立最终模型。

表 2 通过 RMSEV 和 t 检验确立的最终模型*

Table 2 Final models determined with root mean square error of validation (RMSEV) and t -test

烟气成分 Component	模型 Model	烟气成分 Component	模型 Model	烟气成分 Component	模型 Model	烟气成分 Component	模型 Model
烟碱 Nicotine	4L	CO	5P	焦油 Tar	4P	NH ₃	5L
苯并[a]芘 Benzo[a]pyrene	5P	HCN	5L	苯酚 Phenol	5L	巴豆醛 Crotonaldehyde	4P

* : L 和 P 分别表示线性 and 多项式模型;数字 5 和 4 分别表示参数个数。Letters L and P indicate linear and polynomial models, whereas digits 5 and 4 indicate the number of variables, respectively.

3.3 定性分析

将此模型用于定性分析,即明确特定烟气成分释放量的决定性因素。因此,需要分析卷烟辅材参数对烟气成分释放量的影响程度。这种影响程度可以使用模型函数关于各参数的偏导数评价,而且通过线性模型计算比较简单。但是模型函数中各参数的量纲不同,所以不能直接求偏导数以研究其影响程度。为了统一量纲,将模型函数中的自变量 x_i 替换为 $\bar{x}_i \delta_i$,其中 δ_i 为新自变量,含义是该参数关于平均值的百分比, \bar{x}_i 即为平均值。通过此变换,模型函数中对应于各参数的自变量均为百分数,因此可以直接通过偏导数进行比较。此偏导数被定义为相应参数的影响权重,其物理意义是该参数相对于平均值发生 100% 的改变时,引起烟气成分释放量的变化。计算了 5 种参数的影响权重(表 3),其中绝对值最大值和最小值分别用黑体和斜体标出。

表 3 5 种参数对烟气成分释放量的影响权重

Table 3 Impacts of the 5 auxiliary materials on the release of the harmful components into mainstream smoke

烟气成分 Component	卷烟纸克重 Cigarette paper grammage	卷烟纸透气度 Cigarette paper permeability	成型纸透气度 Wrap paper permeability	接装纸透气度 Tipping paper permeability	滤棒吸阻 Pressure drop
烟碱 Nicotine	-0.0281	-0.114	-0.0452	-0.140	-0.368
焦油 Tar	-0.0371	-1.83	-0.804	-2.27	-5.19
苯并[a]芘 Benzo[a]pyrene	-1.45	-2.62	-0.633	-1.50	-4.86
苯酚 Phehol	-4.65	-1.83	-0.527	-1.84	-10.2
CO	3.53	-3.03	-1.01	-2.85	-2.20
NH ₃	-1.24	-0.748	-0.463	-1.16	-3.98
HCN	17.4	-32.2	-9.94	-35.3	-52.7
巴豆醛 Crotonaldehyde	4.27	-0.428	-0.0696	-3.33	0.809

黑体为最大值,斜体为最小值。Boldface is the maximum, itatics is the minimum.

从表 3 可见,对于多数烟气成分而言,滤棒吸阻的影响权重最大,此结果与烟草科学研究中的相关结论一致。有研究表明,烟气颗粒物受滤棒吸阻的影响较大,受通风稀释的影响较小;烟气的相物则受通风稀释的影响较大^[4]。本研究中的焦油、烟碱、B[a]P 和苯酚属于烟气颗粒物;卷烟烟气呈弱酸性,因此烟气中的 NH₃ 主要以 NH₄⁺ 形式存在于烟气颗粒物中;HCN 尽管是挥发性物质,但其分析方法是测定烟气中的氰化物。所以,这些成分的释放量主要受滤棒吸阻的控制。CO 属于烟气的相成分,受通风稀释的影响较大,这与表 3 中数据一致,即卷烟纸和接装纸透气度的影响权重大于滤棒吸阻,但卷烟纸克重的影响权重最大,可能与卷烟纸中 CaCO₃ 填料的高温分解有关。对于巴豆醛,由于卷烟纸克重增加,相应的纤维素用量也增加,而纤维素是巴豆醛的重要前体物之一,因此,卷烟纸克重的影响权重最大。从表 3 还可见,绝大部分情况下成型纸透气度的影响最小,这是由于成型纸只是辅助滤嘴成型,对烟气的稀释作用受外边包裹的接装纸的控制。

References

- Hoffmann D, Hoffmann I. Risks Associated with Smoking Cigarettes with Low Tar Machine-Measured Yields of Tar and Nicotine. *Smoking and Tobacco Control Monographs*, Chapter5. pp13, 2001. U. S. Department of Health and Human Services, National Institutes of Health, National Cancer Institute, Bethesda, MD
- YU Chuan-Fang, LUO Deng-Shan, WANG Fang, XI Nian-Sheng, WANG Bing, SHEN Yu-Jun(于川芳, 罗登山, 王芳, 席年生, 王兵, 申玉军). *Acta Tabacaria Sinica* (中国烟草学报), **2001**, 7(2): 1~7
- JIANG Wei, LI Bin, YU Chuan-Fang, LUO Deng-Shan(江威, 李斌, 于川芳, 罗登山). *Tob. Sci. Tech.* (烟草科技), **2007**, 9: 5~10
- Durocher D F, Mattina C F, Selke W A. *Beitr. Tabakforsch. Int.*, **1978**, 9(4): 201~207
- Morie G P. *Tob. Sci.*, **1976**, 20: 167~169
- Muramatsu M, Mikami T, Natio N, Tomita H. *Beitr. Tabakforsch. Int.*, **1977**, 9(3): 141~146
- Owen W C, Reynolds M L. *Tob. Sci.*, **1967**, 11: 14~20
- The Technical Center of Changsha Cigarette Factory (长沙卷烟厂技术中心). *Tob. Sci. Tech.* (烟草科技), **2001**, 8: 3~5
- XIE Jian-Ping, LIU Hui-Min, ZHU Mao-Xiang, ZHONG Ke-Jun, DAI Ya, DU Wen, XIE Fu-Wei, MIAO Ming-Ming, DENG Jia-Yun, NIE Cong, SHAO Xue-Guang, HU Qing-Yuan, YAO Xiao-Jun, LI Zhong, ZHANG Xiao-Bing(谢剑平, 刘惠民, 朱茂祥, 钟科军, 戴亚, 杜文, 谢复炜, 缪明明, 邓家云, 聂聪, 邵学广, 胡清源, 姚小军, 李忠, 张晓兵). *Tob. Sci. Tech.* (烟草科技), **2009**, 2: 5~15
- Official Method T-101, Determination of Ammonia in Mainstream Tobacco Smoke. Health Canada, December 31, **1999**
- CHU Xiao-Li, YUAN Hong-Fu, LU Wan-Zhen(褚小立, 袁洪福, 陆婉珍). *Spectrosc. Spectral Anal.* (光谱学与光谱分析), **2001**, 21(6): 881~885
- Feudale R N, Woody N A, Tan H, Myles A J, Brown S D, Ferré J. *Chemom. and Intell. Lab. Syst.*, **2002**, 64(2): 181~192
- Bouveresse E, Hartmann C, Massart D, Last I R, Prebble K A. *Anal. Chem.*, **1996**, 68(6): 982~990
- YAN Ke-Yu(闫克玉). *Cigarette Smoke Chemistry* (卷烟烟气化学). Zhengzhou(郑州), Zhengzhou University Press (郑州大学出版社), **2002**: 129~130

Building, Transferring and Validating Multivariate Models Between Properties of Cigarette Auxiliary Materials and Amount of Harmful Components in Mainstream Smoke

NIE Cong¹, XIE Fu-Wei², ZHAO Le¹, PENG Bin¹, YANG Jun², ZHANG You-Jin², SHAO Li-Min^{*2}

¹(Zhengzhou Tobacco Research Institute of CNTC, Zhengzhou 450001)

²(School of Chemistry and Material Science, University of Science and Technology of China, Hefei 230026)

Abstract This paper presented a systematic investigation into the relationship between the properties of cigarette auxiliary materials and the amounts of the harmful components in the mainstream smoke, including tar, nicotine, CO, HCN, benzo[a]pyrene (B[a]P), 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanone (NNK), Crotonaldehyde, phenol, and NH₃. We employed several chemometric methods to build, transfer and validate the multivariate models, and established the best model for each of the harmful components. We also investigated the impact of every auxiliary material on the release of the harmful components into the mainstream smoke, and obtained consistent results with corresponding conclusions in tobacco research. The investigation is not only beneficial to the manufacture of low harm cigarette, but also helps the data analysis protocol in tobacco industry.

Keywords Multivariate calibration; Model transfer; Mainstream smoke; Harmful components

(Received 15 March 2011; accepted 7 June 2011)