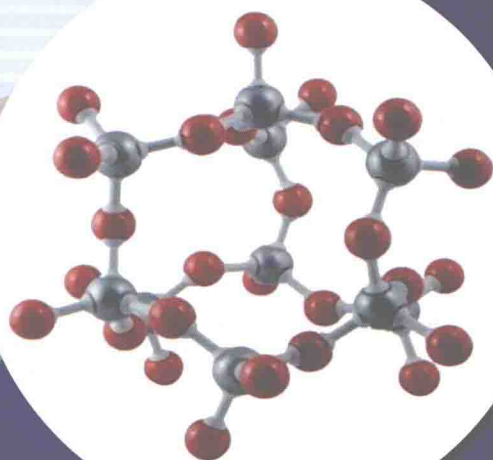


拓扑指数的应用

——烃类物质和持久性有机污染物的
定量结构性质关系研究

焦龙◎著

TUOPU
ZHISHU
DE YINGYONG



中国石化出版社

[HTTP://WWW.SINOPEC-PRESS.COM](http://www.sinopec-press.com)

拓扑指数的应用

——炔类物质和持久性有机污染物的
定量结构性质关系研究

焦龙 著



中国石化出版社

内 容 提 要

本书在介绍定量结构性质关系的基础知识,以及定量结构性质关系研究中常用的拓扑指数和化学计量学建模方法的基础上,重点介绍了可用于建立烃类物质的定量结构性质关系模型的拓扑指数和建立持久性有机污染物的定量结构性质关系模型的拓扑指数,以及如何用这些拓扑指数建立定量结构性质关系模型。此外本书还介绍了拓扑指数在建立其他常见有机物的定量结构性质关系模型中的应用。

本书可以作为化学、计算机和数学专业高年级本科生、研究生的专业课教材,以及高校教师和相关专业技术研究人员的参考书。

图书在版编目(CIP)数据

拓扑指数的应用:烃类物质和持久性有机污染物的
定量结构性质关系研究 / 焦龙著. —北京:中国石化
出版社, 2017.7

ISBN 978-7-5114-4579-7

I. ①拓… II. ①焦… III. ①持久性-有机污染物-
研究 IV. ①X5

中国版本图书馆 CIP 数据核字(2017)第 176794 号

未经本社书面授权,本书任何部分不得被复制、抄袭,或者以任何
形式或任何方式传播。版权所有,侵权必究。

中国石化出版社出版发行

地址:北京市朝阳区吉市口路9号

邮编:100020 电话:(010)59964500

发行部电话:(010)59964526

<http://www.sinopec-press.com>

E-mail:press@sinopec.com

北京富泰印刷有限责任公司印刷

*

700×1000 毫米 16 开本 14.5 印张 276 千字

2017 年 9 月第 1 版 2017 年 9 月第 1 次印刷

定价:50.00 元

前 言

定量结构性质关系研究是化学计量学的重要分支和前沿研究领域。近年来欧洲和美国化学计量学领域的专家学者撰写出版了多部关于定量结构性质关系研究的专著，介绍了该领域的研究进展和主要成果。但目前国内出版的相关专著还非常少，这很大程度限制了国内研究人员和学生对该领域优秀研究成果的了解和学习。

目前国内外出版的化学计量学专著很少专门介绍最新的拓扑指数方法及其在定量结构性质关系研究中的应用，本书选择这个主题，对此内容做了详细地介绍和阐述，介绍了定量结构性质关系以及拓扑指数的基础知识，并主要围绕当前拓扑指数在烃类物质和持久性有机污染物的定量结构性质关系研究中的进展和应用展开论述。全书共分6章：绪论部分对全书内容进行了整体概括介绍；第1章简要介绍了定量结构性质关系的基础知识，以及定量结构性质关系研究中常用的拓扑指数、化学计量学建模方法以及模型的验证方法；第2章主要介绍了可用于建立烃类物质定量结构性质关系模型的拓扑指数，并介绍了如何用这些拓扑指数建立相应的定量结构性质关系模型；第3章主要介绍了可用于建立持久性有机污染物定量结构性质关系模型的拓扑指数，并介绍了如何使用这些拓扑指数建立相应的定量结构性质关系模型；第4章介绍了拓扑指数在建立其他常见有机物定量结构性质关系模型中的应用；第5章介绍了最新的、以拓扑指数为结构描述符建立的混合物定量结构性质关系模型。

本书是一本独特的化学计量学专著，介绍了关于拓扑指数的主要研究方法、新的研究进展以及主要研究成果，可以作为化学、计算机和数学专业高年级本科生和研究生教材，还可以作为高校教师和相关研究人员的参考书，可为国内化学计量学工作者和相关研究人员从事定量结构性质关系研究提供有用的研究资料。

感谢国家自然科学基金项目 No. 21775118、中国石油科技创新基金研究项目 No. 2015D-5006-0407、陕西省青年科技新星计划项目 No. 2016KJXX-16、西安石油大学科研创新团队 No. 2013QNKYCXTD01、西安石油大学校级拔尖人才计划对于本书的支持。

感谢中国石化出版社许倩和王雷两位编辑对本书的撰写和出版给予的帮助。感谢西安石油大学化学化工学院近年来为作者科研工作提供的支持。感谢西安石油大学化学化工学院硕士研究生秦玉翠为本书撰写做出的贡献。

由于水平有限，书中不足之处在所难免，敬请广大读者提出批评和意见，以便改进和提高。

目 录

0 绪论	(1)
1 定量结构性质关系基础知识	(4)
1.1 化学计量学	(4)
1.2 定量结构性质关系的定义及相关概念	(5)
1.3 结构描述符	(8)
1.3.1 拓扑指数	(8)
1.3.2 量子化学参数	(13)
1.3.3 宏观性质参数	(14)
1.4 建立定量校正模型的主要方法	(14)
1.4.1 多元线性回归	(14)
1.4.2 偏最小二乘	(15)
1.4.3 人工神经网络	(15)
1.4.4 变量选择方法	(21)
1.5 QSPR 模型验证方法	(28)
1.5.1 外部测试集验证	(28)
1.5.2 留一交叉验证	(29)
1.5.3 k 折交叉验证	(30)
1.6 QSPR 模型评价参数	(30)
参考文献	(34)
2 烃类物质的定量结构性质关系研究	(41)
2.1 烃类物质的定量结构性质关系研究现状	(41)
2.1.1 从分子水平认识烃类物质的性质	(41)
2.1.2 烃类物质的定量结构性质关系模型	(42)
2.2 烷烃密度的定量结构性质关系模型	(43)

2.2.1	分子连接性指数模型	(43)
2.2.2	电性拓扑状态指数模型	(43)
2.2.3	烷烃密度的定量结构性质关系研究总结	(47)
2.3	烷烃辛烷值的定量结构性质关系研究	(48)
2.3.1	烷烃辛烷值的测定和预测方法	(48)
2.3.2	电性拓扑状态指数模型	(49)
2.3.3	自相关拓扑指数模型	(52)
2.3.4	基团贡献法结合拓扑指数方法	(53)
2.3.5	拓扑指数结合量子化学方法	(54)
2.4	烷烃色谱保留指数的定量结构性质关系研究	(56)
2.4.1	烷烃的色谱保留指数及其计算方法	(56)
2.4.2	分子矩边矢量指数模型	(58)
2.4.3	改进的原子点价指数 mQ 模型	(62)
2.4.4	分子连接性指数结合奇偶指数模型	(69)
2.4.5	原子支化度 N 指数模型	(71)
2.4.6	1R 指数模型	(73)
2.5	烃类化合物爆炸下限的定量结构性质关系研究	(75)
2.5.1	烃类化合物的爆炸下限及其预测方法	(75)
2.5.2	改进的原子点价指数 mQ 模型	(76)
2.6	烃类物质生成焓的定量结构性质关系研究	(80)
2.6.1	烃类物质的生成焓及其预测方法	(80)
2.6.2	分子连接性指数模型	(81)
2.6.3	Z_h 拓扑指数模型	(82)
2.6.4	拓扑指数结合量子化学方法	(84)
2.7	烷烃定量结构性质关系研究中的其他方法	(93)
2.7.1	HQSAR 方法	(93)
2.7.2	CoMFA 模型	(97)
	参考文献	(100)

3 持久性有机污染物的定量结构性质关系研究	(106)
3.1 持久性有机污染物的概念及其性质测定存在的问题	(106)
3.2 多溴联苯醚色谱相对保留时间的定量结构性质关系模型	(107)
3.2.1 多溴联苯醚及其色谱相对保留时间	(107)
3.2.2 <i>MDEV</i> 指数	(108)
3.2.3 数据集	(109)
3.2.4 建立的定量结构性质关系模型	(109)
3.3 多溴联苯醚正辛醇/空气分配系数的定量结构性质关系模型	(116)
3.3.1 多溴联苯醚的正辛醇/空气分配系数及其预测方法	(116)
3.3.2 多元线性回归模型	(117)
3.3.3 人工神经网络模型	(117)
3.4 多溴联苯醚过冷液体蒸气压的定量结构性质关系模型	(119)
3.4.1 多溴联苯醚的过冷液体蒸气压及其预测方法	(119)
3.4.2 多元线性回归模型	(120)
3.4.3 人工神经网络模型	(120)
3.5 多氯联苯气/粒分配系数的定量结构性质关系模型	(123)
3.5.1 多氯联苯及其气/粒分配系数	(123)
3.5.2 多氯联苯的 <i>MDEV</i> 指数	(124)
3.5.3 数据集	(125)
3.5.4 多元线性回归模型	(125)
3.5.5 线性人工神经网络模型	(126)
3.6 多氯联苯水溶解度的定量结构性质关系模型	(128)
3.6.1 多氯联苯的水溶解度及其预测方法	(128)
3.6.2 分子连接性指数-分子价键连接性指数模型	(129)
3.6.3 定位拓扑指数-基团对应指数模型	(129)
3.6.4 自相关拓扑指数模型	(130)
3.6.5 <i>MDEV</i> 指数模型	(131)
3.6.6 结构信息价连接性指数模型	(134)

3.7	多氯联苯正辛醇/水分配系数的定量结构性质关系模型	(135)
3.7.1	多氯联苯的正辛醇/水分配系数及其预测方法	(135)
3.7.2	定位拓扑指数-基团对应指数模型	(135)
3.7.3	结构信息价连接性指数模型	(136)
3.7.4	拓扑距离特征指数模型	(137)
3.7.5	MDEV 指数模型	(137)
3.8	多氯联苯沉积物/水分配系数的定量结构性质关系模型	(139)
3.9	多氯代二苯并二噁英/呋喃光解反应半衰期的定量结构性质关系模型	(141)
3.9.1	多氯代二苯并二噁英/呋喃及其光解反应的半衰期	(141)
3.9.2	PCDD/Fs 的 MDEV 指数	(142)
3.9.3	数据集与模型	(143)
3.10	PCDD/Fs 沸点的定量结构性质关系模型	(155)
3.10.1	PCDD/Fs 的沸点及其预测方法	(155)
3.10.2	数据集与模型	(155)
3.11	PCDD/Fs 水溶解度的定量结构性质关系模型	(164)
3.11.1	PCDD/Fs 的水溶解度及其预测方法	(164)
3.11.2	数据集与模型	(165)
	参考文献	(174)
4	烃类物质与其他常见有机物的定量结构性质关系研究	(182)
4.1	烃、醚及醇类物质闪点的定量结构性质关系模型	(182)
4.1.1	烃、醚及醇类物质的闪点及其预测方法	(182)
4.1.2	数据集与电性拓扑状态指数	(182)
4.1.3	多元线性回归模型	(188)
4.1.4	PLS 模型	(189)
4.2	苯及其衍生物毒性的定量结构性质关系模型	(190)
	参考文献	(191)
5	混合物的定量结构性质关系模型研究	(192)

5.1 混合物的定量结构性质关系研究现状 (192)

5.2 常见有机混合物的闪点定量结构性质关系模型研究 (195)

5.2.1 常见有机混合物的闪点及其预测方法 (195)

5.2.2 数据集与混合物描述符 (195)

5.2.3 多元线性回归模型 (211)

5.2.4 逐步回归模型 (212)

5.2.5 人工神经网络模型 (213)

5.2.6 常见有机混合物的闪点定量结构性质关系模型研究总结 (215)

5.3 苯及其衍生物所形成混合物的毒性定量结构性质关系模型 (215)

5.4 全氟羧酸混合物的毒性定量结构性质关系模型 (217)

参考文献 (219)

0 绪论

定量结构性质关系(Quantitative Structure-Property Relationship, QSPR), 又称为定量构效关系(Quantitative Structure-Activity Relationship, QSAR), 是研究物质结构与性质之间的定量数学关系, 并由此预测物质性质的一类化学计量学方法的总称, 是化学计量学的主要研究领域之一。作为一种重要的研究方法, 定量结构性质关系已在药物化学、环境化学、分析化学等领域中得到了广泛的应用。

拓扑指数是定量结构性质关系研究中常用的一种结构描述符, 能够用数字定量表示物质的结构。作为一种能够反映分子结构的特征变量, 拓扑指数利用了拓扑学的理论和方法, 将物质的分子结构图抽象成一种可以用数字来表示的数值量, 以此表示其分子结构特征。拓扑指数的实质是一类来自分子结构图的拓扑学参数, 是分子结构图的拓扑不变量, 其优点主要有形式简明、计算简单、能够区分结构不同的物质, 以及建模实用性好等。拓扑指数在定量结构性质关系研究中已有长期和广泛的应用, 起着非常关键的作用。近年来欧洲和美国化学计量学研究领域的学者撰写出版了多部关于定量结构性质关系研究的专著, 其中介绍了该领域近年来包括拓扑指数在内的相关研究进展和主要成果。但目前国内出版的相关专著还非常少, 这些专著很少涉及到拓扑指数在烃类物质和持久性有机污染物定量结构性质关系研究中的应用, 这很大程度限制了国内研究者和学生对该领域优秀研究成果的了解和学习。

烃类物质是结构最简单的一类基本有机化合物, 也是石油天然气的主要组成物质。烃类物质的物理化学性质是石油天然气加工和相关化学工业领域生产和研究工作中必需的基本数据。现代分析化学能够通过多种手段测定烃类物质的性质, 但烃类物质数量众多(需要测定的性质同样非常多), 通过实验方法测定其性质, 工作量非常庞大、所需时间和经济成本较高。此外, 由于标准品不易制得、测定技术手段有限等原因, 一些烃类物质的性质目前尚不能通过实验测定。因此, 仅通过实验方法测定烃类物质的性质还不能满足实际生产和科研的需要。定量结构性质关系方法是一种从分子原子水平上掌握物质结构与性质信息的重要手段, 建立烃类物质的定量结构性质关系模型, 并由此预测其性质是一种能够简单、快速、经济地得到烃类物质性质数据的方法。通过研究定量结构性质关系模型还能够从分子水平上理解烃类物质结构与性质间的关系, 为有机合成、有机分

析、石油天然气化工提供必要的理论指导。因此,定量结构性质关系方法能够作为从分子水平上认识烃类物质结构与性质的一种有效途径。开展烃类物质的定量结构性质关系研究,有非常重要的理论和实际意义。

持久性有机污染物(Persistent Organic Pollutants, POPs)是指具有持久性(长期残留性)、生物累积性、半挥发性和一定毒性,能够通过各种环境介质(大气、水、生物体等)长途迁移并长期存在于环境,对人类健康和环境有严重危害的有机物。持久性有机污染物目前已成为全球性的环境问题,包括我国在内的92个国家于2001年5月签署了《斯德哥尔摩公约》,旨在禁止或限制持久性有机污染物的生产和使用。该公约成员国目前已达到124个。尽管已被禁止或限制使用,但由于持久性有机污染物有很强的长期残留性,加之一些持久性有机污染物的产生不受人类控制,持久性有机污染物将在环境中长期存在,对环境构成严重的威胁。因此,必须全面深入研究这类物质对环境和生物的影响,解决由此带来的各种生态问题。有关持久性有机污染物的研究现已成为环境科学的重点研究领域之一,是一个受到广泛关注、十分活跃并有非常大发展潜力的领域。持久性有机污染物的性质测定目前主要是通过实验手段进行的。实验方法直观、可靠,是不可替代的取得持久性有机污染物性质数据的方法。但实验方法又不可避免地存在有一定的局限性,需要其他方法的协作和补充。建立持久性有机污染物的定量结构性质关系模型,并由此预测其性质是一种能够简单、快速、经济地得到持久性有机污染物性质数据的方法。为了更准确地预测与测定持久性有机污染物的性质及环境行为,评估其生态效应、环境风险与健康危害,预测与测定持久性有机污染物的毒性,解释相应的致毒机理,以及研究建立更高效的POPs污染控制与消除技术手段,必须从分子水平上认识持久性有机污染物结构与性质之间的内在关系,为相关研究提供必要的理论指导。通过对持久性有机污染物定量结构性质关系模型的研究,能够充分解释其分子原子结构与性质之间的内在联系,因此,开展持久性有机污染物的定量结构性质关系研究,有非常重要的意义。

除了研究纯物质的定量结构性质关系,在实际工作中研究者往往会更多地遇到需要研究混合物性质的问题,这需要建立混合物的定量结构性质关系模型,由混合物的组成和结构信息预测其性质。可以说,研究建立由若干烃类物质形成的混合物或者由若干持久性有机污染物形成的混合物有重要的实践意义。此外,建立混合物的定量结构性质关系模型需要综合考虑混合物的组成、组分结构以及组分间的相互作用这三方面的因素,这比仅考虑物质结构的纯物质的定量结构性质关系模型更为复杂,能够发展化学计量学定量结构性质关系的基本方法和理论,有着重要的理论意义。

目前,烃类物质和持久性有机污染物的定量结构性质关系研究领域取得了

很多有价值的成果，建立了一批有很好实用价值的定量结构性质关系模型。这些模型可以用于这些物质的性质预测以及相关理论研究。鉴于此，本书主要介绍了拓扑指数在烃类物质和持久性有机污染物定量结构性质关系研究中的研究进展和主要成果，介绍了当前烃类物质和持久性有机污染物的定量结构性质关系研究中常用的拓扑指数、相关建模方法及建立的定量结构性质关系模型，并以部分定量结构性质关系模型为例介绍了以拓扑指数为结构描述符建立烃类物质和持久性有机污染物定量结构性质关系模型的实例；这是目前国内定量结构性质关系和化学计量学专著都没有涉及到的内容。此外，本书还解释了常用的定量结构性质关系建模方法，并指导读者如何进行相关的计算；这可以作为读者进行相关建模计算的实用指导。

1 定量结构性质关系基础知识

1.1 化学计量学

分析化学作为一门获取物质定性和定量信息的科学，长期以来在科学和生产实践中发挥着重要的作用。近年来，分析化学正经历巨大的变革，化学家逐渐认识到分析化学是通过化学量测获取化学信息的科学。随着各类科学技术的发展，数学、物理、计算机和新仪器被引入分析化学，这使得分析化学家可以在相对短的时间内得到大量的原始分析数据。如何处理和利用分析过程中得到的原始数据，提取有用的相关化学信息，这是现代分析化学向分析化学家提出的新要求和新的挑战。化学计量学正是为了适应这个要求和迎接这个挑战而逐渐发展起来的一门化学分支学科^[1-3]。

“化学计量学(chemometrics)”是由瑞典化学家 S. Wold 于 20 世纪 70 年代初类比于生物计量学(biometrics)和经济计量学(econometrics)提出的名词。S. Wold 提出，化学计量学是研究从化学实验产生的数据中提取相关化学信息的化学学科分支。这一建议得到了美国化学家 B. R. Kowalski 的响应。他们于 1974 年倡导成立了国际化学计量学学会(International Chemometrics Society, ICS)。简而言之，化学计量学就是运用数学、统计学、计算机科学以及其他相关学科的理论与方法，优化化学量测过程，并从化学量测数据中最大限度地提取有用的化学信息的一门科学^[1-3]。也可以说化学计量学是一门关于化学量测的基础理论与方法学。

化学计量学的研究内容涵盖了化学测量的全过程，包括采样理论与方法、实验设计与优化方法、分析检测理论和信号处理方法、分析信号的多元校正与分辨、化学模式识别、定量构效关系(即定量结构性质关系)、数据库检索、人工智能与化学专家系统等丰富的学科分支^[2]。这些化学计量学理论和方法的发展为现代分析化学研究提供了新的手段，同时也大大丰富了化学学科的基础理论和方法。研究化学计量学的基础理论和算法及其在各种领域中的应用，对现代分析化学的发展具有重要意义。因此，化学计量学方法越来越得到广大化学工作者的重视。

当前化学计量学研究主要分为两个方面：一方面是化学计量学基础理论与方法研究，即把其他领域特别是数学中的许多有用算法和工具，如小波分析、人工神经网络(Artificial Neural Network, ANN)、支持向量机(Support Vector Machine, SVM)、随机森林(Random Forest, RF)等引入化学计量学并结合化学体系的独有特征，建立适合于解决化学问题的化学计量学算法。另一方面，将化学计量学方法应用到化学及其他相关领域，如药物化学、环境科学、食品科学、生物科学等，解决依靠传统分析化学方法无法处理的问题。这两方面相辅相成，互相促进。实际问题的解决，促进了化学计量学在各化学相关领域的应用；而随着研究问题的复杂化，反过来又对化学计量学提出了更高的要求，促使化学计量学新方法的出现。

1.2 定量结构性质关系的定义及相关概念

定量结构性质关系，又称为定量构效关系，是研究物质结构与性质之间的定量数学关系，并由此预测物质性质的一类化学计量学方法的总称。定量结构性质关系是化学计量学的主要研究领域之一，已在药物化学、环境化学、分析化学等领域中得到了广泛的应用。定量结构性质关系研究领域的学术论文主要集中发表在 J. Chem. Inf. Model., Chemometr. Intell. Lab. Systems., Chemoshpere, Anal. Chem., Anal. Chim. Acta, J. Chromatogr. A, J. Harzard. Mater., SAR QSAR Environ. Res., Chem. Phys. Lett., J. Med. Chem., J. Chemometr., J. Pharm. Sci., J. Mol. Graph. Model. 等期刊。

定量结构性质关系研究的基本思想是：分子(包括原子、离子)是构成物质的基础结构，化合物的分子结构决定了其宏观性质，因此可以根据化合物的结构预测其性质。从本质上说，定量结构性质关系研究就是在分子水平上研究物质的组成结构与其性质之间的定量数学关系^[2, 5]，将物质的性质表达为结构的函数，即：

$$\text{性质} = f(\text{结构}) \quad (1-1)$$

显然，物质的性质可以表达为一个数字。但物质的结构不能直接作为式(1-1)所示函数的自变量。为了建立 QSPR 模型，即建立式(1-1)所示的函数，必要条件之一是用数字定量表示物质的结构。定量结构性质关系研究中将这种用于表示物质结构的数字(数据)称为结构描述符(structural descriptors)。常用的结构描述符有物理化学性质参数、拓扑指数、量子化学参数和场参数等^[6-9]。通过一定的数学方法进行数据分析，可以建立结构描述符与性质之间定量关系的数学模型，即将性质表达为结构描述符的一个函数，即：

$$\text{性质} = f(\text{结构描述符}) \quad (1-2)$$

这个函数就称为是所研究物质的 QSPR 模型。因此 QSPR 模型的建立也可以用图 1-1 来说明。

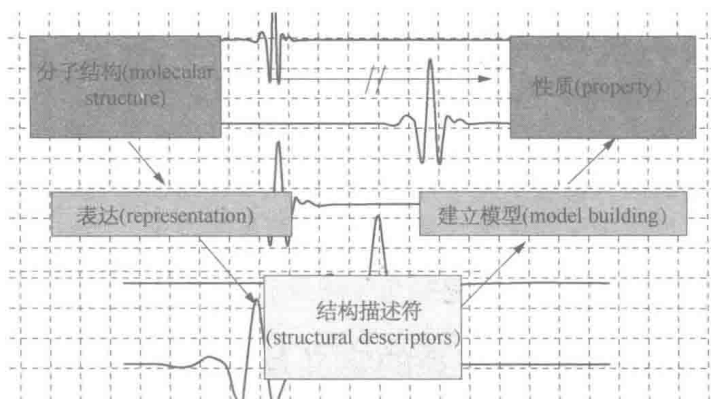


图 1-1 QSPR 模型与结构描述符

建立模型以后，还需要对建立的 QSPR 模型进行模型检验，经检验合格的 QSPR 模型才可以用于预测其他物质的性质。

定量结构性质关系的研究可以追溯到 20 世纪中期，有机反应相关性分析(Correlate Analysis of Organic Chemistry, CAOC)这一研究领域可以看作是定量结构性质关系研究的前身^[4]。Hansch 等^[6]于 20 世纪 60 年代提出的 Hansch 模型，用多元线性回归方法研究了化合物的性质与其结构之间的关系，这普遍被认为是经典 QSPR 研究的开始。随后 Free 和 Wilson 又提出了基团贡献法，被称为 Free-Wilson 基团贡献法。这种方法认为，任何基团对活性的贡献依赖于所处的位置，只要在特定的位置上，不管分子中是否还有其他基团存在，其作用不变，因此可以直接以分子结构作为自变量对物质的活性进行回归分析。之后的 QSPR 研究中又提出了拓扑指数方法和量子化学方法^[7-9]。拓扑指数方法和量子化学方法与 Hansch 方法的主要不同是根据分子结构计算被研究物质的结构描述符，能够更好地直接反映物质的结构信息。这些方法的结构描述符都是从被研究物质的结构计算得到，都可以归为二维定量结构性质关系(2D-QSPR)方法。

近年来随着化学计量学及相关量子化学、计算机和数学方法的发展，定量结构性质关系研究也得到了很大地发展，研究水平逐渐提高。QSPR 研究领域又提出了三维、四维、五维，甚至更高维的定量结构性质关系方法^[10-12]。

20 世纪 80 年代 Hopfinger 等将分子形状与定量结构性质关系结合起来，提出了第一种三维定量结构性质关系(3D-QSPR)方法——分子形状分析法(Molecular

Shape Analysis, MSA)。与 2D-QSPR 方法相比, 3D-QSPR 方法的物理化学意义更为明确, 所建立的模型也更准确。因此, 3D-QSPR 方法提出之后取得了很大的成功, 得到了化学计量学家的广泛关注, 建立了众多的方法, 包括: Apex-3D^[13]、距离几何方法 (Distance Geometry, DG)^[14]、CoMMA^[15]、EVA^[16]、MS-WHIM^[17]、比较分子力场分析 (Comparative Molecular Field Analysis, CoMFA)^[18, 19]、比较分子相似因子分析 (Comparative Molecular Similarity Indices Analysis, CoMSIA)^[20]等方法。其中最受欢迎、最成功, 应用最为广泛的就是 CoMFA 和 CoMSIA 方法。3D-QSPR 方法的问题之一是每个化合物只能用一种最低能量构象进行叠合。这个最低能量构象的选择, 特别是柔性结构化合物的构象选择会显著影响模型的准确性。此外, 实际上化合物不一定是以最低能量构象参加化学反应的。这使得 3D-QSPR 方法对于一些物质的预测准确性受到很大的限制。

为了克服 3D-QSPR 方法的问题, 1997 年, Hopfinger 提出了四维定量结构性质关系 (4D-QSPR) 的概念^[21]。4D-QSPR 的基本方法是基于 3D-QSPR 方法建立的。所谓的第四维实际指的是化合物分子各个构象、取向等的集合, 以考虑第四维来消除进行构象选择时带来的误差。与 3D-QSPR 不同的是, 4D-QSPR 中使用的构象并非最低能量构象, 而是最优构象。并且, 4D-QSPR 方法不采用传统的构象系统搜索, 而是在进行分子热力学计算后, 生成构象集成参数 (Conformational Ensemble Profile, CEP)。对与化合物最低能量构象的能量差在 $2\text{kcal} \cdot \text{mol}^{-1}$ (注: $1\text{cal} = 4.18\text{J}$) 以内的所有构象进行采样, 以得到的构象集合作为化合物的最优能量构象集合。在此过程中, 由于化合物的最优能量构象是以集合的形式表现的, 也就不再存在分子的排列问题。将所得每个化合物的最优构象放在与 CoMFA 类似的网格中, 计算形成网格占据描述符 (Grid Cell Occupancy Descriptors, GCOD), 然后根据 GCOD 进行偏最小二乘法和遗传算法建模计算后, 得到最终的 4D-QSPR 模型。

2002 年, Vendani 和 Dobler 提出了五维定量结构性质关系的概念^[22-24]。此方法的第四维和前面所讲的 4D-QSPR 一样, 也是构象的集合, 而第五维则是各种诱导契合的集合。诱导契合主要包括有六种类型: 线性适应、对立体场的适应、对静电场的适应、对氢键场的适应、对立体场方向的能量最低适应和对分子亲脂势能的适应。在分别生成构象集合与诱导契合集合后, 换算成相应的描述符, 再进行建模计算, 以得到模型。

目前, 还有研究者提出了更高维的定量结构性质关系方法, 但是相关理论和方法仍然需要完善。因此, 这里不对这些方法进行介绍。有兴趣的读者可以自行阅读相关文献。