



**Textbook of**  
**化工数据教程**  
**Chemical Engineering Data**

马沛生 主编

# 化工数据教程

Textbook of Chemical Engineering Data

马沛生 主编



## 内 容 提 要

本书包括各种物性项的化工数据,书中系统介绍这些化工数据的定义、测定、收集、整理、评价、关联及估算,特别是其数据查找指引和大量估算方法评述等,全书包括总论、纯组分的基本物性、蒸气压和相变焓、热力学函数、 $pVT$ 及相平衡、黏度和导热系数、扩散系数和表面张力、网上化工数据资源获取途径和总览9章。

本书可作为化工及相关专业本科生教材及相应的研究生参考教材,也可供化工设计、研究、生产工作者进行化工设计及计算时的重要参考书。

## 图书在版编目(CIP)数据

化工数据教程/马沛生主编. —天津:天津大学出版社,  
2008.9

ISBN 978-7-5618-2788-8

I. 化… II. 马… III. 化学工业—数据—高等学校—  
教材 IV. TQ015

中国版本图书馆 CIP 数据核字(2008)第 149573 号

出版发行 天津大学出版社  
出 版 人 杨欢  
地 址 天津市卫津路 92 号天津大学内(邮编:300072)  
电 话 发行部:022-27403647 邮购部:022-27402742  
网 址 www.tjup.com  
短信网址 发送“天大”至 916088  
印 刷 天津泰宇印务有限公司  
经 销 全国各地新华书店  
开 本 185mm × 260mm  
印 张 14.25  
字 数 356 千  
版 次 2008 年 9 月第 1 版  
印 次 2008 年 9 月第 1 次  
印 数 1 - 4 000  
定 价 23.00 元

---

凡购本书,如有缺页、倒页、脱页等质量问题,烦请向我社发行部门联系调换

版权所有 侵权必究

# 前 言

在化工学科越来越趋于量化的当代,由于化工计算能给出定量的结果,计算水平的高低因此成为化工学科发展的重要标志。

化工数据是化工热力学的一个重要分支,是化工计算的基础之一,是化工设计、研究、生产、管理的必要依据。要实现严格的化工计算,需要探求化工过程规律,并以方程表达,而在求解过程中,化工数据是必不可少的,由此可总结为,化工计算的两大支柱是过程模型和化工数据。由于化工数据的重要性,它是化工高等教育中必不可少的部分。

化工数据应包括各项数据的定义、测定、收集、整理、评审、关联、估算。本书作为教材,其篇幅和对应的学时均有限,以上各项不能面面俱到,也考虑到大多数学生不需要“发展”化工数据学科,因而着重在其应用。本课程对上述内容作如下处理。

(1)化工数据的定义是简单的,基本上也已为读者所掌握,本书只是作一提示。

(2)了解化工数据的测定方法对准确评价化工数据的可靠性是有益的,也有助于读者优选数据。鉴于学时的限制,本书略去了这一部分。

(3)数据的收集、整理和评审主要由化工数据工作者进行,其成果主要反映在化工数据手册及数据库中。本书在第1章给出数据查找指南,又在各章更详细地介绍及评论了这些手册,在书末集中列出了这些手册,供本课程教学及其他课程参考,也便于各类化工工作者查找。这是本书重点内容之一,其中包括化工数据库的使用。

(4)化工数据的关联是重要内容,但比较简单,占篇幅不多。

(5)当缺乏实测化工数据时,为使化工计算顺利进行,化工数据的估算是必不可少的。估算方法多,内容复杂,占有本书的最大部分。作为教材,只能介绍一些最重要、最可靠、最通用、最方便的方法,同时指明其优缺点及可能的误差。

(6)本课程是化工热力学的后续课程,凡是在化工热力学教材中已详细讨论过的内容就不再重复了,例如涉及 $pVT$ 及活度系数时,只作为化工数据进行讨论。

本书的文献标注有其独特性,所有收载大量化工数据的手册、专著、期刊论文都以“数据查找指引”形式附于书的最后,并分为A、B、C、E四类。A类为有大量数据的综合型手册,但其中也有其他非化工数据的内容,重点不在数据;B类指该手册列出的化工数据是多种类型的,即该手册是综合型数据手册;C类指单项或单类

化工数据手册或含大量数据的期刊论文,并在每篇文献后注明数据项目;E类指估算方法专著。每章后所附“方法查找指引”主要是关于估算方法的。

作者在天津大学讲授本课程 20 余年,也曾编写过一些简明的讲义,供本校及兄弟院校使用,不足之处是内容过于简单。2003 年,出版的专著《化工数据》(中国石油化工出版社)内容较多且过细,不宜作教材。

本书作为教材,注意调整和压缩内容。课程的目的只要求引导学生入门,在解决具体问题时,还需参阅其他书籍。本教材的第 1 章(总论)和第 2 章(纯组分的基本物性)还是详实的,这是其重要性及作为实例而决定的,即使是少学时的课程教学计划也应详细讲解。

化工数据学科在较快地发展,新的数据、新的计算方法、新的数据手册在陆续发表,本教程注意收入近几年(至 2007 年底)发表的新内容;同时,也有许多新数据手册纳入书中。从网上寻找是获得化工数据的重要方法,这种方法有优点,也有明显的不足,本书新加一章,对有关内容作初步的介绍。在本书最后,列入“总览”(第 9 章),总结全书,也供学生一览本教程的要点。

本书可作为化工类本科或硕士生教材,同时也可供化工设计院、研究院、企业、管理部门的工程技术人员使用,推动化工计算的开展,并提高其可靠性。

作者未见国内外有类似的教材可供参考,虽然力求达到科学性与教学要求的统一,但因知识有限,不足之处恳请读者指正。

本书由马沛生主编,并撰写了第 1、2、4、5、7、9 章,夏淑倩撰写了第 3、6 章,丛玲撰写了第 8 章,梁英华参加了第 2 章的编写。

编者

2008 年 8 月



# 目 录

第1章 总论 .....	(1)
1.1 概述 .....	(1)
1.1.1 化工数据的定义 .....	(1)
1.1.2 化工数据的主要内容 .....	(1)
1.1.3 各类物性项的基本情况 .....	(2)
1.1.4 化工数据的应用 .....	(4)
1.2 化工数据的收集、整理和评价 .....	(4)
1.2.1 数据收集(处理)的一般过程 .....	(4)
1.2.2 数据的收集 .....	(5)
1.2.3 数据的评估 .....	(5)
1.2.4 数据机构 .....	(6)
1.2.5 化学化工数据库 .....	(6)
1.3 化工数据的回归 .....	(8)
1.3.1 目的和要求 .....	(8)
1.3.2 回归方程的选择 .....	(8)
1.4 数据手册 .....	(9)
1.4.1 化工数据手册的分类 .....	(9)
1.4.2 主要综合型化学化工数据手册 .....	(11)
1.5 数据查找指南 .....	(13)
1.6 化工物性的估算 .....	(15)
1.6.1 概述 .....	(15)
1.6.2 对应状态法 .....	(16)
1.6.3 基团贡献法 .....	(18)
1.6.4 用量子化学和统计热力学计算物性 .....	(19)
1.6.5 分子拓扑与物性 .....	(20)
1.6.6 估算方法专著 .....	(21)
1.7 不同类化合物对化工数据的使用 .....	(21)
第2章 纯组分的基本物性 .....	(23)
2.1 沸点和熔点 .....	(23)
2.1.1 概述 .....	(23)
2.1.2 数据 .....	(23)
2.1.3 估算方法 .....	(24)
2.2 临界参数 .....	(40)
2.2.1 概述 .....	(40)



2.2.2	数据	(40)
2.2.3	估算方法	(41)
2.2.4	临界压缩因子( $Z_c$ )	(53)
2.3	偏心因子	(53)
2.4	微观参数	(54)
2.4.1	偶极矩	(54)
2.4.2	Lennard-Jones12-6 参数	(55)
2.5	光学性质	(56)
	方法查找指引	(56)
<b>第3章</b>	<b>蒸气压和相变焓</b>	<b>(57)</b>
3.1	蒸气压的数据及关联式	(57)
3.1.1	温度关联式	(57)
3.1.2	数据	(58)
3.2	蒸气压的估算	(59)
3.2.1	对应状态法	(59)
3.2.2	基团贡献法	(61)
3.2.3	参考物质法	(65)
3.2.4	推荐意见	(69)
3.3	相变焓的数据	(70)
3.4	沸点下蒸发焓的估算	(70)
3.4.1	对应状态法	(70)
3.4.2	基团贡献法	(72)
3.4.3	推荐意见	(75)
3.5	$\Delta_v H_{298}^\ominus$ 的估算	(76)
3.6	$\Delta_v H_T$ 的关联和估算	(78)
3.6.1	$\Delta_v H_T$ 的关联方程	(78)
3.6.2	$\Delta_v H_T$ 的估算	(78)
3.7	溶度参数	(81)
3.8	混合热和超额焓	(81)
	方法查找指引	(81)
<b>第4章</b>	<b>热力学函数</b>	<b>(83)</b>
4.1	热容	(83)
4.1.1	概述	(83)
4.1.2	温度关联式	(83)
4.1.3	数据	(84)
4.1.4	$C_{p,C}^{\text{id}}$ 的估算	(85)
4.1.5	$C_{p,l}$ 的估算	(90)
4.2	焓和熵	(100)



4.3 燃烧焓、生成焓和生成 Gibbs 自由能 .....	(101)
4.3.1 燃烧焓(热) .....	(101)
4.3.2 生成焓(热) .....	(102)
4.3.3 生成 Gibbs 自由能 .....	(103)
4.4 超额焓 .....	(104)
4.5 闪点 .....	(104)
方法查找指引 .....	(108)
<b>第 5 章 <math>pVT</math> 及相平衡 .....</b>	<b>(109)</b>
5.1 纯液体的 $pVT$ .....	(109)
5.1.1 关联式(状态方程) .....	(109)
5.1.2 数据 .....	(111)
5.1.3 沸点下液体密度( $d_b$ )或摩尔体积( $V_b$ )的估算 .....	(111)
5.1.4 饱和液体密度( $d_s$ )或饱和液体摩尔体积( $V_s$ )的估算 .....	(112)
5.1.5 过冷液体密度的估算 .....	(118)
5.2 混合液体的 $pVT$ .....	(119)
5.2.1 数据 .....	(119)
5.2.2 混合规则 .....	(119)
5.2.3 液体混合物在泡点下密度的估算 .....	(120)
5.2.4 过冷混合液体密度的估算 .....	(120)
5.2.5 超额体积 .....	(121)
5.3 流体相平衡 .....	(121)
5.3.1 数据 .....	(121)
5.3.2 辛醇/水分配系数 .....	(123)
方法查找指引 .....	(128)
<b>第 6 章 黏度和导热系数 .....</b>	<b>(130)</b>
6.1 黏度关联方程和数据 .....	(130)
6.1.1 气体 .....	(130)
6.1.2 液体 .....	(131)
6.2 气体黏度的估算 .....	(131)
6.2.1 低压纯气体黏度估算 .....	(131)
6.2.2 低压混合气体黏度估算 .....	(136)
6.2.3 加压下气体黏度估算 .....	(138)
6.3 液体黏度的估算 .....	(141)
6.3.1 纯液体黏度估算 .....	(141)
6.3.2 混合液黏度的计算 .....	(148)
6.4 导热系数关联式和数据 .....	(148)
6.4.1 温度关联式 .....	(148)
6.4.2 导热系数数据 .....	(149)
6.5 气体导热系数的估算 .....	(149)





6.5.1 低压气体导热系数的估算 .....	(149)
6.5.2 低压气体混合物导热系数的估算 .....	(152)
6.5.3 加压气体导热系数的估算 .....	(153)
6.6 液体导热系数的估算 .....	(154)
6.6.1 纯液体导热系数的估算 .....	(155)
6.6.2 液体混合物导热系数的估算 .....	(157)
方法查找指引 .....	(158)
<b>第7章 扩散系数和表面张力</b> .....	<b>(159)</b>
7.1 扩散系数 .....	(159)
7.1.1 数据 .....	(159)
7.1.2 低压下二元气体 $D_G$ 的计算 .....	(160)
7.1.3 压力、温度及组成对 $D_G$ 的影响 .....	(161)
7.1.4 无限稀释下液体扩散系数 $D_{L,AB}^0$ .....	(161)
7.1.5 浓度、温度、压力对 $D_{L,AB}$ 的影响 .....	(163)
7.2 表面张力 .....	(164)
7.2.1 数据与温度关联式 .....	(164)
7.2.2 纯液体表面张力估算 .....	(165)
7.2.3 溶液的表面张力 .....	(175)
方法查找指引 .....	(176)
<b>第8章 网上化工数据资源获取途径</b> .....	<b>(178)</b>
8.1 获取网上数据信息的方法 .....	(178)
8.1.1 通过搜索引擎获取数据信息 .....	(178)
8.1.2 利用网络导航获得数据信息 .....	(179)
8.1.3 政府及专业网站平台 .....	(180)
8.2 在线化学化工数据网站 .....	(182)
8.2.1 物性数据库 .....	(182)
8.2.2 在线电子手册 .....	(187)
8.2.3 数据查询导航指南 .....	(189)
8.3 Internet 上的大型检索系统和文献型数据库 .....	(190)
8.3.1 STN 检索系统( <a href="http://www.chinastn.com/LINK.HTM">http://www.chinastn.com/LINK.HTM</a> ) .....	(190)
8.3.2 Dialog 数据检索系统( <a href="http://www.dialog.com">http://www.dialog.com</a> ) .....	(192)
8.3.3 Beilstein/Gmelin CrossFire .....	(193)
8.3.4 美国《化学文摘》数据库( <a href="http://www.cas.org">http://www.cas.org</a> ) .....	(194)
8.3.5 Knovel 电子工具书( <a href="http://www.knovel.com">http://www.knovel.com</a> ) .....	(195)
8.3.6 Landolt-Börnstein 网络工具书( <a href="http://www.landolt-börnstein.com">http://www.landolt-börnstein.com</a> ) .....	(195)
方法查找指引 .....	(197)
<b>第9章 总览</b> .....	<b>(198)</b>
<b>附录 符号表(通用)</b> .....	<b>(200)</b>
<b>数据查找指引</b> .....	<b>(202)</b>



# 第1章 总论

## 1.1 概述

### 1.1.1 化工数据的定义

化工数据有广义和狭义两种理解。广义是指与化工生产有关的各种数据,包括热力学数据、微观数据、传递性质数据、与化工安全和环保等内容有关的数据及反应速度数据等。狭义的理解所包括的范围要小些,它不包括反应速度数据,一般也不包括毒性、爆炸范围等数据。由于化学品对环保影响愈来愈受重视,与化学品有关的环保数据也已开始纳入化工数据范围内。被这样理解的化工数据也称为化工基础数据,化工数据中的绝大部分是各种纯物质或混合物的物理或化学性质,因此也被称为物化性质,或简称物性。本书只讨论狭义的也是最广泛被理解的化工数据。

### 1.1.2 化工数据的主要内容

#### 1) 定义

定义与其用途有关,也与其测定方法有关,当然是重要的,但一般说使用者是不难掌握定义的,因此不构成本书的重点。

#### 2) 测定方法

使用者大多不需要了解测定方法,特别是测定的技术、设备及其他细节,只有作数据评价时,上述内容才是重要的,因此,此内容不作为本书的重点。

#### 3) 数据的收集及评审

使用数据者希望在最短时间内找到基本可靠的数据,若要求不高,最方便的方法是使用数据手册,有时也要从原始测定文献或者从网上查找数据,如何找到这些信息本身就是一种“技能”。不同数据手册或文献中的数据有时很一致,但更可能有差别,甚至相差悬殊,因此数据评价是必需的,正确地选择数据更是一种技能。本书除介绍通用的评审规则外,也将在各章中针对各物性特点分别介绍。以上内容对使用者很重要,是本书重点内容之一。

#### 4) 数据的关联

绝大部分化工数据是温度、压力及组成的函数,实验测定结果总是逐个的数据点系列,为求得在工程计算中所需要的不同条件下的数据值,需要通过关联式把这些零碎的数据点表达为一种函数式,当输入温度、压力或组成后,就能方便而可靠地内插得出该条件下的物性值,还

可方便数据在计算机程序或过程模拟中使用。关联的主要工作是回归方程的选择和回归系数的求得,此项内容也是化工数据中的重要内容之一。

5) 估算方法

对纯物质摩尔质量又不大的物质而言,化工数据比较好找,但若要在大的温度或压力范围内,数据就很难找全。而在化工生产中,所处理的更多是混合物,混合物的物性数据是十分缺乏的,因此,在大多数设计或计算过程中,缺实验物性数据更常见,为此不得不寻找一种估算方法。在可接受误差前提下,选择合适的估算方法是化工数据的重要内容之一,并占有本书的最大篇幅。

1.1.3 各类物性项的基本情况

不同物性项的化工数据情况不同,到目前还没有公认的分类方法,按物性本身的规律而提出的分类目录见表 1-1。

表 1-1 主要化工数据项目表

分类	项 目	数据测定	数据收集	数据评审	估算方法	
					多少	可靠性
基本物性常数	三相点( $T_{tr}$ )	少	少	少	少	差
	熔点或凝固点( $T_m$ 或 $T_f$ )	多	多	少	少	差
	(常压)沸点( $T_b$ )	多	多	少	少	差
	临界温度( $T_c$ )	较少	多	多,好	多	好
	临界压力( $p_c$ )	较少	多	多,好	多	较好
	临界体积( $V_c$ )	较少	多	多,好	多	好
	临界压缩因子( $Z_c$ )	少	多	少	少	较差
	偏心因子( $\omega$ )	较少	较多	少	较少	较好
	常温液体密度或相对密度( $d^{20}$ 、 $d^{25}$ 或 $d_4^{20}$ 、 $d_4^{25}$ )	多	多	少	少	差
	常温折射率( $n_D^{20}$ 、 $n_D^{25}$ )	多	多	少	少	差
微观参数	偶极矩( $\mu$ )	多	多	较多	少	差
	Lennard-Jones 参数( $\sigma$ 和 $\epsilon/k$ )	较少	多	较少	少	差
热力学性质	气体 $pVT$ (包括超临界区 $pVT$ )	较少	较少	少	多	好
	气体第二维里系数( $B$ )	较多	较多	较少	较多	好
	液体密度( $d$ )	多	少	少	较多	较好
	蒸气压( $p^*$ )	多	多	多	多	好
	气体比热容( $C_{pG}$ 或 $C_{pG}^d$ )	多	多	较多	多	好
	液体比热容( $C_{pL}$ )	多	较多	较少	较多	较好
	固体比热容( $C_{pS}$ )	较多	较多	较少	少	差
	超额体积( $V^E$ )	较多	多	少	少	差
	超额焓( $H^E$ )	较多	多	少	少	差
	蒸发焓( $\Delta_v H$ )	多	多	较多	多	好
	熔化焓( $\Delta_m H$ )	多	较多	少	少	差
	升华焓( $\Delta_s H$ )	少	少	少	少	差
	标准燃烧焓(热)( $\Delta_c H^\ominus$ )	多	多	多,好	少	差
	标准生成焓(热) $\Delta_f H^\ominus$	多	多	多,好	多	好
	标准 Gibbs 生成自由能( $\Delta_f G^\ominus$ )	多	多	多	多	较好
	标准熵( $S$ )	多	多	多	多	好



续表

分类	项 目	数据测定	数据收集	数据评审	估算方法	
					多少	可靠性
相平衡数据	汽(气)液平衡(vle)	多	多	多,好	少	较差
	气液平衡(gle)	多	多	较少	少	较差
	液液平衡(lle)	多	多	较多	少	较差
	固液平衡(sle或lse)	多	较多	较多	少	差
	固固平衡(sse)	多	较少	较少	少	差
	气固平衡(gse)	较少	少	少	少	差
	超临界萃取(sfe)	较少	少	少	少	差
传递性质	气体黏度( $\eta_G$ )	少	较多	较多	较多	较差
	液体黏度( $\eta_L$ )	多	较多	较少	较多	较差
	气体导热系数( $\lambda_G$ )	少	较少	少	较少	差
	液体导热系数( $\lambda_L$ )	少	较少	少	较少	差
	表面张力( $\sigma$ )	多	多	较多	较多	较差
	气体扩散系数( $D_G$ )	少	少	少	较少	差
	液体扩散系数( $D_L$ )	少	少	少	较少	差

### 1) 基本物性常数

从表 1-1 可见,这些物性项中大部分是属于热力学性质的,又是唯一的(例如临界点、偏心因子、三相点)或是指定温度点( $d_4^{20}$ 、 $d_4^{25}$ 等)、指定压力点(常压沸点),因为具有单一值的性质,常被称为物性常数。

### 2) 微观参数

随着化工从经验科学向计算科学的发展,统计热力学、量子力学和分子热力学计算愈来愈成为化工计算的重要组成部分,因此作为上述计算所必需的微观参数也成为化工数据的一部分。目前已广泛进入实用计算的是偶极矩和势能函数,今后将有更多的微观参数进入化工计算。微观参数是不能用一般的物理或化学方法直接测出的。

### 3) 热力学性质

热力学性质的范围极其广泛,它可以包括已列入基本物性常数的三相点、沸点、凝固点、临界点等数据,也可以包括各种相平衡数据。考虑到相平衡数据量极大而且十分重要,本书把它们单独列为一类。

在表 1-1 中,热力学性质主要包括由  $pVT$  测量和量热法或热分析法测定所求得的数据。 $pVT$  数据包括气相  $pVT$  和第二维里系数数据,液体  $pVT$  数据,也包括高压下超临界区  $pVT$  和超额体积数据,蒸气压数据作为一种特殊的  $pVT$  也应列入其中。由量热法或热分析法可测定热容、各种相变焓、超额焓,结合化学反应还可测定燃烧焓和生成焓。对气体热容而言,数据主要来自结合统计热力学的光谱法。

热力学性质都是温度、压力的函数,因此蒸气压、蒸发热、气体热容、液体热容等热力学性质都需要温度函数式,而气液密度和热容更是温度、压力、组成的函数。

在大量热力学性质数据中,298.15 K 下标准态值最常用,也是列入数据手册最多的。

### 4) 相平衡数据

相平衡数据是在指定温度、压力下不同相组成间的关系,因此相平衡数据基本上都是混合物的数据。由于相平衡数据很难估算,因此不得不更依赖于实验测定,在不同混合物组分,不

同混合物组成间进行测定数据量是极大的,虽然相平衡数据可以是热力学数据中的一部分,但把相平衡数据作为单独一部分更合适些。相平衡数据量极大,甚至可大于其他热力学数据的总和;相平衡数据很重要,它是精馏、吸收、萃取、结晶等分离操作计算或设计的依据,因此在化学工程领域中,寻找相平衡数据常成为化工计算的关键。相平衡关系的核心常常是活度系数与浓度关系式,再配合各相组成间的关系式,而有关的化工数据主要就是活度系数关联式中的系数值。相平衡关系也可以用状态方程来表达,此时除需要可靠的临界数据外,还需要交互作用参数,该数据也需要由实验值关联求得。

在环境热力学数据中最重要的就是化学污染物在气、液、固相中的分布,其分布规律是要符合相平衡关系的,而相应的数据也是相平衡数据的一部分。

#### 5) 传递性质

黏度、导热系数、扩散系数、表面张力的产生也是基于分子间作用力的,因此可在分子热力学的基础上与其他物性统一起来。某些论文中把这些数据称为热物理数据,但热物理数据的范围更广泛些,它还包括热容、蒸气压、相变热等项热力学性质。

### 1.1.4 化工数据的应用

在化学工程或工艺计算中,可分为物料衡算、热量衡算和设备计算三大部分。在物料衡算中,相平衡计算是许多计算的基础,因而相平衡数据是必不可少的,例如精馏塔塔顶温度、塔釜温度都需要相平衡计算确定,塔顶分凝器及塔釜再沸器的汽液组成也是典型的汽液平衡组成,若塔顶基本上是纯组分,则塔顶温度与压力的关系就是该物质蒸气压关系,也就是需要蒸气压温度关联系数。在热量衡算中,物理过程需要热容或相变热的数据,而反应过程热计算则需要生成热。在设备计算中,一系列的准数值是必不可少的,以许多计算中十分重要的  $Re$  数为为例,就需要密度和黏度值,而且是不同温度、压力和组成下的值。总之,整个化工计算是离不开计算方法、计算模型、计算方程的,但这些方法、模型、方程都需要数据才能运行,因此化工数据在化工计算中同样是起关键作用的,没有化工数据就没有化工计算。应该指出,所需要的数据要适应各种化合物、各种组成、很大的温度应力范围,这就说明,化工数据不但重要,也很复杂,同时也说明在过程模拟计算中所花的机时绝大部分是用在有关化工数据部分的原因。

## 1.2 化工数据的收集、整理和评价

### 1.2.1 数据收集(处理)的一般过程

化工数据量极大,这首先是因为化学工业涉及的化合物极多,更因为需要更多的混合物数据,再考虑不同温度、压力条件下,数据量就难于计数了。以 CA 为例,每年 200 多万篇论文中,有关化学化工数据的论文至少在 10 万篇以上。一般的数据使用者是难于直接面对这样大量的数据的,因而在文献值与使用者之间需要一个中间媒介,其基本关系见图 1-1。

从图 1-1 可见,对一般使用者,寻找数据的主要途径是使用数据手册,或者从网上(变相的手册)寻找。为了便于内插使用,也为了使用计算机时方便,需要把实测的数据变成温度、压力和组成的函数关系式,因此回归工作或回归方程及回归系数的寻找也是数据处理的重要部分。

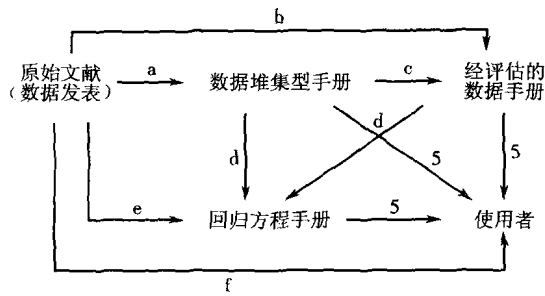


图 1-1 数据使用者取得数据的途径

a—数据收集;b—数据收集和评估;c—数据评估;  
d—数据回归;e—数据收集和回归;f—使用者寻找数据

## 1.2.2 数据的收集

数据处理中的第一步是数据的收集,最基本的也是最可靠的方法是查检索性文献(例如CA),也可以查找一些检索性巨型手册,例如 Beilstein 手册<sup>[A1]</sup>、Gmelin 手册<sup>[A2]</sup>等。如果在同样条件下,找到较多的数据时,就应该建立数据档案,把所查得数据按年代排列,并注明作者、方法、数值、可能的误差范围等,以便于比较和评估。这样的工作对物性的测定者是必须做的,以便检查自己工作的可靠性,有时还有助于确定误差范围。对一般数据使用者,没有能力也没有必要或者花精力去做这样的系统工作,因此本书只在讨论个别物性时对数据档案作举例介绍。在大多数场合下,比较方便或实用的方法是查找数据手册,或是在多本手册中寻找及评比。还应该说明,选择合适的手册或从网上查找,一般只是得到比较可靠的数据,不能保证得到更新更可靠的数据。

## 1.2.3 数据的评估

数据评估是在数据收集的基础上进行的,重要的化工原料或产品的化工数据往往被重复测定,对于沸点、熔点之类的数据,即便是工业上不重要的化合物,多次测定也是常见的。由于测定方法、技术的不同,温度和压力条件的控制精度不同,也由于原料纯度有异,甚至由于所使用的温标或摩尔质量不同,也可能引起测定值的差异,使用者应该学会对数据进行评价及选择,而不要查到一个就用。评价数据,要遵循一定的规则,以期得到较为可靠的推荐值。

主要有如下通用评估规则。

(1)一般选用由经典的实验方法所得的数据,若是新的测定方法,数据也要用与经典方法一致性来考核。若选用非标准方法应该有一定理由。

(2)采用较新年代的实验(或评选)数据。新的数据往往意味着新的测定方法、技术及设备,意味着较高的物料纯度。与测定方法相比,测定年代不是第一依据,当测定方法相同时,才更注意测定年代。

(3)宜采用经过评估的数据,也就是说可以信任其他人的评估。同样,对经过评估的数据手册或综述性数据文献可有更大的信任度,特别是信任有评估过程资料的手册或文献。

(4)优先采用高知名度的测定者或实验室的数据,因为他们一般具有丰富的经验和可靠的设备,可靠性高。

(5)注意作者自己公布的误差,作为参考。同时注意温度、压力的测定方法及精度,注意恒温装置的可靠性,注意分析方法的合理性。

(6)优先选择做过该项物性测定者的评估成果。

(7)了解实验的目的,优先采用专为本物性测定的文献,为了其他目的而附带测定的数据应尽量少用。

对每一项具体物性,还可使上述规定具体化,但原则不变。

经评估后,对所推荐的值,最好还能给出数据的可能误差,也可给出质量码,即分为 A、B、C、D 等等级,每一级相当于一定范围的可能误差。

#### 1.2.4 数据机构

从 19 世纪末就有数据工作者进行化学化工数据的整理工作,当时这些工作主要集中在德国。数据收集工作从个人开始,通过出版手册,逐步形成团体。在国际纯粹和应用化学联合会(Int. Union of Pure and Applied Chemistry)的推动下,在 20 世纪 20 年代出版了著名的“国际评选数据表”(International Critical Tables, 简称 ICT),整理并提供了大量数据,是 20 世纪 50 年代前化学化工者必用的手册。

一些数据机构在发展化学化工数据的收集及评价工作中起了重要作用。20 世纪 40 年代美国石油协会(API)在其 44 号计划推动下,组织了大量有关烃类的化学化工数据的收集、整理、测定、评估和回归工作,在 1953 年出版的手册也是很著名的,经常被引用。所建数据机构于 20 世纪 60 年代与 API 脱钩后成立热力学研究中心(TRC),一直从事烃类数据的系列工作,并逐步增加非烃类数据的研究。美国的热物理研究中心(TRRC)以处理热物理数据为中心,其中包括大量固体的物性数据,后来改名为 CINDAS,出版了许多数据专著。美国国家标准局(NBS)在组织美国数据机构活动中发挥了很好的作用,CINDAS 就是 NBS 所属的国家标准文献数据系统(NSDRS)成员之一。德国化工学会(DECHEMA)从 20 世纪 70 年代组织了一系列的化学化工数据收集和评估工作,出版了许多重要的物性数据手册,例如汽液平衡手册背靠强大的数据库,已成为世界著名的汽液平衡手册。另外,还有一些强大的系列手册系统,例如历史悠久、涉及面广、作者遍布世界的 LB 手册<sup>[R21][R22]</sup>。

数据机构早已向国际化发展,例如在 1969 年成立的“国际科技数据委员会”(Committee on Data for Science and Technology, 简称 CODATA)在组织和协调数据工作中起重大作用,该委员会每两年组织一次全球会议,CODATA 组织涉及各类数据,而化学化工数据组是 CODATA 工作中的重点工作组之一。

#### 1.2.5 化学化工数据库

##### 1. 概述

系统地收集、整理、评估及回归已发表的大量化学化工数据后,出版了大量数据手册。由于数据量极大,即使采用各项数据分类出版的方法,数据手册也难于容纳,更大的问题是时间上难于及时更新。为解决此难题,出现了数据库。另外,数据库在加入了人工智能及计算功能后,还能按使用者要求,作出内插、外推甚至组合的计算,还可能与一些化工软件连接,直接参与化工过程模拟计算。总之,数据库有许多优点是胜于数据手册的。

化工数据库是 20 世纪 60 年代以后,随着计算机及软件的发展而发展起来的,最初发展的



数据库大都属于管理型的,即用软件有组织地存储大量数据,供检索使用。这样的数据库与图书馆很类似,数据相当于图书,外存相当于书库,用户相当于读者,数据的组织方法相当于图书的存放办法。

在管理型数据库发展5~10年后,科学数据库开始发展,其特点是:

- (1)可针对个别学科或项目,不一定是大型的;
- (2)重点在数据的收集、评价及加工上,所提供数据常有质量码;
- (3)在数据处理、计算及使用方法上具有一定的人工智能功能;
- (4)与更多的数学运算相结合,包括数据的回归、方程计算等。

在各种学科中,化学化工要涉及最多的化合物品种和各类物性项目,需要更多的信息,因此建立数据库的必要性最大,有关的活动也最活跃。

## 2. 一般要求

化学化工数据库的一般功能如下。

### 1) 数据检索

一个好的数据库首先在于存入大量可靠的数据,对于化学化工数据库这一点尤显重要,有时以化合物数作为评比数据库的主要依据之一。所存入数据应该经严格的核对和评估,又应该能方便地查到数据源,并定时更新,对目前大部分在运行中的数据库上述要求都是有差距的。总体说,数据库中所涉化合物及所存数据量在不断增加。

化学化工数据库所存数据项有:

(1)基础物性和热力学性质是化学化工数据库中的主要组成部分,主要差别在于化工数据库的热力学性质要关注不同温度、压力条件下的数据,而化学数据库一般只存入标准态下的数据;

(2)相平衡数据是化工数据库中的重要组成部分,由于数据量极大,相平衡数据常以单独数据包出现,在化学数据库中一般不包括此项数据;

(3)传递性质数据也只出现在化工数据库中,且应该包括不同温度和压力条件;

(4)光谱、质谱、核磁及多类微观性质可列入化学数据库中,也可单独建库或专用数据库,并常被称为结构数据库。

以上是化学化工数据库总体的可能内容,但化学化工数据库也有综合型和专项型数据库之分,不同数据库质量是不同的。对综合型数据库,数据面广,总量又大,每项物性数据未必全,质量也不一定高,因此一般说用专项性数据库好些。

### 2) 数据的运算

数据库需要能提供不同温度、压力或组成下的物性值,一般的数学运算是简单的,但库中需要有回归方程及相应的系数值,部分系数值可从手册或文献中查得后存入库中,另外库中也要能选择回归方程和进行回归。此外,数据库还要能进行更复杂的运算,汽液平衡数据的热力学一致性校验是一例,不同方法的汽液平衡组成计算是另一例,状态方程的计算是又一例。为了应付不同的数学计算,库中必须存有一大批数学计算软件。

### 3) 人工智能

当缺乏实验值时,数据库应能提供估算值,估算方法可由使用者选择,或由数据库根据不同情况(不同化合物、不同温度压力条件、不同精度要求)下代替使用者选择恰当方法。此外,在数据评价等方面若能建立一些专家系统,把专家的经验计算机化就是更高的要求了。





#### 4) 操作性能、连接条件和管理系统

有些化工数据库是专门与过程模拟计算配套的,这样的数据库重视反应迅速、连接恰当。许多大型化学化工数据库重点在于数据检索,这种检索型数据库不求检索速度快,重点在于数据的质量。一个好的数据库应能适应不同规模的机器,应能适应各种操作系统,具有图形处理功能,具有良好的多种接口软件,并能支持实现网络化。

### 3. 数据库的国际化

个人力量有限,数据库涉及面又广,因此常常要依靠一个团体进行建设。数据库要进行联网,并走向国际化的使用及建设。在 CODATA 的工作中,数据库的交流安排为重要内容。

### 4. 化学数据库与化工数据库

两者有共同点,也有较大差别。这也表明,化学数据与化工数据的异同,化学数据有更多甚至大量的微观数据,甚至形成单项数据手册或数据库,例如各种光谱或色谱数据;化工数据包括更多的相平衡数据及传递性质数据。化工数据库要包括更多的计算功能也是重要特色。

### 5. 大型数据手册发展的数据库

大型数据手册有很强大的工作队伍,为方便使用,也常常开展了数据库的工作,例如著名的 Beilstein 手册就同时有 Cross Fire Beilstein 数据库,其信息来源为该手册从第 4 版正编到第 4 补编的全部内容和 1960 年以来原始文献数据,涉及 800 万种以上化合物,900 万个以上反应,有物理性质、化学性质、药效学、环境毒理学等,功能强大,使用方便,被称为世界上最大的化学信息库。以物理性质而言,其不足之处是数据整理过程不甚透明。

## 1.3 化工数据的回归

### 1.3.1 目的和要求

实验所测得的只是在不同温度、压力或组成下分散的实验点。在早期人们习惯于用作图的方式联系这些物性点,并作内插使用,这样的方法虽然非常直观,但当数据曲线有复杂变化时,误差可能较大,外推时更可能有很大的偏差。更好的方法是把这些数据变成温度、压力或组成的函数,这样的方程叫回归方程。使用回归方程后不但内插时误差小些,外推时成功的可能也有增加,此外还可直接用于计算机计算,在过程模拟计算中更是不可缺少的。

对回归方程的一般要求:

- (1) 尽可能反映实验数据点,回归误差小;
- (2) 尽可能有较强的理论基础,这样的模型及方程适用范围大,有时还能适度外推,当理论基础虽好而误差却偏大时,一般更重视误差的大小;
- (3) 一般选用比较简单而又通用于各种化合物的回归方程,以求减少过程模拟计算中所占机时,由实验值确定的回归系数也不要太多。

### 1.3.2 回归方程的选择

回归方法和方程可按如下原则进行。

- (1) 若有大量实验数据,且要求很高和使用范围很宽的回归精度,可以选用很复杂的回归方程,水蒸气表中的物性回归方程就是一例。这种方程中可以有十几个甚至几十个方程系数,