

摘要

随着沿海地区经济的快速发展和人口的激增,使得近岸海域的污染日趋严重,为了有效地控制近岸海域的污染,就必须掌握其特征和规律,并在此基础上合理确定近海水环境容量。本文以天津市近岸海域为研究对象,完全依据环境监测数据,通过对复杂系统理论的研究和计算机软件的应用,成功实现了对近海水质的分析、预测和评价,并提出了计算近海水环境容量的新方法,为近海污染总量控制和环境管理提供了决策依据。

首先,通过对分形理论的研究,采用重标极差分析方法对近海各水环境监测点的 COD 和无机氮时间序列进行分析,计算其 Hurst 指数 H ,结果显示其 Hurst 指数 H 均在 0.80 左右,表明近海水污染物时间序列变化具有明显的分形特征;然后,采用多重分形维数谱方法对近海水污染物空间分布特征进行分析,通过多重分形矩方法计算近海 20 个水环境监测点的 COD 监测数据的多重分形维数谱函数 $f(\alpha)$,计算结果显示其 $\alpha(q)-f(\alpha)$ 曲线为一连续上凸曲线,表明近海水污染物的空间分布为连续多重分形分布。

其次,在证明近海 COD 时间序列具有分形特征的基础上,根据分形拼贴定理,由基于仿射变换的分形插值方法求取各历史时间阶段水质变化的迭代函数系,然后根据近海水质变化的年周期性,对上述求得的迭代函数系加权求和,得到预测年份水质变化的统计意义上的迭代函数系,从而建立近海水质的分形预测模型,最后应用随机迭代算法求得预测年份水质变化曲线的吸引子,对近海水质进行预测,预测结果显示平均预测误差为 24.4%。

再次,本文通过对非线性映射理论及技术进行深入学习,以 DPS 软件为工具,建立了近海水质综合评价的新模型。首先,通过非线性映射理论对高维水质监测数据进行非线性变换,得到二维映射值并将其描绘在平面坐标上;然后,根据映射点之间的位置关系进行分类,从而实现对近海水质的准确评价。

最后,在分析近海有机物空间分布具有连续多重分形分布特征的基础上,本文通过对分形曲面插值算法的研究,提出了近海有机物浓度估值的分形曲面插值算法模型,并在 MATLAB 上编制相关程序对近海有机物浓度进行空间插值,从而得到整个研究海域的有机物空间分布情况;然后,通过确定研究海域的最不利区域,依据质量守恒原理,最终确定了其在最不利条件下的水环境容量。

关键词: 近岸海域; 监测数据; 分形理论; 预测; 评价; 水环境容量

ABSTRACT

With the rapid economic development and the population booming in coastal zone, the coastal water quality was deteriorated. In order to control the coastal water pollution effectively, its characters and regulations must be mastered, based on which the coastal water environmental capacity could be calculated. In this dissertation, through the integrated using of complex system theory and computer software, the analyzation, prediction and assessment of the water quality in Tianjin coastal marine were studied totally based on the environmental monitoring data. Furthermore, a new calculation method for coastal water environmental capacity was proposed, which could supply the decision support for coastal pollution total amount control and environmental management.

Firstly, based on fractional theory, the rescaled range analysis was used to compute the Hurst exponents H of the COD and inorganic nitrogen time series. The results showed that the Hurst exponents of the two water quality indexes were around 0.80, it proved that the coastal water quality pollution presents fractional characteristics. Then, based on multi-fractal theory, a typical multi-fractal method, which called the method of moments, was applied to compute the multi-fractal dimension spectrum function $f(\alpha)$ of the 20 monitoring points. From the results, we could found the curve of $\alpha(q) - f(\alpha)$ was a continuous curve. The results indicated that the spatial distributing mode of coastal water pollution was a continuous multi-fractal distribution.

Secondly, according to the fractional collage theory, the fractal interpolation which based on affine transform was used to find the iterated function systems of historical water quality. The weighted summation method was used to find the iterated function system of predicting period according to the above iterated function systems, which was based on periodicity. Then the fractal predicting model was established according to the iterated function system of predicting period. Finally, the random iterated algorithm was used to find the attractor of each predicting periods which could provide the predicting data according to the time values. The predicting results showed that the average predicting error was 24.4%.

Thirdly, a new assessment model for coastal water quality was proposed through the study on the nonlinear mapping theory and DPS software. Firstly, the high-dimension water quality data were mapped to two-dimension plane by nonlinear mapping method. Secondly, the water quality assessment model was established according to location of the mapping points. So the water quality assessment could be simplified and improved.

At last, based on the results obtained in the chapter 2 and fractal surface interpolated algorithm, a new method for estimation of organic substance concentration in coastal marine was proposed. The MATLAB program of the fractal surface interpolation algorithm was made to estimate the concentration of the organic substance in coastal marine. Then, the COD concentration of organic substance in the whole area was gotten. Finally, the water environmental capacity under the worst condition was calculated according to the conservation of mass.

KEY WORDS: coastal marine; monitoring data; fractional theory; prediction; assessment; water environmental capacity

第一章 前言

1.1 课题的提出及研究意义

海岸带地区是自然界的特殊区域,一方面,这里有着丰富且便于开发的能源、生物和矿产等资源,以及适宜生产、生活的自然环境,具有对当地和区域乃至全球都十分重要的“三生”功能(生活、生产和生态功能);另一方面,这里是海陆交互地带,不但受到海洋、陆地、大气等自然环境的综合影响,更受到人类活动剧烈、频繁、持久的作用。但是,近年来,由于海岸带人口的过度集中,经济的高速发展以及人类对近海资源掠夺式的开发、利用,加之全球气候变化所带来的海平面上升,使得本来就相当敏感和脆弱的海岸带环境承受着日益严重的压力,出现了近岸海域富营养化、生物多样性丧失、近海湿地的减少等多种环境问题,而这些问题反过来都会影响海岸带的人类活动,会大大削弱人类在这一区域追求经济发展所赖以生存的物质基础。世界上许多国家都曾报道过海湾及近岸海域的污染情况^[1-3],包括近海水体中营养类物质的逐渐增多和营养物质比例的变化,比如波罗的海、北海、黑海、亚得里亚海、荷兰海、日本海、泰国湾以及印度洋等,可以说遍布整个世界的近岸海域。但是,在过去 30 年里,只有少数的发达国家开始了海洋环境的保护工作,并且这些工作也有很大的局限性。但是,海洋环境不仅没能改善,还在进一步恶化。不过,人们对于海洋污染的观念有了很大的变化,并开始重视海洋污染,特别是近海污染的问题^[4]。

近年来,随着我国沿海地区人口的增长和经济的发展,大量的工业废水和生活污水直接或间接的排入近岸海域,导致沿海水体受到不同程度的污染,由此导致的经济损失十分可观。据 2003 年的数据统计^[5],我国全海域未达到清洁海域水质标准的面积约为 14.2 万平方公里,比上年减少了 3.2 万平方公里,总体污染趋势有所减缓。其中:较清洁海域面积约为 8.0 万平方公里,减少 3.1 万平方公里;轻度污染海域面积约为 2.2 万平方公里,增加 0.3 万平方公里;中度污染海域面积约为 1.5 万平方公里,减少 0.3 万平方公里;严重污染海域面积约为 2.5 万平方公里,减少 0.1 万平方公里。我国近岸海域污染依然严重,但中度污染和严重污染海域的面积均比上年有所减少;近海大部分海域符合清洁海域水质标准;远海海域水质继续保持良好。海水中的主要污染物依然是无机氮、活性磷酸

盐和铅。严重污染海域主要分布在鸭绿江口、辽东湾、渤海湾、长江口、杭州湾、珠江口和部分大中城市近岸局部水域。其中，渤海污染范围减小，未达到清洁海域水质标准的面积由 2002 年的 3.2 万平方公里降至 2.1 万平方公里，占渤海总面积的 27.3%。污染范围减小的海域主要包括莱州湾和辽东湾。主要污染物是无机氮、活性磷酸盐和铅。天津近岸海域中度和严重污染海域范围较小，主要集中在北塘近岸水域，轻度污染海域主要集中在渤海湾近岸海域。主要污染物为无机氮和铅。

本文研究的海域为天津市近岸海域。天津市地处渤海湾西岸、海河水系与永定新河水系的尾间，东临渤海湾，北依燕山，西接首都北京，南北与河北省接壤，地处华北平原的东北部；是海河五大支流南运河、子牙河、大清河、永定河和北运河的汇合处和入海口。天津市海岸线长约 153 km，滩涂面积约 370 km²，管辖海域面积约 3000 km²。2003 年的监测结果表明：天津市大部分近岸海域海水环境质量符合较清洁海域水质标准，永定新河和海河河口之间 5 米等深线以内海域，属于轻度污染海域，局部海域属于中度污染海域，主要污染物为无机氮、铅和铜；永定新河和海河河口区污染严重，属于严重污染海域，污染物为硫化物和活性磷酸盐。

天津市近岸海域一直是天津市发展的经济中心，国家“十一五”规划确定了重点发展天津市滨海新区的战略，天津市近岸海域在经济发展方面的功能更得以体现，近岸海域水质的优劣是其更好发挥作用的前提和基础。但是近年来由于地理位置和自然条件的影响、污染负荷的日益加重，其水质有加快恶化的趋势，也持续出现了水体富营养化等环境问题，给天津市近岸海域的环境带来更大的压力，尽管采取了大量的技术手段，其污染问题仍然呈现日益严重的趋势^[6]；2003 年中度和重度污染海域面积比 2002 年均有所增加，2004 年的赤潮次数由 2003 年的 1 次增加至 4 次，赤潮发生面积由 2003 年的 100 km² 增加至 3230 km²。而经过环境管理实践发现，只有进行容量总量控制才能有效改善水环境质量，所以国家环保总局在地表水环境容量及总量控制的工作基础上，准备进行近海水环境容量和总量控制的研究。

本课题正是针对近海水环境容量总量控制的重要性及其研究的先导性开展研究。力求借助天津市滨海新区辐射周边的区极效应和示范效应，对有关揭示区域尺度的近海水环境污染规律的理论和方法进行整合与创新，制定一套符合地区特点和中国国情的海岸带污染规律研究模式、水环境容量与总量控制的理论方法体系和陆海协调发展的系统对策，指导研究区域内经济格局调整、污染治理策略的制定，促进天津市及周边区域的社会、经济和环境可持续发展。通过本研究，制定一套科学的污染控制方案，实现最大限度的利用近海水域的水环境容量；同

时本研究采用的方法具有普遍性意义,可以推广到其它地区近岸海域水环境管理的研究和应用中。

1.2 国内外研究现状

无论在发达地区还是发展中地区,海洋,尤其是近岸海域正遭受着人为活动带来的污染。如何有效的控制近海的污染,使其达到功能区划提出的水质目标,是世界各国环境管理和研究机构的工作焦点。随着环境管理工作的摸索,大家一致认为必须与浓度管理相结合的基础上,实行容量总量控制,才能真正实现近海水质的改善。所以,如何科学的确定近岸海域的水环境容量,成为大家研究的重心。国内外学者对近海水质预测、评价和水环境容量计算这几个方面都做了大量的研究工作。

国外对近岸海域污染问题的研究开始较早,很多专家和学者对于近海水环境污染问题进行了大量的理论研究工作。James 在对近岸海域的水质、生态系统和污染物扩散模拟研究这一文指出,由于最近 10 年来包含多种学科的模式逐渐增多,并且计算机的并行计算能力达到了一定高度,使得多学科模型进入了实用化阶段,能够用于预测和模拟近海水质^[7]。Nina Gjerde Winther 等人通过三个不同垂直坐标设计的实验验证了适用于北海和斯卡格拉克海峡之间循环模拟的 HYCOM 模型同样适用于海岸带的模拟^[8]。Minna Kuusisto 等人采用三维水质和水流模型模拟芬兰湾中的浮游植物数量和溶解性营养物质的浓度^[9]。Qiuwen Chen 等人采用 Delft-WAQ 模块计算荷兰海湾水体中光照因子、营养物质等非生物条件,利用模糊元胞自动机模拟了荷兰海湾水体中藻类生物的数量^[10]。Dong S Ko 等人利用卫星数据建立了一个对海洋海平面变化、温度、盐度等因子的实时短期预测系统,并对实例进行了预测研究^[11]。Michael D Krom 等人通过测定以色列地中海海岸线受汞污染海域鱼的体重和含汞量之间的关系,拟合出一条关系曲线,并依据此关系曲线计算出了海域的环境容量^[12]。

在国内,陶建华等人用物理模型实验和数学模型计算相结合的方法,研究了均匀缓坡岸滩上,规则波和不规则波浪作用下形成的沿岸流及其对岸边排放污染物输移扩散的影响^[13]。沈永明等人在分析海洋环境中各种物理、化学和生物过程基础上,建立了统一考虑三种过程综合作用的近岸海域污染物迁移转化的三维水质动力学模型^[14]。牛志广等人通过确定近海水质数据符合 log-logistic 的概率密度函数,运用神经网络法和傅立叶级数法对近海水质进行了组合预测,提出一种完全依据环境监测数据的近海水质预测模型^[15]。娄安刚等人通过对胶州湾水质现状的数值模拟,预测了胶州市城市污水处理厂建成后,在正常运行和事故情况下

大沽河口邻近海域的水质^[16]。栗苏文等人利用 Delft3D 数学模型对大鹏湾现状和未来条件下的水质进行了数值模拟, 并采用模型试算法估算出大鹏湾的水环境容量^[17]。郭森等人选用多源线性迭加的方法确定纳污海域水环境容量计算的几个重要条件, 并以此计算纳污海域水环境容量更加便利和有效^[18]。

在近海水质评价方面, 常用的水环境评价指数包括单因子评价指数, 多因子评价指数(包括均值型、加权型和内梅罗指数三种)^[19]。近年来, 随着模糊数学、灰色系统、物元分析、遗传算法、人工神经网络等学科的兴起与发展, 国外学者提出了新的水环境质量评价方法体系, 而国内研究的发展亦大致呈现这一趋势^[20]。谢宏斌采用权重贡献率分析法优选因子和确定因子权值的基础上, 应用 B-P 神经网络对南湖水质富营养化进行综合评价^[21]。王晓鹏等选择湟水流域内空间跨度大及污染差异分布广泛的水质断面作为研究评价对象, 建立河流水质评价的多级模糊模式识别模型, 并以级别特征值方法对湟水流域的主要水质断面就污染程度进行定量化评价和分级^[22]。罗定贵等针对模糊综合评价法的评价结果具有模糊性, 简单依据最大隶属度原则确定评价结果存在的不足, 提出了模糊综合评价与模糊聚类分析相结合进行地下水水质评价与分区的复合模型, 并将模型应用于抚州市地下水环境质量评价与分区。而且, 针对 BP 神经网络结构和初始权值确定的人为性, 应用 MATLAB6.5 神经网络工具箱中径向基函数网络对研究区域的地表水质进行了评价^[23-24]。

由上述可知, 大多数学者在研究近海污染的规律时, 多以机理性研究为主, 即以二维或三维水动力学方程为基础, 配合物理、化学、生物等模型, 并将这些模型组合在一起, 力图模拟出各种污染物的时空变化特征。但是近海水质数据变化剧烈, 规律性不强, 海洋污染的复杂性使得这些工作举步维艰, 不太适合于环境管理的实际, 而且模拟模型所需基本数据很难获得。所以, 随着各种基础学科和计算机软、硬件的发展和环境监测设施的逐步完善, 环境监测数据的可得性较好, 近海水环境越来越多的研究可借助于交叉学科的优势, 可以应用机理性研究的部分成果, 依据环境监测数据开发出科学性和实用性并重的研究方法。

1.3 本文主要研究内容和方法

本文主要探讨了复杂系统理论用于近海水污染物时空变化特征的分析、水质的预测、水质综合评价、有机物浓度的估值和近海水环境容量计算方面的研究, 通过实例应用, 验证了复杂系统理论用于近海水环境研究方面的实用性和科学性。

(1) 首先, 采用分形理论中的重标极差分析方法(亦称 R/S 法)对 COD 等

水质参数的时间序列进行分析,分析水质时间序列变化的单重分形特征;然后,采用多重分形理论中的多重分形维数谱方法对天津市近海 20 个水环境监测点的 COD 监测数据进行分析,分析 COD 空间分布的连续多重分形分布特征。

(2) 其次,针对我国近海水质污染呈现的年周期性和自相似性特点,在分析近海水质时间序列具有分形特征的基础上,通过对分形理论的研究,充分利用分形时间序列具有大小时间尺度规律可以互用的特点,提出了基于分形拼贴定理和分形插值函数迭代的近海水质污染变化的分形预测方法。

(3) 再次,通过对非线性映射理论及技术进行研究,借助 DPS 数据处理系统软件,将高维水质信息通过非线性映射过程转换为二维水质信息,根据映射点之间的位置相互关系,建立近海水质综合评价的新模型,并实现对近海高维水质信息的可视化。

(4) 最后,在分析近海有机物空间分布具有连续多重分形分布特征的基础上,通过对分形曲面插值算法的研究,提出了一种完全基于环境监测数据的近海有机物空间变异特性和浓度估值研究的新方法,并在此基础上,依据质量守恒原理,通过确定天津市近岸海域的最不利区域,成功的实现了近海水环境容量的计算。

通过以上研究工作,本文形成了一套完整的基于复杂系统理论的近海水质分析、预测、水质综合评价、有机物浓度估值和近海水环境容量计算的方法体系,能够为近海水环境管理提供全方位决策支持,具有一定的理论意义和实用价值。

第二章 近海水污染物的时空变化特征分析

2.1 概述

对于近岸海域的水环境管理来讲,由于管辖区域和监测难度较大,不可能设置高密度的水环境监测点和进行高频率的海上水质监测,也就造成对近海水环境在时间和空间上的变化存在盲区。建立近海水环境功能区,并在每个功能区当中设置合理数目的水环境监测点,可以在一定程度上解决这个矛盾。进行近岸海域环境功能区划的目的是为了优化海洋环境和资源的配置,解决海洋资源开发利用上的矛盾,建立近岸海域环境功能利用的正常秩序,对海域实行按水质分类管理,这样也就可以根据各个功能区的人为影响强度和重要程度合理设置水环境监测点,监测各个环境功能区的水质变化,为近海水环境管理提供数据依据。根据水环境监测点上某个水质指标的变化,可以体现出该水环境监测点所代表的功能区的水质变化情况,但是不能反映近岸海域各个不同的子区域之间的水质空间变化关系。事实上,由于近岸海域受到各种作用的影响,不同水环境监测点的水质指标呈现一定的关联性,在长期变化过程中分为不同的变化区域,体现相同或相近的变化趋势,这将为近海的区域化管理提供决策支持;同时,由于某些水环境监测点水质受局部因素的影响,而导致水质变化在空间上又存在变异特征。

为了对近海水环境进行有效的管理,这就需要了解其污染的时空变化规律,为近海水污染的防治提供理论依据,如何利用少数的几个水环境监测点的水质监测数据分析近海水污染变化的规律性是一大难题。根据以往的研究^[25-26],通过对所有水期和不同水期水质监测数据的检验,已经排除水污染物正态分布的可能,采用灰色系统理论和统计学方法分析表明,在不同的陆源入海河流的影响下,天津市近岸海域的20个水环境监测点按照COD的变化规律被分成了北塘区域、大沽区域、以及独流减河区域。在上述分析近海水环境监测点水质空间分布关联关系的基础上,本文尝试采用分形理论对近海水污染物时间和空间上的变化规律和变异特征进行分析,力图确定近岸海域水质污染的分形变化特性,为近海水质的预测和水环境的管理提供理论依据。

本文将天津市近岸海域COD和无机氮(IN)监测数据为例,采用分形理论中的重标极差分析方法(亦称R/S方法)对单个水环境监测点的水质时间序列

进行分析,采用多重分形理论中的多重分形维数谱方法对 20 个水环境监测点水质监测数据的空间变异特征进行分析。

2.2 分形理论

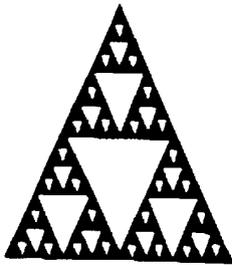
2.2.1 分形理论介绍

分形理论是非线性科学研究中十分活跃的一个分支,它的研究对象是自然界和非线性系统中出现的不光滑和不规则的几何形体^[27]。分形理论的基本思想起源于 20 世纪初,发展于 20 纪 60 年代后,创立分形理论的代表人物为美国数学家 B.B.Mandelbrot,其数学基础是分形几何。1960 年, B.B.Mandelbrot 发现自己研究的经济模式中高低收入的分布图与利塔卧经济中心大厦黑板上的棉花价格变动图是一样的,之后,他转向为研究棉价变化。他采用一种特殊的变换尺度的方式考察棉价变化的内在序列,惊人的发现了:尽管棉价的每一次变化是随机的,但价格的日变化和月变化的曲线竟然完全一致,也就是说,在大量的无序的数据里存在一种出乎意料的有序。随后, B.B.Mandelbrot 又对“海岸线到底有多长”这个问题进行了研究,他发现在不同的尺度研究下,任何一段海岸线都是无限长的。虽然一条曲折的海岸线不能精确测量,但它却有某种特征量,就是分形所揭示的分形维数,可以对分形对象内部的不均匀性、层次结构性的整体数量特征进行刻画。1967 年他在《科学》杂志上发表了一篇题为“英国的海岸线有多长?统计自相似性与分数维数”的文章,这通常被认为是“分形”学科诞生的标志,他在随后两本著作《自然界的分形几何学》和《分形:形状、机遇与维数》中第一次提出了 Fractal 这个英文词^[28-29],其原意是“不规则的”、“分数的”、“支离破碎的”,并阐述分形的基本思想,即分形研究的对象是具有自相似性的无序系统,其维数的变化是连续的。应该认识到的是,自相似性是自然界中一个普遍的规律,小到树叶的叶脉,大到宇宙天体,自相似性普遍存在于物质系统的多个层次上。这位被科学界尊为“分形之父”的 B.B.Mandelbrot 的最大贡献在于提出“物体或几何图形的维数的变化可以是连续的”这一惊人的论断,即其维数可以不是整数,分形理论也就是用分数维数来描述自然界中千奇百怪的事物取得了重大成功,这也正是与采用整数维数对自然界中的事物进行描述的欧氏几何的最基本的区别。分形理论借助相似性原理洞察隐藏于混乱现象中的精细结构,为人们从局部认识整体、从有限认识无限提供了新的方法论,为不同学科发现的规律提供了崭新的语言和定量的描述,为现代科学提供了新的思想方法^[30]。

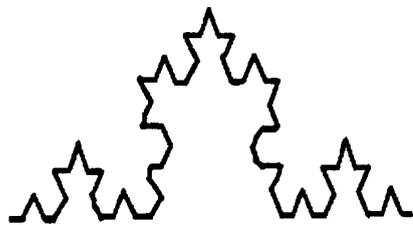
分形具有两个基本的性质:自相似性和标度不变性。自相似性是指局部是整

体成比例缩小的性质。形象地说，就是当用不同倍数照相机拍摄研究对象时，无论放大倍数如何改变，看到的照片都是相似的（统计意义），而从相片上也无法判断所用相机的倍数，即标度不变性。严格按一定的数学方法生成的许多经典的分形具有严格的自相似性，称之为有规分形（见图 2-1）。而一般情况下的分形是无规分形，即自相似性并不是严格的，只是统计意义下的自相似性，其局部经放大或缩小可能得到与整体完全不同的表现形式，但表征自相似结构或系统的定量参数如分形维数并不因此变化，也就是说，分形体内任何一个相对独立的部分，在一定程度上都是整体的再现和缩影。目前对分形还没有严格的数学定义，只能给出描述性的定义。粗略地说，分形是对没有特征长度（所谓特征长度是指所考虑的集合对象所含有的各种长度的代表者，例如一个球，可用它的半径作为它的特征长度），但是具有一定意义下的自相似图形和结构的总称。原则地说：分形是一些简单空间上，如 R^2 、 C^2 上的一些“复杂”的点的集合，这种集合具有某种特殊的性质，首先它是所在空间的紧子集，并且具有下列典型的几何性质的集合 F [31]：

- (1) F 具有精细的结构，即在任意小的比例尺度内包含整体。
- (2) F 是不规则的，以至于不能用传统的几何语言来描述。
- (3) F 通常具有某种自相似性，近似的或统计意义下的自相似性。
- (4) F 在某种方式下定义的“分维数”通常大于 F 的拓扑维数。



(a) Sierpinski 三角形



(b) Koch 曲线

图 2-1 两种简单、规整的分形图形

2.2.2 分形理论基础

2.2.2.1 分形空间

设 (X, d) 是完备度量空间，用 $\mathfrak{R}(X)$ 表示由 X 的非空紧子集的全体组成的空间。令点 $x \in \mathfrak{R}(X)$ ，集合 $B \in \mathfrak{R}(X)$ ，定义：

$$d(x, B) = \inf\{d(x, y) : y \in B\} = \min\{d(x, y) : y \in B\} \quad (2-1)$$

为 x 到集合 B 的距离, \inf 表示取下确界。由于集合 $B \in \mathfrak{R}(X)$ 是非空紧集, 可以证明实数集 $\{d(x, y) : y \in B\}$ 包含一个最小值, 也就是说, 一定存在一个 $y \in B$ 使 $d(x, y) = d(x, B)$ 。令 $A, B \in \mathfrak{R}(X)$, 定义:

$$d(A, B) = \sup\{d(x, B), x \in A\} = \max\{d(x, B), x \in A\} \quad (2-2)$$

为集合 A 到集合 B 的距离, \sup 表示取上确界, 定义:

$$h(A, B) = d(A, B) \vee d(B, A) \quad (2-3)$$

为 $\mathfrak{R}(X)$ 中两个点 A 和 B 之间的 Hausdorff 距离。符号 \vee 表示两个实物中取较大者。

我们称 $(\mathfrak{R}(X), h(d))$ 为分形空间, 可以证明, 分形空间具有完备性。

2.2.2.2 分形维数

分形维数是分形理论的重要内容, 经典的拓扑维数只能取非负整数, 而分形维数可以是任意非负实数。

分形理论的研究对象是由非线性系统产生的不光滑和不可微的几何形体, 分形结构的明显特点是局部与整体形态间的相似性, 对应的特征指标参数是分形维数。分形维数不是通常意义下欧氏空间维数的简单补充, 而是赋予了很多暂新的内涵: 分形维数不仅量化了分形结构自相似性质, 同时也可作为区分混沌和其它运动形式的主要依据之一。分形维数可以是分数, 也可以是整数, 并有多种计算方法。常见的分形维数有: Hausdorff 维数 (盒子维数)、相似维数、计盒维数、信息维数、关联维数以及广义维数等。下面给出最常用的 Hausdorff 维数和计盒维数的定义。

(1) Hausdorff 维数

设 F 为 \mathbb{R}^n (n 维欧氏空间) 中的任意子集, S 为一非负实数, 对任意给定的 $\delta > 0$, 设:

$$H_\delta^S(F) = \inf \left\{ \sum_{i=1}^{\infty} |U_i| : \{U_i\} \text{ 为 } F \text{ 的 } \delta \text{ 覆盖} \right\} \quad (2-4)$$

考虑 F 的所有直径不超过 δ 的覆盖, 并使式 (2-4) 右端出现的和式达到最小, 当 δ 减小时, $H_\delta^S(F)$ 随之增加, 当 $\delta \rightarrow 0$ 时, 记极限值为:

$$H^S(F) = \lim_{\delta \rightarrow 0} H_\delta^S(F) \quad (2-5)$$

称 $H^S(F)$ 为 F 的 S 维 Hausdorff 测度, 此测度是对长度、面积和体积概念的推广, 更详细的推导可以说明, 存在 S 的一个临界值, 使得 $H_\delta^S(F)$ 在这个值处从无穷大跃变到 0, 这个临界值就称为 F 的 Hausdorff 维数, 记为 $d_H(F)$ 。Hausdorff 维数定量描述了一个点集规则或不规则的程度, 同时其整数部分反映出图形的空

间规模, 对动力系统来讲, Hausdorff 维数大体上表示独立变量的数量。

(2) 计盒维数

设 F 为 R^n 上任意非空的有界子集, $N_\delta(F)$ 是直径最大为 δ , 可以覆盖 F 的最少个数, 则 F 的下、上盒维数分别定义为:

$$\underline{Dim}_B F = \lim_{\delta \rightarrow 0} \frac{\log N_\delta(F)}{-\log \delta} \quad (2-6)$$

$$\overline{Dim}_B F = \lim_{\delta \rightarrow 0} \frac{\log N_\delta(F)}{-\log \delta} \quad (2-7)$$

如果这两个值相等, 则称这个共同的值为 F 的计盒维数, 记为:

$$Dim_B(F) = \lim_{\delta \rightarrow 0} \frac{\log N_\delta(F)}{-\log \delta} \quad (2-8)$$

2.2.2.3 压缩映射

设 f 是距离空间上的一个自身映射, 即 $f: X \rightarrow X$, 如果存在一个常数 $0 \leq s \leq 1$, 使得对一切 $x, y \in X$, 都有:

$$d(f(x), f(y)) \leq sd(x, y) \quad (2-9)$$

则称 f 为压缩映射, s 为 f 的压缩因子。

2.2.2.4 仿射变换

(1) 在 R^1 上的仿射变换 (Affine transformation) 是形如:

$$f(x) = ax + b \quad (2-10)$$

的变换。其中 a 、 b 是实常数, a 是伸缩因子, b 是平移因子。

(2) 在 R^2 上的仿射变换 $w: R^2 \rightarrow R^2$ 形式为:

$$w(x_1, x_2) = (ax_1 + bx_2 + e, cx_1 + dx_2 + f) \quad (2-11)$$

称为二维仿射变换。其中 a 、 b 、 c 、 d 、 e 、 f 是实数。

2.3 近海水污染物时间序列的分形特征研究

2.3.1 基础数据

本文以天津市近岸海域为研究对象, 该海域共设有 4 个水环境功能区 20 个水环境监测点 (水环境监测点位置见图 2-2)。本文选择其中的 8 个水环境监测点 (1、3、5、6、7、9、10、16) 为例进行研究, 研究采用的原始数据来自 1996 年至 2003 年每年枯水期 (5 月)、丰水期 (8 月)、平水期 (10 月) 3 个水期的 COD 和 IN 监测数据, 共有 384 ($8 \times 24 \times 2$) 个数据样本。由于监测时间为每年的枯、丰、平 3 个水期, 单个水环境监测点的 COD 和 IN 时间序列不是严格的

等间隔时间序列，但是监测值能够代表 3 个水期水质的变化趋势，所以在此将 COD 和 IN 时间序列看为等间隔时间序列，顺次分别用 $\{x_i\}$, $i=1,2,\dots,24$; $\{y_i\}$, $i=1,2,\dots,24$ 表示。其中水环境监测点 6 的 COD 和 IN 监测数据见图 2-3，可以看出 COD 和 IN 监测值变化范围比较大，表面上看监测值随时间的变化不规则、遵循随机游动模型。

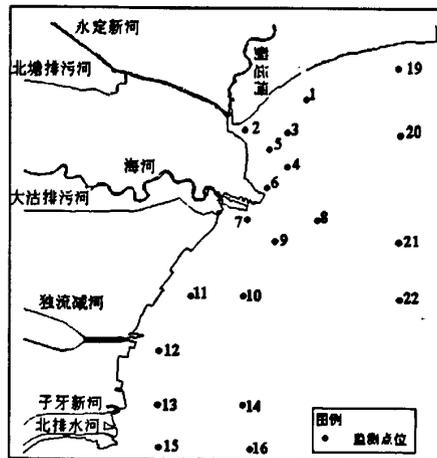


图 2-2 天津市近海水环境监测点位置示意

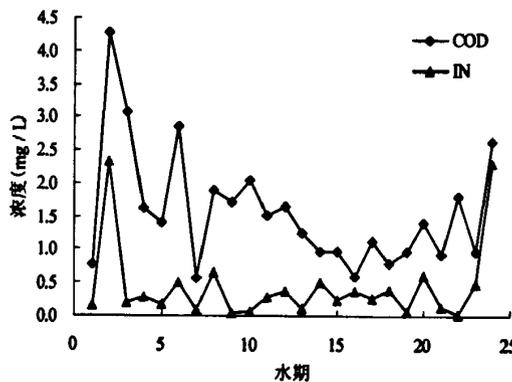


图 2-3 水环境监测点 6 的 COD 和 IN 监测数据

由于近海水质的变化表现出复杂的非线性特征，使研究的 COD 和 IN 时间序列数据前后关系很不明显，单个水环境监测点的单个水质指标的变化规律变得难以把握，这使得以往通常采用的需要大样本数据的时间序列分析方法不能很好的分析其变化规律。而近海水质监测数据的特点正好符合 R/S 分析方法可以大小时间尺度规律互用的基本思想，所以本文尝试采用 R/S 分析方法对近海水污染物时间序列的分形特征进行研究，为进一步应用基于分形理论的水质预测方法提供

理论基础。

2.3.2 基于R/S分析方法的近海水污染物分形特性计算

R/S 分析方法是目前最广泛也最成熟的分形统计方法之一，在 R/S 分析中，不要求序列是高斯过程，该方法最初由水文专家 Hurst 在研究尼罗河水坝工程时提出^[32]，是基于自仿射分形衍生而来的对随机时间序列进行分析的一种统计方法，该方法可用于时间序列的分形特征分析，在分形理论中有着重要的作用，是现代非线性科学中一个十分重要的研究方法，该方法运用于自然科学领域^[33]和社会科学领域^[34]都取得了重大成果。

(1) Hurst 指数 H 的计算

Hurst 指数是分形理论中的一个重要参数，它可以用来考察时间序列的序列相关性以及描述对象是否符合分形分布。利用 R/S 分析方法可以估计 Hurst 指数 H ，从而研究时间序列的分形特征，R/S 分析方法的基本思想是，改变所研究对象的时间尺度的大小，研究其统计特性变化规律，从而可以将小尺度的规律用于大的时间尺度范围，或者将大尺度的规律用于小的时间尺度范围，该方法在分形理论中具有广泛的应用。本文利用 R/S 分析方法计算天津市近岸海域 8 个水环境监测点的 COD 和 IN 时间序列的 Hurst 指数 H ，以水环境监测点 6 为例，其基本步骤如下^[35]：

首先将水环境监测点 6 的 COD 时间序列 $\{x_i\}$ ， $i=1,2,\dots,24$ ，分成 a 个长度为 n ($2 \leq n \leq 12$) 的相邻子区间，且 $an=24$ 。记第 k 个子区间为 I_k ($k=1,2,\dots,a$)， I_k 中的第 j 项记为 $x_{j,k}$ ($j=1,2,\dots,n$)。令

$$e_k = \frac{1}{n} \sum_{j=1}^n x_{j,k} \quad (k=1,2,\dots,a) \quad (2-12)$$

$$y_{j,k} = \sum_{i=1}^j (x_{i,k} - e_k) \quad (j=1,2,\dots,n) \quad (2-13)$$

$$R_k = \max_{1 \leq j \leq n} \{y_{j,k}\} - \min_{1 \leq j \leq n} \{y_{j,k}\} \quad (2-14)$$

$$S_k = \left[\frac{1}{n-1} \sum_{j=1}^n (x_{j,k} - e_k)^2 \right]^{\frac{1}{2}} \quad (2-15)$$

式中： e_k 和 S_k 分别为 I_k 中 COD 的均值和标准差，mg/L； $y_{j,k}$ 为 I_k 中 COD 的累积离差，mg/L； R_k 为 I_k 中 COD 的极差，mg/L。对每个子区间计算 R_k/S_k ，可得 a 个值，将它们的平均值记为 $(R/S)_n$ ，即：

$$(R/S)_n = \frac{1}{a} \sum_{k=1}^a (R_k/S_k) \quad (2-16)$$

为了比较不同的时间序列，Hurst 对式 (2-16) 进行了修正，得到了如下关系式：

$$\ln(R/S)_n = H \ln n + \ln b \quad (2-17)$$

式中： b 为常数； H 为 Hurst 指数，且 $0 \leq H \leq 1$ 。对于不同的 n ，可以得到不同的 $(R/S)_n$ ，以它们为式 (2-17) 的观测值，利用最小二乘法进行拟合，便可以得到式 (2-17) 的直线回归方程，则直线的斜率便是 COD 时间序列 $\{x_t\}$ 的 Hurst 指数 H 值。

表 2-1 水环境监测点 6 的 COD 和 IN 时间序列的 $(R/S)_n$ 值计算结果

n	$(R/S)_n$		n	$(R/S)_n$	
	COD	IN		COD	IN
2	0.7071	0.7071	3	1.1098	1.1270
4	1.3815	1.3618	6	2.0113	1.7751
8	2.7145	2.3752	12	3.4053	3.0422

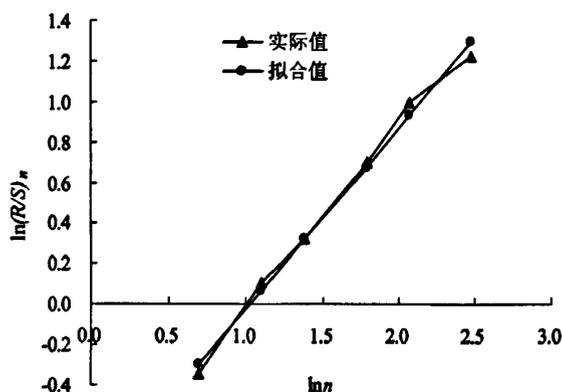


图 2-4 水环境监测点 6 的 COD 时间序列的 R/S 分析图

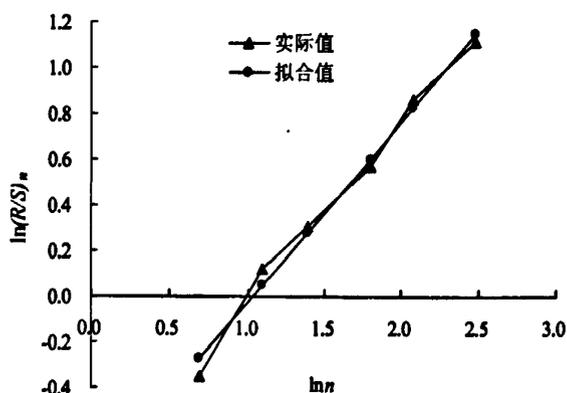


图 2-5 水环境监测点 6 的 IN 时间序列的 R/S 分析图

按照上述 R/S 分析方法的基本步骤, 计算 $n=2, 3, 4, 6, 8, 12$ 时, 水环境监测点 6 的 COD 和 IN 时间序列的 $(R/S)_n$, 计算结果见表 2-1, 采用最小二乘法对式 (2-17) 进行拟合, 拟合直线见图 2-4 和图 2-5。水环境监测点 6 的 COD 和 IN 时间序列的 Hurst 指数 H 分别为 0.8868 和 0.7962。

时间序列的 Hurst 指数 H 介于 0~1 之间, 其值的大小能够表明时间序列的特性。根据统计学分析有以下几点特性^[36]: 当 $H=0.5$ 时, 这标志时间序列为标准的随机游走, 不同时间的值是不相关的, 即现在不会影响将来。当 $0 \leq H < 0.5$ 时, 时间序列具有反持续性, 常被称为“均值回复”, 这种反持续性行为的强度依赖于 H 的大小, H 越小, 反持续性就越强。即如果某一时刻, 序列向上(下)变化, 则下时刻它很可能反转向下(上)变化。当 $0.5 < H \leq 1$ 时, 时间序列具有持续性特征, 常被称为分形时间序列, 即如果某一时刻, 序列向上(下)变化, 则下一时刻它很可能继续向上(下)变化, H 越接近于 1, 这种状态持续性越强。计算结果显示水环境监测点 6 的 COD 和 IN 时间序列的 H 值均大于 0.5, 接近于 1, 表明 COD 和 IN 随时间的变化具有较强的状态持续性, 具有明显的分形特征。

(2) 记忆长度 V_n 的计算

时间序列长期记忆的长度, 即某一时刻的状态值对将来的影响经过多长时间才会减少到度量不到的程度, 可以通过下式计算的统计量 V_n 的值来确定:

$$V_n = \frac{(R/S)_n}{\sqrt{n}} \quad (2-18)$$

由该统计量的定义知, 当序列呈现持续性 ($H > 0.5$) 时, V_n 单调上升; 当序列呈现随机游走或反持续性 ($H = 0.5$ 或 $H < 0.5$) 时, V_n 近似为常数或单调下降。 V_n 由上升转为常数或下降的分界点, 即为序列长期记忆的消失点。以天津市近岸海域水环境监测点 6 为例, 按照式 (2-18) 计算水环境监测点 6 的水质指标 COD 和 IN 时间序列的 V_n , 计算结果见表 2-2。计算结果显示水环境监测点 6 的 COD 和 IN 时间序列的 V_n 值随着 n 的增大一直是递增的, 这说明在统计年份内, 初期水质对以后水质的影响一直没有消失, 表明 COD 和 IN 随时间的变化具有长期记忆性的特点。

表 2-2 水环境监测点 6 的 COD 和 IN 时间序列的 V_n 值计算结果

n	V_n		n	V_n	
	COD	IN		COD	IN
2	0.5000	0.5000	3	0.6407	0.6507
4	0.6908	0.6809	6	0.8211	0.7247
8	0.9597	0.8398	12	0.9830	0.8782

(3) 关联度 C 的计算

时间序列上的一个观测值对后面观测值的影响度跟 H 有关, 可以用关联度 C 来度量:

$$C = 2^{2H-1} - 1 \quad (2-19)$$

根据关联度 C 的定义, 当 $H = 0.5$ 时, $C = 0$, 该时间序列不同时间的值是不相关的; 当 $0 \leq H < 0.5$ 时, $-0.5 \leq C < 0$, 表明一个观测值对后面观测值有负方向的影响; 当 $0.5 < H \leq 1$ 时, $0 < C \leq 1$, 表明一个观测值对后面观测值有正方向的影响。时间序列关联度 C 的绝对值越大, 表明一个观测值对后面观测值的影响度越大。

以天津市近岸海域水环境监测点 6 为例, 按照式 (2-19) 计算水环境监测点 6 的水质指标 COD 和 IN 时间序列的关联度 C , 结果分别为 0.7095 和 0.5078。计算结果均为正值且较大, 表明 COD 和 IN 在某一时刻的观测值对后期水质有较大的正方向的影响。

按照上述计算水环境监测点 6 的 COD 和 IN 时间序列的分形参数的方法, 分别计算天津市近岸海域其它 7 个水环境监测点的 COD 和 IN 时间序列的 Hurst 指数 H 、记忆长度 V_n 和关联度 C , COD 时间序列的计算结果见表 2-3, IN 时间序列的计算结果见表 2-4。

表 2-3 各水环境监测点 COD 时间序列的 Hurst 指数 H 、 C 和 V_n 值计算结果

监测 点位	所属功 能区	H	C	V_2	V_3	V_4	V_6	V_8	V_{12}
3	TJ001A	0.8312	0.5827	0.5000	0.6465	0.7064	0.8145	0.8816	0.9107
16	TJ001A	0.8801	0.6937	0.5000	0.6444	0.6918	0.7809	0.7948	1.0883
1	TJ002B	0.8706	0.6716	0.5000	0.6340	0.7568	0.8985	0.9309	0.9558
10	TJ002B	0.8321	0.5847	0.5000	0.6165	0.7212	0.8275	0.8698	0.8969
7	TJ003C	0.8967	0.7332	0.5000	0.6334	0.6891	0.7791	0.8771	1.0630
9	TJ003C	0.8702	0.6706	0.5000	0.6459	0.7616	0.8720	0.9163	0.9811
5	TJ004D	0.8507	0.6261	0.5000	0.6454	0.7188	0.8705	0.8908	0.9379
6	TJ004D	0.8868	0.7095	0.5000	0.6407	0.6908	0.8211	0.9597	0.9830

表 2-4 各水环境监测点 IN 时间序列的 Hurst 指数 H 、 C 和 V_n 值计算结果

监测 点位	所属功 能区	H	C	V_2	V_3	V_4	V_6	V_8	V_{12}
3	TJ001A	0.7907	0.4963	0.5000	0.6440	0.7446	0.7909	0.8032	0.8831

监测 点位	所属功 能区	H	C	V_2	V_3	V_4	V_6	V_8	V_{12}
16	TJ001A	0.7768	0.4677	0.5000	0.6311	0.6781	0.7163	0.7969	0.8494
1	TJ002B	0.8242	0.5674	0.5000	0.6420	0.6843	0.7238	0.7579	0.9879
10	TJ002B	0.8471	0.6180	0.5000	0.6457	0.6962	0.7769	0.8955	0.9447
7	TJ003C	0.9143	0.7760	0.5000	0.6514	0.7301	0.8599	0.9226	1.0854
9	TJ003C	0.8043	0.5248	0.5000	0.6415	0.7131	0.7484	0.8475	0.8870
5	TJ004D	0.8271	0.5737	0.5000	0.6533	0.6978	0.8386	0.8607	0.9103
6	TJ004D	0.7962	0.5078	0.5000	0.6507	0.6809	0.7247	0.8398	0.8782

(4) 结果分析

(i) 近海水质变化的持续性分析：从表 2-3 和表 2-4 可以看出，天津市近岸海域所选的 8 个水环境监测点的水质指标 COD 和 IN 时间序列的 H 值均在 0.80 左右，接近于 1，这说明近海水质污染存在明显分形特征，近海水质变化具有较强的状态持续性，即如果某一时刻，水质时间序列向上（下）变化，则下一时刻它很可能继续向上（下）变化，通过应用 R/S 分析方法对 COD 和 IN 时间序列的 H 值进行计算，揭示了近海水质变化的深层次规律，这为以后的近海水质预测提供了理论依据。

(ii) 近海水质变化的长期记忆性分析：从表 2-3 和表 2-4 可以看出，所研究的 8 个水环境监测点的水质指标 COD 和 IN 时间序列的 V_n 值随着 n 的增大一直是递增的，表明近海水质变化存在长期记忆性的特点，这说明在统计年份内，初期水质对以后水质的影响一直没有消失，即统计最初年份 1996 年的水质对 2003 年的水质仍有影响。

(iii) 近海水质变化的关联性分析：从表 2-3 和表 2-4 可以看出，所研究的 8 个水环境监测点的水质指标 COD 和 IN 时间序列的关联度 C 均大于 0.45，关联度较大，表明近海水质某一时刻的观测值对下一时刻观测值有较大的正方向的影响，说明近海水质的变化并不是随机的，而是遵循一定的非线性变化规律，这与近海水质变化具有较强的状态持续性和长期记忆性分析结果一致。

上述分析说明近海水质变化具有状态持续性和长期记忆性等明显的分形特征，水质的变化不是一个随机过程，而是随机性和确定性相统一的非线性动力学过程，其时间前后变化互相关联，有着很强的相关性，是有偏的随机游动，具有趋势性又受一定噪声的影响，以前水质的状态对现在和将来水质有很大的影响。正是近海水质分形特性的存在，使表面上看来只具有不规则、随机等特点的 COD 和 IN 变化曲线又具有一定的规律性。而实际上各陆源污染源排放污染物对近海

水质的影响有一个时间上的延迟, 污染物的变化不能直接表现在水质的变化上, 这也意味着各影响因子以非线性方式对近海水质产生影响, 这与 R/S 分析方法得到的结论相吻合。

2.3.3 结论

由于天津市近岸海域的水环境监测点每年只有枯、丰、平 3 个水期的监测数据, 这给水质变化规律的分析带来了很大的困难, 以往的时间序列分析方法不能很好的分析其规律性。本文通过对分形理论的研究, 充分利用了 R/S 分析方法可以将小(大)的时间尺度的规律用于大(小)的时间尺度范围的基本思想, 将 R/S 分析方法用于分析近海水质污染变化规律, 利用时间上低频率的水质监测数据进行大尺度的规律性分析, 从总体上把握近海水质系统的复杂特征, 把复杂的具有多重耦合关系的多变量系统转变为研究简单的单变量系统。以天津市近岸海域 8 个水环境监测点 COD 和 IN 两种水质指标为例进行计算, Hurst 指数 H 均在 0.80 左右, 且 V_n 值随着 n 的增大一直是递增的, 表明近海水质变化具有状态持续性和长期记忆性等明显的分形特征, 可进一步采用基于分形理论的预测方法对近海水质进行预测。并且, 利用该方法大小尺度规律可以互用的特点, 还可以对低频率水质监测数据进行补充, 为机理性研究服务。

2.4 近海水污染物空间分布特性研究

随着我国沿海地区人口的增长和经济的发展, 大量的工业废水和生活污水直接或间接的排入近岸海域, 导致沿海水体受到不同程度的污染。为了对近海水环境进行有效的区域化管理, 单凭近岸海域上环境监测点的水质监测数据很难把握近海水污染物的空间变化规律, 这就需要近岸海域的水污染物区域性特征进行研究。近海水污染物分布规律的研究是揭示水污染程度及空间变化规律的重要途径之一。近海水污染物监测数据的统计特征常常用来描述和刻划水污染物的分布规律。统计方法之所以能用于研究近海水污染物的分布规律, 不仅是由于近海水污染物取样常具有不确定性, 而且近海水污染物的分布本身就具有不均匀性和区域随机性。但是, 普通的统计方法并不考虑样品的空间分布和统计特征随空间度量尺度的变化性, 不具备刻划异常值的功能^[37]。而多重分形理论可以有效地克服统计方法的这些不足, 其用于研究和解释许多统计分布和空间统计分布中的疑难问题的研究成果也层出不穷^[38]。本文通过对多重分形理论进行研究, 依据环境管理部门提供的近海水环境监测数据, 建立近海水环境监测点水质数据的多重分形维数谱函数, 研究近岸海域水污染物的空间分布规律与多重分形维数谱函数形态

特征之间的关系, 确定近岸海域水污染物的空间分布特征。为进一步采用基于多重分形理论的分形曲面插值算法计算近海水环境容量提供理论依据。

2.4.1 多重分形与谱分析方法

分形理论由 B.B.Mandelbrot 于 20 世纪 70 年代提出, 其核心是认为由系统各部分组成的几何形体间具有自相似性和标度不变性, 分形维数是反映分形结构复杂程度的一个重要特征参数, 该理论用于自然科学领域和社会科学领域都取得了重大成果^[33-34]。但是, 近年来, 人们发现仅仅用单一的取决于整体特征的分形维数来描述复杂的分形及其特征是远远不够的, 自然界中许多分形往往需用多个分形维数进行描述, 这种描述即为多重分形。多重分形在空间上可以看作是单一分形的相互缠接, 是单一分形的推广, 能再现复杂分形在生长过程中不同层次的特征。多重分形所描述的往往是定义在二维或三维区域中的一种度量 μ , 如果 μ 具有空间自相似性或统计自相似性, 就可以称为具有多重分形分布。通过这种度量值 μ 的奇异性可将所描述的区域分解成一系列空间上相互缠接的子区域, 各个子区域具有不同的空间分布结构, 均构成单个分形。各个子区域除具有分形维数外, 还具有各自度量的奇异性。这些分形维数和奇异性将构成整个区域度量 μ 的多重分形维数谱函数 $f(\alpha)$, $f(\alpha)$ 一般具有倒“U”字形或“钟”形。

2.4.2 基础数据

本文以近海水环境管理中比较典型的水质指标 COD 为研究对象, 随机选择两个水期: 2000 年 8 月(丰水期)、2002 年 10 月(平水期)的水质监测数据为例进行研究, 分析近海水污染物的空间分布特征。天津市近岸海域的污染主要来自于陆源, 包括北塘排污河、大沽排污河、海河、蓟运河等^[39]。根据以往的研究^[25-26], 通过对所有水期和不同水期水质监测数据的检验, 已经排除污染物正态分布的可能, 采用灰色系统理论和统计学方法分析表明, 在不同的陆源入海河流的影响下, 天津市近岸海域的 20 个水环境监测点按照 COD 的变化规律被分成了北塘区域、大沽区域、以及独流减河区域。本文尝试采用多重分形理论对近海水污染物的空间分布特征重新进行分析, 找到污染物的空间分布规律和变异特征, 为近海水环境的区域化管理和水环境容量的计算提供理论依据。近年来, 多重分形理论用于地球化学领域许多测度分布规律方面的研究已经取得了显著的成效^[40-41], 这些成果的取得为多重分形理论在环境学领域的应用奠定了很好的基础。

2.4.3 近海水污染物空间分布的多重分形维数谱函数计算

近海水污染物在空间上的分布受物理、化学、生物以及入海河流等多种因素的影响,其中入海河流的位置及水力条件是主要的影响因素。根据近几年的近海水质监测数据可知,在入海河口处的水污染物浓度最高,而在远离入海河口海域的水污染物浓度相对较低,近海水污染物的分布是不均匀的。水污染物高浓度区只分布在入海河口处,其分布范围小且零散,所以其分布区具有分形结构,分形维数较小。同样,污染物较低浓度区的分布也具有分形特征,但是其分布范围要大的多,几乎遍及整个近海区域,其分形维数较大,接近于区域的空间拓扑维数 2。所以,近海水污染物在空间上的分布不能用一个单一的分形维数来描述,水污染物不同浓度区的分布对应于不同的分形维数和奇异性,这些分形维数和奇异性构成整个近海区域的多重分形维数谱函数 $f(\alpha)$,近岸海域水污染物空间分布场不是严格数学意义上的分形场,而是一个多重分形场。

由于近海水深较浅,COD 纵向分布可以认为是均匀的,本文只对 COD 空间分布进行二维分析,这样近岸海域水污染物的多重分形分布模型可以简要地概括如下:假设从某近海监测区域内得到一组水环境监测点的 COD 监测数据,通过对这些 COD 数据进行空间插值可形成覆盖该区的 COD 分布的网格数据。记第 i 个网格内 COD 数据的平均值为 $\rho_i(n)$, n 表示网格的边长,则第 i 个网格内的 COD 总量为: $u_i(n) = n^2 \rho_i(n)$,如果 COD 在研究海域内具有多重分形特征,那么 COD 总量 $u_i(n)$ 与网格大小 n 之间将服从幂率关系: $u_i(n) \propto n^{\alpha_i}$, Holder 指数 α_i 是某一有限指数。由于每个网格中可求得一个 α 值,且不同的 α 值将对应一组网格,如果用 $N_\alpha(n)$ 表示在网格大小为 n 时具有 α 的单元数,则: $N_\alpha(n) \propto n^{-f(\alpha)}$,则 $f(\alpha)$ 相当于度量尺度 u 为 n^α 的海域的分形维数,即多重分形维数谱函数。

随着计算机应用技术的发展,目前已有许多计算多重分形维数谱函数 $f(\alpha)$ 的方法^[42],如矩方法、直方图法、小波方法、乘法法、二次维矩方法等。在这些计算方法中,矩方法是最常用的方法之一。用矩方法计算 $f(\alpha)$ 的基本步骤如下^[43]:

设近岸海域上 20 个水环境监测点的原始 COD 监测数据点为 $\{P_j\}_{j=1}^{20} = \{(x_j, y_j, z_j)\}_{j=1}^{20}$,其中: (x_j, y_j) 表示水环境监测点 j 的平面坐标, z_j 表示水环境监测点 j 的 COD 浓度, mg/L。将诸点所分布的矩形区域 $S = \{(x, y), \min(x_j) \leq x \leq \max(x_j), \min(y_j) \leq y \leq \max(y_j)\}$,用尺码 n 沿 x 方向和 y 方向自点 $(\min(x_j), \min(y_j))$ 起分别作等距分划,则矩形区域 S 除了沿其右边界和上边界各有一列和一行边长不超过 n 的矩形之外,将被分成若干个小正方形。称这些被剖分出来的小正方形和小矩形为单元,令 $u_i(n)$ 代表第 i 个单元上的 COD 总量。

(1) 建立质量分配函数 $r_q(n)$ ，在双对数坐标图上绘出 q 的质量分配函数 $r_q(n)$ 和网格大小为 n 的映射图：

$$r_q(n) = \sum_i \mu_i(n)^q \tag{2-20}$$

式中： q 为 $u_i(n)$ 的统计矩阶数，通常取整数。

(2) 计算质量指数 $t(q)$ ：若 $u_i(n)$ 满足多重分形条件，则对于任何 q 值， $r_q(n)$ 与 n 之间具有如下的指数关系：

$$r_q(n) \propto n^{t(q)} \tag{2-21}$$

式中： $t(q)$ 为质量指数。对式 (2-21) 取对数， $\lg n$ 与 $\lg r$ 呈线性关系， $t(q)$ 是直线的斜率。实际计算时，可以利用不同长度的尺码 n 对区域 S 作多次分划，取 q 的部分值，利用最小二乘法进行线性拟合，可容易地求得 $t(q)$ 。

(3) 计算 Holder 奇异指数 $\alpha(q)$ ：利用差分法由式 (2-22) 可以求出 $\alpha(q)$ ，式 (2-22) 中一阶导数由相邻数据对 $q \pm 0.005$ 的估计值 $t(q)$ 决定，把这些数据中任意两个的差值除以 0.01 便可以求出 $\alpha(q)$ ：

$$\alpha(q) = \frac{\partial t(q)}{\partial (q)} \tag{2-22}$$

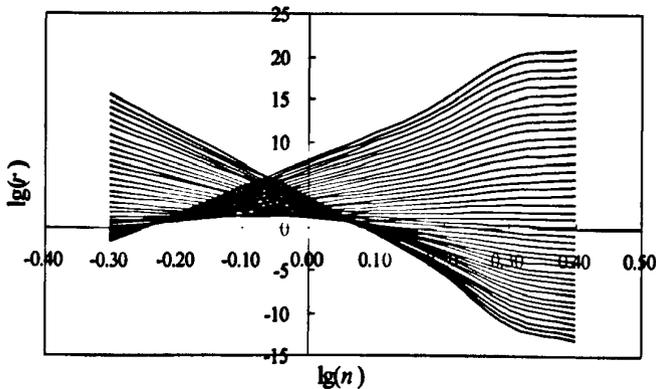
(4) 计算多重分形维数谱函数 $f(\alpha)$ ：对式 (2-22) 进行 Legendre 变换便可以得到多重分形维数谱函数 $f(\alpha)$ ：

$$f(\alpha) = q\alpha(q) - t(q) \tag{2-23}$$

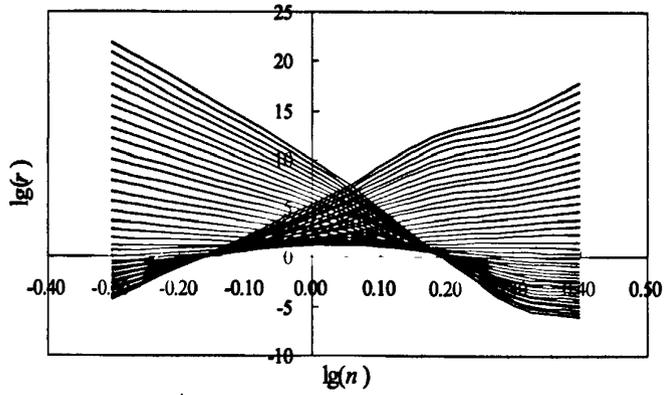
(5) 绘制 $\lg n - \lg r$ 图、 $q - \alpha(q)$ 图、 $q - t(q)$ 图、 $\alpha(q) - f(\alpha)$ 图。

按照上述方法步骤，以天津市近岸海域 20 个水环境监测点 2000 年 8 月（丰水期）、2002 年 10 月（平水期）的 COD 监测数据进行分析，以步长为 1 取统计矩阶数 q 从 20 变化到 -20 共 41 个值进行计算，计算结果见图 2-6—图 2-9（其中：

(a) 对应 2000 年 8 月的计算结果，(b) 对应 2002 年 10 月的计算结果）。

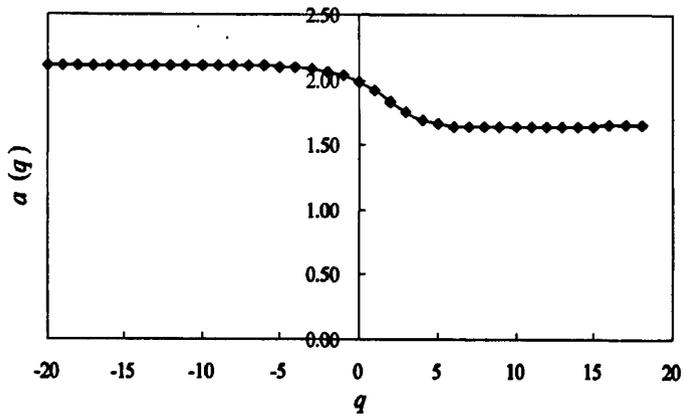


(a)

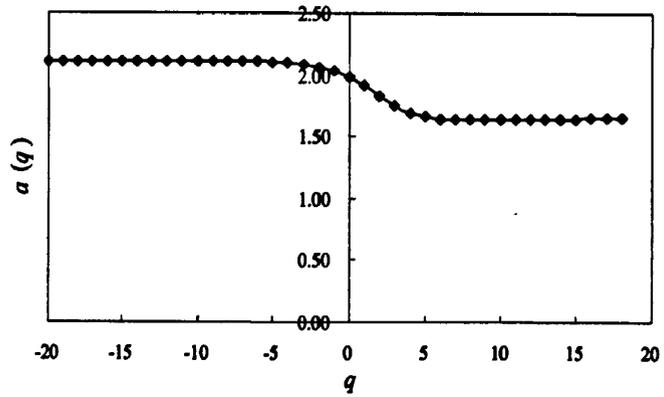


(b)

图 2-6 质量分配函数 r 与网格尺寸 n 的对数图 (其中从上到下分别对应 $q = 20, 19, \dots, -20$)

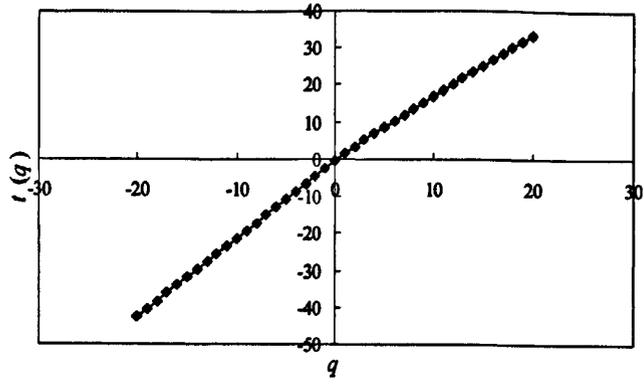


(a)

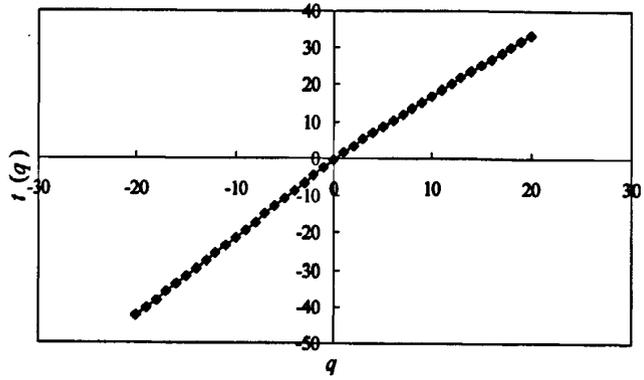


(b)

图 2-7 奇异指数 $\alpha(q)$ 与统计矩阶数 q 的关系

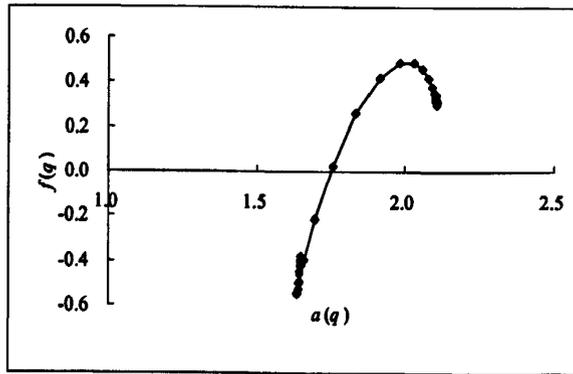


(a)

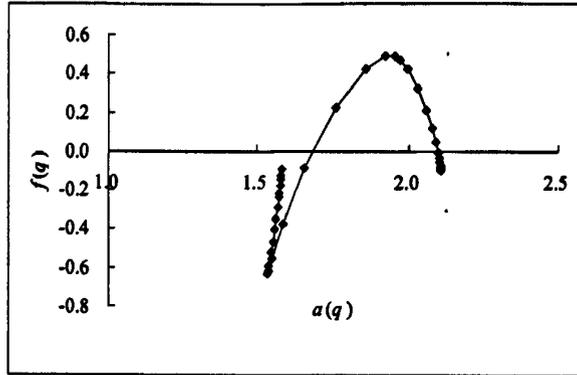


(b)

图 2-8 质量指数 $t(q)$ 与统计矩阶数 q 的关系



(a)



(b)

图 2-9 多重分形维数谱 $f(\alpha)$ 与奇异指数 $\alpha(q)$ 的关系

2.4.4 近海水污染物空间分布规律

采用多重分形矩方法计算近海 20 个水环境监测点在 2000 年 8 月（丰水期）和 2002 年 10 月（平水期）的 COD 分布的多重分形维数谱函数 $f(\alpha)$ ， $f(\alpha)$ 与 $\alpha(q)$ 的关系分别见图 2-9 中的 (a) 和 (b)。由图 2-9 可知，随机选择的两个不同水期 COD 的多重分形维数谱函数 $\alpha(q)-f(\alpha)$ 曲线均为连续上凸曲线，表明近海水污染物 COD 的空间分布具有连续多重分形特征，连续多重分形是近海水污染物空间分布的一种普遍规律。以顶点为分界点， $\alpha(q)-f(\alpha)$ 曲线的左半段主要反映了 $q \geq 0$ 的特征，能刻划分形的基本特征，实际上重点突出了较高浓度的特征，由图 2-9 可知， $\alpha(q)-f(\alpha)$ 为一不对称的上凸曲线，表明近海水污染物多重分形场曾经历过不同程度的局部叠加， $\alpha(q)-f(\alpha)$ 曲线的左半段宽于右半段，相对来说近海水污染物高浓度部分的奇异性较大，这与天津市近岸海域陆源入海河流较多，引起近海河流入海口位置多处污染物浓度监测值局部过高的实际相一致。从这个角度出发，天津市近海水污染物在空间上的分布可以理解为，从区域到局部水污染物的分布可以划分出背景区与异常区两个或多个不同的分形分布模式，近海整个区域水污染物在空间上的分布刚开始可以看作为具有相同的浓度值（背景区），随后受到不同入海河流排入污染物的不同程度影响的叠加，而导致近海水污染物在空间上的分布出现差异，即近海水污染物多重分形场曾经历过不同程度的局部叠加。另一方面，由图 2-9 可见，图 2-9-(a) 对应的 2000 年 8 月（丰水期）的 $\alpha(q)-f(\alpha)$ 曲线窄而尖，而图 2-9-(b) 对应的 2002 年 10 月（平水期）的 $\alpha(q)-f(\alpha)$ 曲线宽而缓，图 2-9-(a) 的 $\Delta\alpha = \alpha_{\max} - \alpha_{\min}$ 值大于图 2-9-(b)，表明平水期近海水污染物空间分布的奇异性较丰水期大，而实际上，天津市近岸海域所属的渤海湾是一个近封闭的海域，与外海的交换速度特别慢，丰水期河流的入海流量比平水期的入海流量大得多，加上海洋本身也处于丰水期，季节性的

洋流运动较强, 这些都有助于水污染物向外部扩散而不是滞留在入海口附近, 促进水污染物在空间上分布的一致性, 使得近海水污染物空间分布的变异性较小, 理论分析的结果与实际情况相一致。

2.4.5 结论

通过对多重分形理论进行研究, 首次尝试采用多重分形维数谱方法对近海水污染物在空间上的分布特征进行分析。以天津市近岸海域为例, 应用分析表明: 近海水污染物在空间上的分布具有连续多重分形分布的特征, 其中水污染物高浓度部分的变异性较大; 由于入海河流和海洋运动季节性规律的影响, 丰水期水污染物空间分布的变异性较平水期小。这些研究结果的取得, 为进一步采用基于多重分形理论的分形曲面插值算法计算天津市近岸海域的水环境容量奠定了理论基础。同时, 本文得到的结论具有普遍的意义, 连续多重分形是近海水污染物空间分布的一种普遍规律, 一般情况下, 从区域到局部污染物的分布可以划分出背景区与异常区两个或多个不同的分形分布模式, 即多重分形嵌套模式; $\alpha(q)-f(\alpha)$ 曲线两侧开口的大小具有重要的意义, 开口越大, 显示水污染物在空间上的分布变异性越大, 水污染物的多重分形维数谱函数的形态特征能为评价近岸海域的污染状况提供重要的线索。这些结论可以为其它海域的区域化管理提供指导意义。

2.5 小结

本文通过对分形理论和多重分形理论的研究, 采用分形理论中的重标极差分析方法 (R/S 分析) 对各个水环境监测点 COD 和无机氮时间序列进行分析, 计算其 Hurst 指数 H , 计算结果显示其 Hurst 指数 H 均在 0.80 左右, 表明近海水水质变化具有明显的分形特征; 然后, 采用多重分形理论中的多重分形维数谱方法对近海水污染物空间分布特征进行分析, 利用多重分形矩方法计算 20 个水环境监测点 COD 空间分布的多重分形维数谱函数 $f(\alpha)$, 计算结果显示其 $\alpha(q)-f(\alpha)$ 为一连续上凸曲线, 表明近海水污染物空间分布具有连续多重分形分布特征。本章对近海水污染物时空分布特征的分析结果与天津市近海实际情况相一致, 这些结果的取得为下一步基于分形理论的近海水水质预测和水环境容量计算提供了理论依据。具体来讲, 本章内容有以下创新点:

(1) 本文首次尝试采用分形理论用于近海水污染物时间序列变化特征的分析, 提出了近海水污染物时间序列是一具有可持续性的分形时间序列。

(2) 本文首次尝试采用多重分形理论用于近海水污染物空间分布特征的分

析,提出了近海水污染物空间分布场是一多重分形场的概念,近海水污染物空间分布场经历了不同程度的叠加,连续多重分形是近海水污染物空间分布的一种普遍规律,一般情况下,从区域到局部水污染物的分布可以划分出背景区与异常区两个或多个不同的分形分布模式,即多重分形嵌套模式。

(3) 该研究方法克服了以往统计方法分析近海水污染物时空变化特征的不足,研究结论的取得为后续水质预测和水环境容量计算的研究提供了理论依据。

第三章 近海水污染物浓度预测方法的研究

3.1 概述

近海水质系统是一个开放的、复杂的、非线性的动力学系统，近海水质的变化受物理、化学、生物以及人类活动等多种因素的影响，表现出复杂的非线性特征。为了让决策者了解近海水环境的变化趋势，有效的防治近海水质的污染，这就需要对近海水环境质量做出准确预测，从而提出针对性强的应对措施。根据前面章节的介绍不难发现，近海水环境质量的预测和模拟多以高维水动力学方程为基础，结合潮波方程，对污染物在近海的输移和扩散进行数值模拟。同时，还要综合考虑物理、化学、生物等因素，并在模型中体现出来。这些机理性研究能够为其它预测方法提供理论基础，但是近海环境的复杂性使得上述机理性研究工作举步维艰，模拟模型所需基本数据很难获得，如果直接将这些研究应用于环境管理的实际，特别是发展中国家的环境管理，存在许多不足：机理性研究的原理过于复杂、在模型使用过程中，参数确定困难、花费巨大等。

但是，环境管理部门每年对近岸海域都要进行常规和非常规的监测，这些监测数据虽然存在时间和空间跨度大的问题（天津市近岸海域常规监测每年进行三次，集中在5、8、10月），但是这些是环境管理部门进行管理的依据，比较有说服力和法律效力。所以，在进行环境管理的研究时，要充分利用这些环境监测数据。

本文在第二章的内容中，对天津市近岸海域的水环境监测点的COD和无机氮时间序列进行了分析，并验证了近海水质变化具有明显的分形特征。在本章内容里，力求通过对分形理论的研究，依据环境管理部门和水利部门提供的水环境监测数据，以COD为例，对近海水环境监测点的COD浓度变化进行预测。针对我国近海水质污染呈现的年周期性和自相似性特点，本文通过对分形插值算法进行研究，在分析近海水质COD时间序列具有分形特征的基础上，根据分形时间序列具有长期持续性和大小时间尺度规律可以互用的特点，首次尝试采用分形理论对近海水污染物浓度进行预测。首先，根据分形拼贴定理，由基于仿射变换的分形插值方法求取各历史时间阶段水质变化的迭代函数系，然后根据近海水质变化的年周期性，对上述求得的迭代函数系加权求和，得到预测年份水质变化的

统计意义上的迭代函数系, 建立分形预测模型, 最后应用随机迭代算法求得预测年份水质变化曲线的吸引子, 对近海水质进行预测, 预测结果显示平均预测误差为 24.4%。应用表明, 该方法预测精度较高、实用性强, 能够为近海水环境管理提供决策支持。

3.2 分形理论基础

3.2.1 迭代函数系 (Iterated Function System, IFS) 及吸引子定理

到目前为止, 用迭代函数系去解析地构造、研究自然界中具有自相似结构的分形最为成功^[44]。一个迭代函数系 IFS 由一个完备的度量空间 (X, d) 和一组有限的压缩映射集 $W_n: X \rightarrow X$ 组成, W_n 的压缩因子为 S_n , 且 $0 \leq S_n < 1 (n=1, 2, \dots, N)$, IFS 用 $\{X; W_n, n=1, 2, \dots, N\}$ 表示, 且 IFS 的压缩因子为 $S = \max\{S_n : n=1, 2, \dots, N\}$ 。

变换 $W: H(X) \rightarrow H(X)$ 定义为: $W(B) = \bigcup_{n=1}^N W_n(B)$, $\forall B \in H(X)$ 是完备度量空间 $(H(X), h(d))$ 上具有压缩因子 S 的压缩映射, 即 $h(W(A), W(B)) \leq Sh(A, B)$, $\forall A, B \in H(X)$ 成立, 则必定存在唯一不动点 $P \in H(X)$ 满足 $P = W(P) = \bigcup_{n=1}^N W_n(P)$, 且 P 可用下式给出, 即对 $\forall B \in H(X)$, 有

$$P = \lim_{n \rightarrow \infty} W^n(B) \quad (3-1)$$

式中: $W^n(B)$ 表示变换 W 的 n 次迭代, 即 $W^n(B) = W(W(\dots W(B)))$, 则不动点 P 为该 IFS 的吸引子^[45]。

3.2.2 分形拼贴定理

在完备度量空间 (X, d) 中, 给定 $L \in H(X)$ 和 $\varepsilon > 0$, 选定一个具有压缩因子 $S: 0 \leq S < 1$ 的 IFS $\{X; W_0, W_1, \dots, W_n\}$, 使得:

$$h(L, \bigcup_{n=0}^N W_n(L)) \leq \varepsilon \quad (3-2)$$

则
$$h(L, P) \leq \varepsilon / (1 - S) \quad (3-3)$$

式中: P 为 IFS $\{X; W_0, W_1, \dots, W_n\}$ 的吸引子, $h(L, P)$ 是 Hausdorff 度量^[45]。

拼贴定理说明必存在一个 IFS $\{X; W_0, W_1, \dots, W_n\}$, 其吸引子 P 近似或相似于一个给定的集合 L 。因为不动点 P 是从它自身的变换 $W(P)$ 构造而来的, 所以对给定的集合 L 作压缩变换, 然后把它们粘贴在一起以重构 L 。

3.2.3 基于仿射变换IFS码的分形插值方法

分形插值与传统的函数插值既有相同点也有不同点：相同点在于都需要根据一组给定的插值点（采样点）通过一定的规则构造插值函数；不同点在于传统方法在整个区间内构造出一个函数或一个分段函数，而分形插值在整个区间内构造出一个 IFS，该 IFS 满足吸引子定理和分形拼贴定理，即从此函数空间内的任一初始点利用 IFS 进行迭代，其吸引子为通过这组插值点的函数点图，该图由 IFS 而不是通常的一个解析函数来确定^[46]。

一般来说，分形插值所需要的 IFS 码可通过下述仿射变换得到：

$$W_i \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} a_i & 0 \\ c_i & d_i \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} e_i \\ h_i \end{bmatrix} \quad (3-4)$$

式中： (x, y) 为某一点的坐标，即插值点； W_i 为仿射变换关系； a_i 、 c_i 、 d_i 为变换矩阵元素； e_i 、 h_i 为 (x, y) 变换后的常数分量。

若要求此 IFS 的吸引子通过给定插值点 $\{(x_i, y_i) : i=1, 2, \dots, N\}$ ，则每个变换必须满足以下条件：

$$w_i \begin{bmatrix} x_0 \\ y_0 \end{bmatrix} = \begin{bmatrix} x_{i-1} \\ y_{i-1} \end{bmatrix} \quad w_i \begin{bmatrix} x_N \\ y_N \end{bmatrix} = \begin{bmatrix} x_i \\ y_i \end{bmatrix} \quad (3-5)$$

即原始插值点区间的左端点 (x_0, y_0) 映射到子区间的左端点 (x_{i-1}, y_{i-1}) ，点区间的右端点 (x_N, y_N) 映射到子区间的右端点 (x_i, y_i) 。首先通过经验法或解析法确定自由参数 d_i 的值^[47]，然后将 (3-4) 式代入 (3-5) 式确定仿射变换的其它参数 a_i 、 c_i 、 e_i 和 h_i 的值：

$$a_i = \frac{x_i - x_{i-1}}{x_N - x_0} \quad (3-6)$$

$$c_i = \frac{y_i - y_{i-1}}{x_N - x_0} - d_i \frac{y_N - y_0}{x_N - x_0} \quad (3-7)$$

$$e_i = \frac{x_N x_{i-1} - x_0 x_i}{x_N - x_0} \quad (3-8)$$

$$h_i = \frac{x_N y_{i-1} - x_0 y_i}{x_N - x_0} - d_i \frac{x_N y_0 - x_0 y_N}{x_N - x_0} \quad (3-9)$$

在求得各参数后，便可以求出 IFS 中第 i 个仿射变换 W_i 。求得各 W_i 后，可用确定型或随机型迭代算法得到 IFS 的吸引子。随着迭代次数的增加，插值获得的曲线与原曲线的拟合程度不断提高，经过多次迭代将形成一条稳定的不变插值曲线，它不仅通过采样的插值点，而且与原采样曲线充分逼近。

3.3 近海水污染物分形预测模型

本文以水环境监测点 9 为例说明近海水污染物分形预测模型建立的基本步骤, 其中水环境监测点 9 的 COD 监测数据见图 3-1, 可以看出, COD 监测值变化范围比较大, 由于近海水质的变化表现出复杂的非线性特征, 使研究的 COD 时间序列数据前后关系很不明显, 即便是单个水环境监测点的单个水质指标, 其变化规律也难以把握, 给以往的预测方法带来很大困难。本文通过对分形理论的研究, 在分析近海水质 COD 时间序列具有明显分形特征的基础上, 利用分形时间序列大小时间尺度规律可以互用的特点, 尝试采用分形理论对天津市近海水质进行预测。

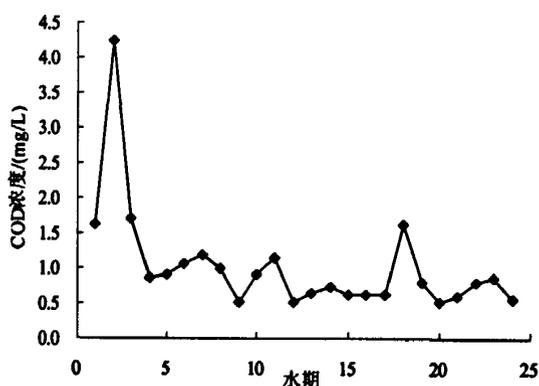


图 3-1 水环境监测点 9 的 COD 监测数据

下面说明利用近海水质变化的年周期性和自相似性, 根据分形时间序列具有大小时间尺度规律可以互用的特点, 采用分形插值方法^[48]对近海水质指标 COD 进行预测的方法, 具体步骤如下:

(1) 将水环境监测点 9 的 COD 监测值从 1996 年至 2003 年每两年组合在一起, 成为 1 个时间阶段, 这样得到 4 个时间阶段, 每个时间阶段内包含 6 个 COD 监测数据, 取离 2004 年最近的时间阶段, 即 2002 年和 2003 年组合为基准阶段。

(2) 获取每个时间阶段的 COD 数据作为样本数据, 确定插值点集合。在这里取每个时间阶段的水质监测时间(水期)为横坐标, 分别用 0 到 5 表示, 取 COD 监测数据值作为纵坐标, 这些点一起构成插值点集合。

(3) COD 样本数据标准化处理。

$$x_i = \frac{X_i - X_{\min}}{X_{\max} - X_{\min}} \quad (i = 0, 1, \dots, 5) \quad (3-10)$$

$$y_i = \frac{Y_i - Y_{\min}}{Y_{\max} - Y_{\min}} \quad (i = 0, 1, \dots, 5) \quad (3-11)$$

式中： X_i 为监测时间， Y_i 为对应该时间点的 COD 监测数据，mg/L； X_{\min} 、 X_{\max} 分别为样本时间点最小值和最大值； Y_{\min} 、 Y_{\max} 分别为 COD 样本内的最小值和最大值，mg/L； x_i ， y_i 分别为标准化后的样本数据。

(4) 求取各时间阶段 COD 负荷曲线的 IFS 码。以第(3)步中标准化后的插值点集合，按照式(3-4)和(3-5)求取 COD 负荷曲线的 IFS 码，由于插值点数据较少，不能用解析法求出垂直尺度因子 d_i ，在这里依据经验取 $d_i=0.2$ ($i=1, 2, \dots, 5$)，然后计算其它参数，确定仿射变换 w_i 。

(5) 对已求得的 4 个 IFS 码进行相应参数加权求和，求得一个统计意义上的 IFS 码，可依据经验对离预测年最近的时间阶段的参数赋予较大的权重。在这里分别取 0.2、0.2、0.3、0.3。按照此方法计算的水环境监测点 9 的 2004 年 COD 变化曲线的 IFS 码见表 3-1。

表 3-1 水环境监测点 9 的 2004 年水质 COD 变化曲线的 IFS 码

	a	c	d	e	h
1	0.2000	-0.2183	0.2000	0.0000	0.4966
2	0.2000	-0.0394	0.2000	0.2000	0.2301
3	0.2000	0.1325	0.2000	0.4000	0.1425
4	0.2000	0.1994	0.2000	0.6000	0.2268
5	0.2000	-0.0742	0.2000	0.8000	0.3780

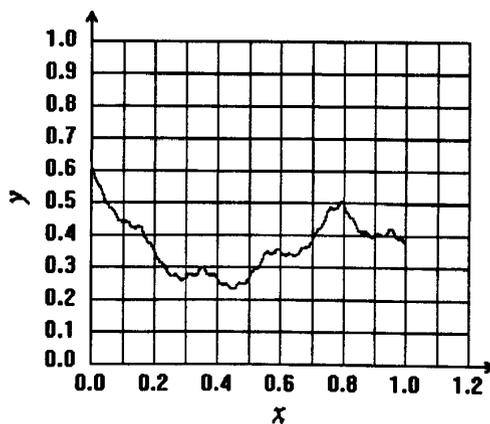


图 3-2 根据表 3-1 中的 IFS 码迭代得到的水环境监测点 9 的吸引子

(6) 应用随机型迭代算法^[49]，通过任意点进行迭代求得上面的统计意义上的 IFS 码的吸引子，此吸引子便可以认为是基于历史水质数据外推一个周期得到

的 COD 负荷曲线。迭代得到的水环境监测点 9 的 2004 年 COD 变化曲线的吸引子见图 3-2。

(7) 水质预测。根据 x_i 的值从图 3-2 中的吸引子中读出所对应的 y_i 的值，依据近海水质 COD 的历史监测数据对 COD 变化范围进行预估，然后根据式 (3-11) 进行反算便可以得到 2004 年的 COD 预测值。水环境监测点 9 在 2004 年枯水期、丰水期、平水期的 COD 预测值分别为：0.97、0.78、0.71，预测误差分别为：49.4%、25.0%、11.2%，平均误差为 28.5%。

按照上述预测步骤，分别预测其它 8 个水环境监测点在 2004 年 3 个水期的 COD 值，预测结果见表 3-2。表 3-2 显示，预测值的最大相对误差为 49.4%，最小相对误差为 2.3%，平均相对误差为 24.4%，能够满足近海水水质预测实际应用的需要。

表 3-2 基于分形理论的近海水水质预测结果

监测 点位	2004 年枯水期			2004 年丰水期			2004 年平水期		
	监测值	预测值	误差	监测值	预测值	误差	监测值	预测值	误差
	(mg/L)	(mg/L)	(%)	(mg/L)	(mg/L)	(%)	(mg/L)	(mg/L)	(%)
01	0.84	0.91	8.3	0.48	0.69	43.7	0.64	0.79	23.4
02	2.40	1.63	32.1	0.76	0.89	17.1	1.04	1.22	17.3
03	0.64	0.73	14.1	0.52	0.66	26.9	1.12	1.06	5.4
06	0.56	0.71	26.8	1.84	1.22	33.7	1.00	1.09	9.0
08	2.96	1.58	46.6	1.36	0.93	31.6	1.28	1.31	2.3
09	1.92	0.97	49.4	1.04	0.78	25.0	0.80	0.71	11.2
10	2.08	1.66	20.2	1.12	1.35	20.5	0.80	1.04	30.0
12	1.58	0.95	39.8	1.48	1.29	12.8	1.00	1.32	32.0
15	1.28	0.91	28.9	1.56	0.99	36.5	1.20	1.03	14.2
平均									
误差			29.6			27.5			16.1

3.4 小结

在本章内容中，本文首先对分形理论的相关内容进行了学习，在第二章分析近海水污染物时间序列具有分形特征的基础上，根据分形拼贴定理，由基于仿射变换的分形插值方法求取各历史时间阶段水质变化的迭代函数系，然后根据近海水水质变化的年周期性，对上述求得的迭代函数系加权求和，得到预测年份水质变

化的统计意义上的迭代函数系,建立分形预测模型,最后应用随机迭代算法求得预测年份水质变化曲线的吸引子,对近海水质进行预测,取得了令人满意的效果。具体来讲,本章主要有以下创新点:

(1) 由于近海环境监测数据具有监测频率低、时间跨度大、数据样本少等特点,这给以往预测方法带来了很大的困难,而本文提出的近海水污染物分形预测模型充分利用了近海水污染物变化的年周期性和分形时间序列具有大小时间尺度规律可以互用的特点,预测精度令人满意。

(2) 首次将分形理论引入到近海水污染物浓度预测中,不仅拓宽了分形理论的应用领域,而且为近海水污染物浓度的预测提供了新的方法,可以尝试采用该方法为近海水环境的管理提供决策支持。

第四章 近海水环境质量评价方法的研究

4.1 概述

随着我国沿海地区经济的发展,使得大量的生活污水和工业废水直接或间接地排入近岸海域,导致沿海水体受到不同程度的污染和生态环境的破坏,由此造成的经济损失十分可观。从 20 世纪 80 年代起,越来越多的科研工作者针对我国近海水域的污染情况,特别是水体富营养化、石油类污染、重金属污染等进行了深入的研究,取得了很多成果。作为这些研究的基础,对各个近海水域水质的监测与分析、污染特点的掌握也是至关重要的。

根据从 1998 年 7 月 1 日起实施的海水水质标准 (GB3097-1997),海洋水质监测的项目有 35 个之多。这些项目分别针对不同的侧重点,包括物理性状、营养化水平、重金属、油类、放射性、毒性物质等。沿海地区根据各自近海污染特征的不同,除了必测的项目外,分别选取其中的部分项目进行监测。本文的研究海域,即天津市近岸海域的监测项目包括 pH 值、COD、无机氮、活性磷酸盐等 23 项指标。监测项目的多样性使得环境管理者能够全面的认识近岸海域的污染特征和污染程度,但同时对环境评价提出了更高的要求。

现在环境管理一般采用的是单因子指数法和综合指数法,这两种方法以数字的形式给出了每种评价项目的污染程度和总体的污染状况。但是,这两个评价指数在应用上存在一些问题:单因子指数法只考虑了单个监测项目的污染情况,而未考虑与之相关的监测项目的污染程度,比如富营养化现象,不是单纯某个指标过高就可以发生的。综合指数法虽然综合考虑了参加评价的众多监测项目的污染情况,但是单从结果看不能体现出各个因子的重要性差异,也不能很好的反映污染特征。同时,以上两种方法的计算过程较为复杂,结果的可视化程度不高,不太适合于近海水环境管理的需要。

本文通过对非线性映射理论进行研究,力求借助于 DPS 数据处理系统将高维水质数据通过非线性映射变换,得到二维映射值并将其描绘在平面坐标上,从而实现高维水质信息的可视化,然后根据映射点的位置关系建立近海水质的综合评价模型,实现近海水环境质量的准确评价。通过将非线性映射理论应用到近海监测数据的分析与评价,提出了一种崭新的近海水环境质量的可视化评价方

法。

4.2 非线性映射理论基础

4.2.1 非线性映射的具体原理

非线性映射方法是 Sammon 于 1969 年提出的一种几何降维的数学方法, 因此分析也是一种降维方法, 即将高维变量综合为少数几个综合变量(原来变量的线性组合), 使综合指标能够最大限度地表达原来多个指标的信息。非线性映射通过某种非线性变换, 将高维空间的几何图象变换为低维(一维、二维等)空间中的图像, 并且变换后仍能近似的保持原来的几何关系。这种方法直观、形象, 使人们能够在低维空间中看到一些高维样本点相互关系的近似图像。该方法在分析过程中只针对数据, 而不考虑数据的类型, 在很大程度上克服了其它聚类分析方法以下方面的不足: 首先, 某种特定的聚类分析方法, 分类结果在很大程度上依赖于一系列人为(主观地)确定的参数, 如样本间相似性度量、各种相似性阈值、聚类方法的选择等, 而且没有一个很好的方法评价所得的聚类结果; 其次, 聚类谱系图并不能表达各个样本间的真实关系, 当两个类群互相接近时, 它们之间的样本有形成一个桥梁的趋势, 并有可能造成虚假的归并。鉴于以上优点, 非线性映射分析方法在数据处理方面得到了广泛的应用^[50]。

4.2.2 非线性映射分析的映射过程

设有 n 个样本, 每个样本包含 p 项观测指标, 则每一个样本点就相当于 R^p 空间中的一个 p 维点: $X_i = \{x_{i1}, x_{i2}, \dots, x_{ip}\} (i=1, 2, \dots, n)$ 。要将 R^p 空间中的 n 个 X_i 映射到低维空间 $R^L (L < p)$ 中, 即通过非线性映射后把 R^p 空间中的 n 个样本点变为低维空间 R^L 中的 n 个点: $Y_i = \{y_{i1}, y_{i2}, \dots, y_{iL}\} (i=1, 2, \dots, n)$ 。为便于画图, 一般取 $L=2$ 。 R^L 空间中 n 个样本点 Y_i 间的距离在经过非线性映射后必须仍然近似于高维空间 R^p 中 n 个样本点间的距离。

$$X_i = \{x_{i1}, x_{i2}, \dots, x_{ip}\} \Rightarrow Y_i = \{y_{i1}, y_{i2}, \dots, y_{iL}\} \quad (4-1)$$

具体的映射原理见式(4-1), 这一映射变化过程是通过一非常复杂的高阶偏微分方程来实现的。为达到上述目的, 需要引入非线性映射从高维变换到低维的约束条件:

$$K = \frac{1}{\sum_{i < j} d_{ij}^*} \sum_{i < j} \frac{(d_{ij}^* - d_{ij})^2}{d_{ij}^*} = \frac{1}{nF} \sum_{i < j} w_{ij} (d_{ij}^* - d_{ij})^2 \quad (4-2)$$

式中: $nF = \sum_{i < j} d_{ij}^*$ 为标准化因子, $w_{ij} = \frac{1}{d_{ij}^*}$ 为权重系数, d_{ij}^* 为原来空间 R^p 中样本点 X_i 与 X_j 间的距离, d_{ij} 为新空间 R^l 中样本点 Y_i 与 Y_j 间的距离, K 的含义是使原来空间距离与新空间距离之差的平方和达到最小时来求得新空间点的几何构形。具体计算步骤如下:

(1) 为了避免不同观测数据数量级之间的差异, 首先对分析的观测数据进行数据转换处理, 通常的方法有中心化、标准化、规格化等。本文选择标准化处理方法。

(2) 计算 R^p 空间内任意两点 X_i 与 X_j 之间的欧氏距离 d_{ij}^* :

$$d_{ij}^* = \sqrt{\sum_{k=1}^p (x_{ik} - x_{jk})^2} \quad (4-3)$$

其中: $i, j = 1, 2, \dots, n$, 得到距离矩阵 D^* :

$$D^* = \begin{bmatrix} d_{12}^* & d_{13}^* & \cdots & d_{1n}^* \\ & d_{23}^* & \cdots & d_{2n}^* \\ & & \ddots & \vdots \\ & & & d_{(n-1)n}^* \end{bmatrix} \quad (4-4)$$

(3) 任取 R^l 空间中的 n 个初值点: $Y_1 = \{y_{11}, y_{21}, \dots, y_{l1}\}$, $Y_2 = \{y_{12}, y_{22}, \dots, y_{l2}\}$, \dots , $Y_n = \{y_{1n}, y_{2n}, \dots, y_{ln}\}$, 并将它们代入 K 中, 则 K 为 $l \times n$ 个变量 y_{ij} 的函数, 用迭代方法求使 K 达到极小值时新空间 R^l 中的 n 样本点 Y_i 值, 用欧氏距离表达有:

$$d_{ij}(m) = \sqrt{\sum_{k=1}^l (y_{ik}(m) - y_{jk}(m))^2} \quad i, j = 1, 2, \dots, n \quad (4-5)$$

式中: m 为迭代次数。采用最速下降法搜索映射误差的极小值, 其迭代公式为:

$$y_{ij}(m+1) = y_{ij}(m) - MF \Delta_{ij}(m) \quad (4-6)$$

式中: MF 为魔力因子, 一般凭经验取 0.3 或 0.4。

$$\Delta_{ij}(m) = \frac{\partial K(m)}{\partial y_{ij}(m)} \bigg/ \left| \frac{\partial^2 K(m)}{\partial (y_{ij}(m))^2} \right| \quad (4-7)$$

$$\frac{\partial K(m)}{\partial y_{ij}(m)} = -\frac{2}{nF} \sum_{\alpha=1}^n \left[\frac{d_{i\alpha}^* - d_{i\alpha}}{d_{i\alpha} \cdot d_{i\alpha}^*} \right] (y_{ij} - y_{i\alpha}) \quad (4-8)$$

$$\frac{\partial^2 K(m)}{\partial y_{ij}^2(m)} = -\frac{2}{nF} \sum_{\alpha=1}^n \frac{1}{d_{i\alpha} d_{i\alpha}^*} \cdot \left[(d_{i\alpha}^* - d_{i\alpha}) - \frac{(y_{ij} - y_{i\alpha}^*)^2}{d_{i\alpha}} \left(1 + \frac{d_{i\alpha}^* - d_{i\alpha}}{d_{i\alpha}} \right) \right] \quad (4-9)$$

为便于画图, 常取 $l=2$ 。为了减少运算时间, 系统先对原始数据进行主成分

分析,找出前两个主成分,构成一个因子面,然后把 n 个样本点 X_1, X_2, \dots, X_n 在此面上的因子得分点 Y_1, Y_2, \dots, Y_n 作为初始构形进行迭代分析。

4.3 近海水环境质量评价模型

4.3.1 基础数据

本文采用的原始数据来自于天津市近岸海域 20 个水环境监测点的水质监测数据。天津市近海的常规环境监测指标包括水温、盐度、溶解氧、pH 值、无机磷、化学需氧量、总氮、叶绿素 A、石油类等 23 项水质指标。水质监测指标非常完善,但是以往普通的水环境质量评价方法不能很好的综合利用这部分水质信息对近海水环境质量做出综合评价:单因子指数法只考虑了单个监测项目的污染情况,而未考虑与之相关的监测项目的污染程度;比如富营养化现象,不是单纯某个指标过高就可以发生的。综合指数法虽然综合考虑了参加评价的几个监测项目的污染情况,但是由于模型的复杂性不能把所有的水质指标纳入进行综合考虑,也不能很好的反映污染特征。本文通过对非线性映射理论进行分析,借鉴该理论在其它领域方面的应用^[51-53],将该理论首次应用到近海水环境质量评价中,通过非线性映射分析方法建立近海水环境质量综合评价模型,并对 2002 年三个水期的水环境监测数据为例进行分析和评价。

4.3.2 近海水环境质量评价模型的建立

考虑到天津市近岸海域的污染主要表现为富营养化、石油类和重金属污染的特点,参考其它水环境质量评价模型常用的水质评价指标,并根据监测数据的完整性和《海水水质标准》(GB3097-1997)中对水质指标的定量关系,本文选择 11 个定量水质指标进行综合分析,包括溶解氧(DO)、无机氮(DIN)、化学需氧量(COD)、无机磷(DIP)、石油类、六六六、铜、锌、铅、铬、汞等 11 个水质指标组成天津市近海水环境质量的评价指标体系。根据国家《海水水质标准》(GB3097-1997)对第一、第二、第三、第四类海水水质的要求确定天津市近海水水质 11 个评价指标的标准值(见表 4-1)。将四类海水水质的标准值与天津市近岸海域 20 个水环境监测点的水质监测数据一起定义成数据矩阵,然后采用非线性映射分析方法对数据进行映射变换,这样可将 11 维的水质指标在保留其空间结构不变的情况下映射到 2 维平面中,得到一系列空间分布的散点图。最后,根据各个水环境监测点对应映射点与四类水质标准值对应映射点之间的位置关系,判定各水质监测数据属于第几类海水水质,从而建立了近海水环境质量的综

合评价模型^[54-56]。

4.3.3 非线性映射分析的映射结果

以天津市近岸海域 2002 年 5 月（枯水期）、8 月（丰水期）、10 月（平水期）的水质监测数据为例，利用非线性映射分析方法进行映射变换。首先指定映射误差，本文设定允许误差为 0.001；其次要设定被映射空间的维数，为了使映射结果能显示在平面上，增加映射结果的可视性，本文将维数设定为 $L=2$ 。整个映射过程借助于 DPS 数据处理系统软件完成^[57]，其中：四类不同水质标准值对应的映射值见表 4-1，20 个水环境监测点在 2002 年 5 月（枯水期）的水质监测值和映射值见表 4-2，映射误差为 $E=0.01335$ 。将表 4-1 和表 4-2 中的 24 个映射值坐标绘于平面坐标上，各映射点位置关系见图 4-1（其中各映射点在平面上 x, y 值只用来衡量各点间的相对位置，并无实际意义）。

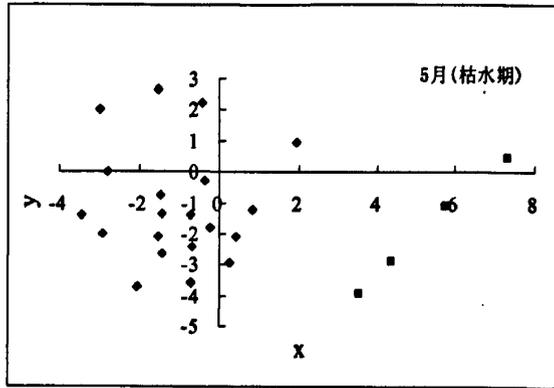
表 4-1 近海水环境质量评价指标标准及映射值

	DO	DIP	DIN	COD	石油类	六六六	铜	锌	铅	镉	汞	x	y
第一类	6	0.015	0.2	2	0.05	0.001	0.005	0.020	0.001	0.010	0.00005	3.51840	-3.94641
第二类	5	0.03	0.3	3	0.05	0.002	0.010	0.050	0.005	0.050	0.00020	4.37583	-2.89555
第三类	4	0.03	0.4	4	0.30	0.003	0.050	0.100	0.010	0.100	0.00020	5.72608	-1.11523
第四类	3	0.045	0.5	5	0.50	0.005	0.050	0.500	0.050	0.100	0.00050	7.34342	0.49302

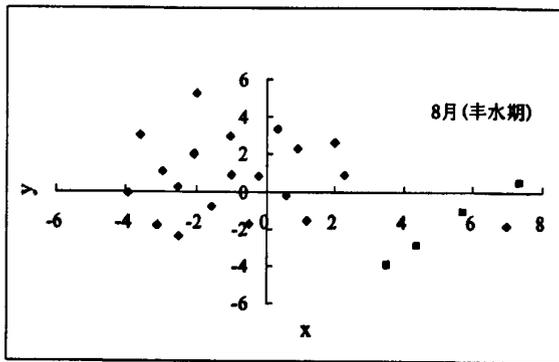
表 4-2 2002 年 5 月（枯水期）水质监测值及对应的映射值

监测点	水质监测值											映射值	
	DO	DIP	DIN	COD	石油类	六六六	铜	锌	铅	镉	汞	x	y
1	8.40	0.008	0.181	1.88	0.025	0.24	1.4	12	4.62	0.21	0.02	0.81014	-1.22585
2	6.70	0.058	0.265	1.76	0.015	0.36	1.9	27	14.73	0.22	0.04	-2.97426	2.01654
3	8.80	0.022	0.405	1.72	0.025	0.24	1.9	47	7.31	0.21	0.07	-1.54564	2.63674
4	8.63	0.035	0.130	1.32	0.028	0.20	0.6	31	7.36	0.24	0.02	-0.35555	-0.30245
5	6.05	0.079	0.074	1.44	0.037	0.20	0.3	19	6.45	0.20	0.02	1.94662	0.95910
6	7.20	0.016	0.473	1.04	0.059	0.24	1.2	27	12.56	0.21	0.08	-0.42418	2.23701
7	7.20	0.003	0.061	0.96	0.020	0.38	0.7	28	11.1	0.24	0.04	-2.90399	-1.98979
8	7.75	0.003	0.090	0.96	0.027	0.20	1.1	27	4.65	0.33	0.02	-0.23646	-1.80490
9	7.20	0.001	0.164	1.00	0.030	0.18	1.2	18	3.64	0.30	0.02	0.40225	-2.08333
10	8.25	0.001	0.094	0.72	0.024	0.20	0.9	16	5.31	0.28	0.04	-0.69430	-2.41538
11	8.40	0.001	0.063	0.92	0.086	0.24	0.5	20	7.24	0.21	0.04	-1.44324	-2.65203
12	7.60	0.008	0.072	1.08	0.024	0.26	0.7	32	8.92	0.18	0.02	-1.49245	-0.76647
13	7.60	0.007	0.102	0.84	0.023	0.22	0.8	5	5.74	0.24	0.02	0.24790	-2.95269
14	8.95	0.005	0.119	0.72	0.020	0.22	0.8	32	7.21	0.22	0.02	-1.45055	-1.36419
15	7.60	0.009	0.067	0.68	0.024	0.24	0.5	27	5.28	0.28	0.04	-1.55763	-2.09006
16	8.65	0.005	0.205	0.80	0.017	0.22	1.7	17	9.31	0.40	0.06	-2.79474	0.00820
19	7.30	0.003	0.127	1.16	0.031	0.16	0.9	21	7.92	0.29	0.04	-0.69877	-1.39293

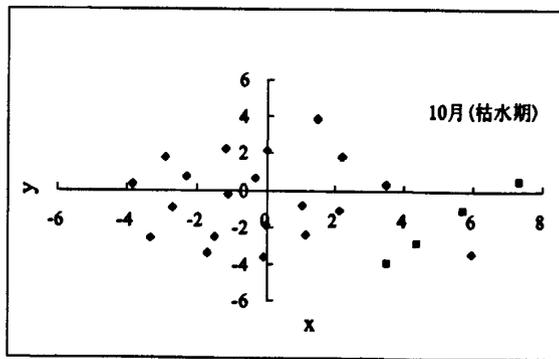
监测点	水质监测值											映射值	
	DO	DIP	DIN	COD	石油类	六六六	铜	锌	铅	镉	汞	x	y
20	7.45	0.016	0.121	0.60	0.023	0.28	0.2	27	13.9	0.34	0.05	-3.41647	-1.39569
21	9.20	0.002	0.131	0.80	0.024	0.24	0.1	17	4.01	0.27	0.02	-0.72760	-3.57370
22	9.00	0.002	0.215	0.80	0.019	0.24	0.1	20	7.92	0.41	0.02	-2.04742	-3.70430



(a)



(b)



(c)

图 4-1 各映射点位置相互关系

从图 4-1 可以清楚的看到，采用非线性映射分析方法进行近海水环境质量综合评价的好处是能够直观地将高维水质信息在低维空间中得到近似图形。然后根据各映射点之间的位置关系，可以对上述水期的水质监测数据进行综合评价，评价结果见表 4-3。

表 4-3 基于非线性映射方法的近海水环境质量综合评价结果

监测点	所属功能区	枯水期		丰水期		平水期	
		评价结果	是否达标	评价结果	是否达标	评价结果	是否达标
1	二	—	达标	—	达标	—	达标
2	三	—	达标	四	不达标	四	不达标
3	二	—	达标	—	达标	—	达标
4	二	—	达标	—	达标	—	达标
5	三	—	达标	—	达标	—	达标
6	四	—	达标	—	达标	—	达标
7	四	—	达标	—	达标	—	达标
8	三	—	达标	—	达标	—	达标
9	四	—	达标	—	达标	—	达标
10	—	—	达标	—	达标	—	达标
11	二	—	达标	—	达标	—	达标
12	二	—	达标	—	达标	—	达标
13	二	—	达标	—	达标	—	达标
14	—	—	达标	—	达标	—	达标
15	二	—	达标	—	达标	—	达标
16	—	—	达标	—	达标	—	达标
19	二	—	达标	—	达标	—	达标
20	—	—	达标	—	达标	—	达标
21	三	—	达标	—	达标	—	达标
22	—	—	达标	—	达标	—	达标

4.4 小结

在本章内容中，结合非线性映射理论和 DPS 数据处理系统软件，提出了近海水环境质量综合评价模型。以天津市近岸海域为例，首先，对环境保护部门提供的环境监测数据分析，依据以往水环境质量评价模型对水质评价指标的选择和定量水质指标监测数据的可得性，选择 COD、溶解氧等 11 个水质指标组成近海水环境质量综合评价的指标体系；然后，根据非线性映射的原理，借助 DPS 数据处理系统将 11 维水质信息映射到二维平面上，实现对多维水质信息的可视化研究；最后，根据水环境监测点监测数据的映射值与海水水质评价标准值对应映

射值之间的位置关系,对近海水环境质量进行综合评价。具体来讲,本文有以下主要创新点:

(1) 首次采用非线性映射分析方法对近海水环境质量进行分析,建立了近海水环境质量综合评价的新模型,成功实现了对整个研究海域水环境质量的准确评价。

(2) 本文提出的基于非线性映射理论的近海水环境质量综合评价新模型克服了以往聚类方法的不足,也克服了以往水环境质量评价模型的不足,最大限度的将尽可能多的水质信息纳入评价中。该模型依靠 DPS 数据处理系统软件,避开了复杂的模型计算,实现起来特别简单,非常有利于近海水环境管理中的应用,有非常显著的理论意义和实用价值。

第五章 近海水环境容量的研究

5.1 概述

了解有机物等近海水质参数空间分布特征是进行近海水环境评价的重要前提,也是治理其污染的基础,而如何利用环境管理部门掌握的水质监测数据对整个管理的近岸海域作出令人满意的估值,是环境管理中的一项难题,也是计算近海水环境容量的基础,而如何实现近海水环境容量的准确计算是进行污染物总量控制的理论依据。以往研究在确定近海水环境容量或纳污能力时,与近海水环境质量的预测和模拟一样,同样多是以高维水动力学方程配合近年来越来越多的机理性研究的成果(营养物质的转化模型、生态系统的演变模型等),首先对污染物在近岸海域的输移、扩散和转化规律进行研究,之后根据各种污染物汇入的途径和强度、近海水体的物化、生物特性,确定近海污染物的平衡浓度场,与功能区或水环境监测点的水环境质量目标进行比较,从而得到研究海域还能够容纳或需要削减多少污染物,也就能够确定在保证近海水环境质量达标的情况下陆源的最大允许排放量。

但是,天津市水环境监测数据具有时间间隔较长(从平水期到下一年的枯水期有7个月之久)、空间跨度大(整个研究海域只有20个水环境监测点,而且大部分集中在近海河流入海口处)等特点,不能满足上述机理性研究的基础数据要求;另一方面,由于各种扰动因素的积累,使得有机物背景值和陆源污染物在近海水体中的演变显得规律性和稳定性很差,所以机理性研究不太适合于近海水环境管理的实际需要。然而,近海有机物空间连续多重分形分布特征是客观存在的,以往的空间插值方法没能很好的利用近海有机物空间分布兼有随机性、确定性和区域结构性这一特征。为了解决上述问题,本文对现已在很多领域成功应用的分形理论进行了深入的研究,立足于环境监测数据,在第二章分析近海水污染物空间分布具有连续多重分形分布特性的基础上,提出了近海有机物等水质参数浓度估值研究的新方法。并在此基础上,根据质量守恒原理,通过确定天津市近岸海域的最不利区域,提出了计算近海水环境容量的新方法。

5.2 分形插值介绍

现实生活中很多事物并不具有光滑性,而是具有分形特性,如山脉、地形、岩石等都是分形曲面的典型例子。在欧氏几何中,由于采用多项式函数、样条函数等建立各种实物的表面模型,因此模型具有局部光滑性。所以传统的方法只能在一定精度意义下描述客观世界中的事物,并不能反映它们的本质。分形几何为我们描述客观世界中的事物提供了有力的工具,用随机方法生成的分形曲面,可以直观和逼真地再现现实生活中的景物。分形插值是美国数学家 M.F.Bamsley 于 1986 年首次提出来的^[58],他利用一组仿射变换产生的迭代函数系统生成分形插值函数,该函数对平面上一组节点进行插值。该方法给出了拟合数据的一种新思想,能够较好地刻划自然界中存在的大量不光滑的曲线和曲面,并且广泛应用于各种数据可视化、模型重建和仿真等问题^[59]。此方法已被用于生成各种二维分形图形,但在三维的情形下,由于计算复杂、子域边界连续性等原因,其应用还相当有限。目前有两种方式可将平面分形插值函数扩展到三维空间。

在对近海有机物空间分布进行研究时,我们只知道部分水环境监测点的水质监测数据,即已知 COD 分形曲面上的部分信息,这就需要通过这部分信息和特征,拟合出 COD 分形曲面的整体形态,从而对整个研究海域 COD 分形体的整体进行研究。这就需要用到分形曲面插值的理论和方法,通过分形插值方法来构造研究海域有机物浓度空间分布的迭代函数系 (Iterated Function System, IFS),也就是通过分形重构的方法构造出研究海域的有机物浓度空间分布曲面。

5.3 分形曲面插值的数学模型

令 $I = [a, b]$, $J = [c, d]$; 设 $D = I \times J = \{(x, y) : a \leq x \leq b, c \leq y \leq d\}$, 以 Δx 、 Δy 为步长,将 D 剖分为网格^[60]:

$$a = x_0 < x_1 < \dots < x_N = b, c = y_0 < y_1 < \dots < y_M = d \quad (5-1)$$

给定一组网格点上的数据 $(x_n, y_m, z_{n,m})$, $n = 0, 1, \dots, N$; $m = 0, 1, \dots, M$ 。欲构造二元分形插值函数 $f: D \rightarrow R$, 且满足: $f(x_n, y_m) = z_{n,m}$, $n = 0, 1, \dots, N$; $m = 0, 1, \dots, M$ 。

令 x 方向的压缩变换为:

$$\phi_n(x) = a_n x + b_n \quad (5-2)$$

y 方向的压缩变换为:

$$\varphi_m(y) = c_m y + d_m \quad (5-3)$$

且满足条件 $\phi_n(x_0) = x_{n-1}$, $\phi_n(x_N) = x_n$, $\varphi_m(y_0) = y_{m-1}$, $\varphi_m(y_M) = y_m$, 可得:

$$a_n = \frac{(x_n - x_{n-1})}{(x_N - x_0)} \quad (5-4)$$

$$b_n = \frac{(x_{n-1}x_N - x_nx_0)}{(x_N - x_0)} \quad (5-5)$$

$$c_m = \frac{(y_m - y_{m-1})}{(y_M - y_0)} \quad (5-6)$$

$$d_m = \frac{(y_{m-1}y_M - y_my_0)}{(y_M - y_0)} \quad (5-7)$$

z 方向的压缩变换为:

$$F_{n,m}(x, y, z) = e_{n,m}x + f_{n,m}y + g_{n,m}xy + a_{n,m}z + k_{n,m} \quad (5-8)$$

式中: $a_{n,m}$ ($n \in \{1, 2, \dots, N\}, m \in \{1, 2, \dots, M\}$) 为垂直比例因子, 决定分形插值曲面分形维数(粗糙程度)的自由参数, 且满足 $0 \leq a_{n,m} \leq 1$ 。

由式(5-3)可得:

$$g_{n,m} = \frac{[z_{n-1,m-1} - z_{n-1,m} - z_{n,m-1} + z_{n,m} - a_{n,m}(z_{0,0} - z_{N,0} - z_{0,M} + z_{N,M})]}{(x_0y_0 - x_Ny_0 - x_0y_M + x_Ny_M)} \quad (5-9)$$

$$e_{n,m} = \frac{[z_{n-1,m-1} - z_{n,m-1} - a_{n,m}(z_{0,0} - z_{N,0}) - g_{n,m}(x_0y_0 - x_Ny_0)]}{(x_0 - x_N)} \quad (5-10)$$

$$f_{n,m} = \frac{[z_{n-1,m-1} - z_{n-1,m} - a_{n,m}(z_{0,0} - z_{0,M}) - g_{n,m}(x_0y_0 - x_0y_M)]}{(y_0 - y_M)} \quad (5-11)$$

$$k_{n,m} = z_{n,m} - e_{n,m}x_N - f_{n,m}y_M - a_{n,m}z_{N,M} - g_{n,m}x_Ny_M \quad (5-12)$$

令 $W_{n,m}(x, y, z) = (\phi_n(x), \varphi_m(y), F_{n,m}(x, y, z))$ ($n \in \{1, 2, \dots, N\}, m \in \{1, 2, \dots, M\}$), 那么就定义了一个迭代函数系, $F_{n,m}(x, y, z)$ 为分形插值函数 f 的隐函数。

5.4 插值结果

上述提出的分形曲面插值的数学模型针对的是矩形区域, 对插值点的位置要求比较苛刻, 而本文研究的天津市近岸海域形状不是很规则, 水环境监测点位置的布设也比较随机(见图 2-2), 不符合分形曲面插值数学模型的要求, 所以本文只对分形曲面插值方法在近海有机物浓度估值方面的应用做初步探讨, 随着计算机应用技术的发展和分形曲面插值算法的不断完善, 分形理论用于曲面插值的优势将更加显著。实际应用时, 做了一定的简化, 在 MATLAB 软件上编写分形曲面插值算法的程序^[60-61], 根据 20 个水环境监测点的 COD 监测数据进行插值, 得到整个研究海域 COD 的空间浓度分布, 其中 2000 年 8 月(丰水期)和 2002 年 10 月(平水期)的插值结果分别见图 5-1 和图 5-2(其中:(a)图表示 COD 初始值所形成的表面;(b)图表示 COD 的分形插值曲面)。

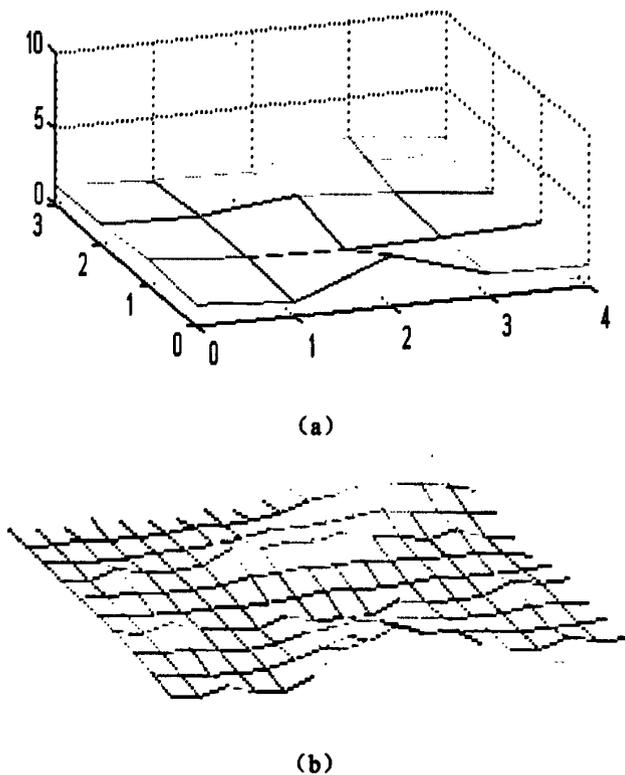


图 5-1 2000 年 8 月的 COD 的插值结果

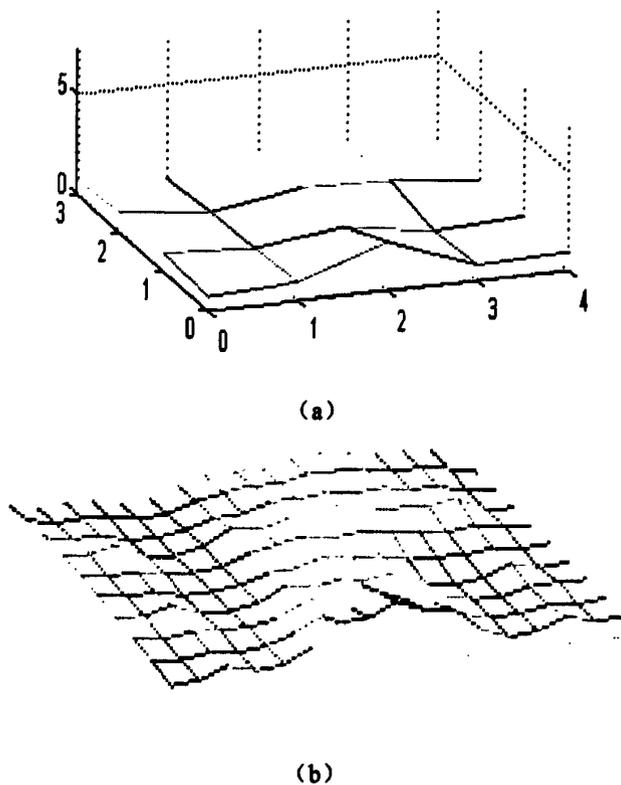


图 5-2 2002 年 10 月的 COD 空间插值结果

5.5 近海水环境容量的计算

5.5.1 最不利区域的确定

根据 5.4 的插值结果, 我们便可以计算天津市近岸海域的水环境容量了。在计算水环境容量时, 首先要根据天津市近岸海域的功能分区确定不同海域对水质的要求(功能区域的划分见图 5-3), 有些区域的污染可能较为严重, 但是该功能区域对水质要求较低, 这样仍然可以满足近海水环境管理的要求; 相反, 某些区域水质较好, 但是这部分区域是重点保护区域而对海水水质要求较高, 而可能出现水质要求不达标现象。所以我们首先确定近海水污染的最不利区域, 即最容易超标的区域。参考以往的研究^[26], 本文认为, 距离现时越近的数据对于环境管理的参考价值越大, 而较早的监测数据对于决策的制定也有一定的意义。所以按照以下步骤确定最不利区域。

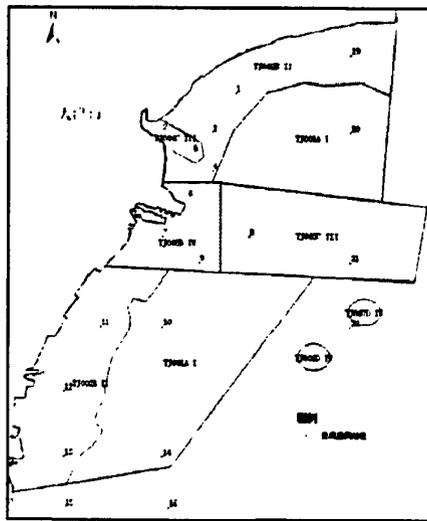


图 5-3 天津市近岸海域水环境功能区图

(1) 给每个水期的插值结果赋予不同的权重, 设定权重与时间早晚成线性关系, 根据分形时间序列的长程相关性, 在本文研究的 1996 年至 2002 年的监测数据之间均具有相关性, 所以确定从 1996 年枯水期到 2002 年平水期的权重分别为 $1/21$ 到 1 等距增长。

(2) 提出各水期中每个矩形多水期污染指数 P 的计算方法如下:

$$P = \sum_{i=1}^{21} \left(\frac{C_i}{S} \times \alpha_i \right) \quad (5-13)$$

式中: C_i 为 i 水期的 COD 浓度; S 为 COD 水质标准值; α_i 为 i 水期的权重。

(3) 按照上述计算方法计算 20 个水环境监测点的多水期污染指数 P ，计算结果见表 5-1，将具有最高污染指数 P 的水环境监测点 2 所在的矩形区域定义为整个研究海域的最不利区域。而事实上根据第四章近海水环境质量综合评价的结果可知，只有水环境监测点 2 出现过水质超标的现象。

表 5-1 20 个水环境监测点的多水期污染指数 P

监测点位	P	监测点位	P	监测点位	P	监测点位	P
1	1.471	2	1.946	3	1.535	4	1.618
5	1.241	6	1.003	7	0.880	8	1.192
9	0.765	10	1.917	11	1.252	12	1.322
13	1.526	14	1.572	15	1.108	16	1.435
19	1.647	20	1.696	21	0.750	22	1.528

5.5.2 近海水环境容量的确定

本文根据质量守恒原理，依据上面确定的最不利区域为立足点，确定影响水环境容量的各个参数的最不利状态，通过分形曲面插值计算在最不利状态下整个研究海域的水环境容量，即陆源最大允许排放量。为了简化问题，本文将整个研究海域的水深定义为 5 米，并认为 COD 浓度在纵向上分布是均匀的。这样，本文按照下列方法计算天津市近岸海域的水环境容量^[26]。

(1) 确定最不利区域的面积

天津市近岸海域的面积为 3712.72 平方公里，根据各个水环境监测点所在区域的矩形数量可以计算出具有不同污染指数区域的面积，其中最不利区域的面积为 77.348 平方公里。

(2) 计算整个研究海域的 COD 容纳量

通过对各个水期 COD 进行空间分形插值后，整个研究海域被分成了 144 个矩形区域。根据每个水期插值结果，可以得到每个矩形区域上的 COD 浓度的平均值，从而计算出每个水期整个研究海域的 COD 容纳量 (T_i)。这样也就确定了各水期之间整个研究海域容纳的 COD 量的变化 (ΔT_i)。

(3) 确定每个水期之间的陆源排放量

根据分形时间序列的长程相关性，在求取多水期污染指数 P 时，本文考虑了 1996 年至 2002 年期间的水质监测值，但是本研究只收集到 2000 年至 2002 年期间天津市近岸海域主要入海河流的陆源排放量数据，所以在此确定每个水期之间的陆源排放量时只考虑近期的值。根据水利部门提供的资料，独流减河、永定新

河在 2000 年至 2002 年期间的入海流量为 0, 所以, 只统计北塘排污河、蓟运河、大沽排污河、海河的各水期的入海流量和 COD 浓度, 确定每个水期之间陆源 COD 的排放总量 (S_i)。

(4) 确定每个水期由其它原因引起的 COD 变化量 (残差量 O_i):

$$O_i = \Delta T_i - S_i \quad (5-14)$$

如果 $O_i > 0$, 表示其它原因给研究海域的 COD 带来了增量; 反之亦然。然后, 按照最不利区域的面积比例将 O_i 分配给最不利区域 O_{wi} 。

(5) 根据各水期的插值结果, 得出最不利区域的 COD 平均浓度以及最不利区域每个水期的 COD 的含有量 T_{wi} , 从而得出每个水期之间最不利区域 COD 的变化量 ΔT_{wi} , 然后根据下式确定由陆源排放引起的最不利区域 COD 的变化 S_{wi} 。

$$S_{wi} = \Delta T_{wi} - O_{wi} \quad (5-15)$$

这样, 计算每个水期到达最不利区域的陆源排放占总陆源排放的比例。

(6) 根据以上计算的结果, 确定以下几个最不利量:

(i) 最不利区域最容易超标的水期, 即最不利区域浓度最高的水期, 得出其距离超标还能容纳多少 COD (w_1);

(ii) 陆源排放到达最不利区域的最大比例 (w_2);

(iii) 月均最小陆源排放强度 (w_3);

(iv) 根据上面 3 个最不利量, 得到最不利条件下的陆源排放强度 (w_4)。

$$w_4 = w_3 + w_1/w_2 \quad (5-16)$$

(7) 由于海水水质标准采用的是 COD_{Mn} , 即碱式高锰酸盐指数, 所以陆源排放同样采用高锰酸盐指数, 但是地表水的管理以 COD_{Cr} 为考察指标, 这样还需要确定 COD_{Cr} 和 COD_{Mn} 之间的换算关系。研究表明, 当在一定的水污染时期, 水质类型较稳定时, COD_{Cr} 和 COD_{Mn} 之间应该存在非常显著的线性关系。所以, 最后根据天津市环境保护局提供的天津市地表水 COD_{Mn} 和 COD_{Cr} 的监测数据, 确定二者之间的换算关系如下:

$$COD_{Cr} = 4.34 \times COD_{Mn} + 38.61 \quad (5-17)$$

最终确定陆源的最大允许排放量为 18763.54 吨/年。该环境容量值只包括北塘排污河、蓟运河、大沽排污河以及海河的入海 COD 量。而根据以往的研究^[26], 采用 GIS 和地统计学方法计算的天津市近海水环境容量为 37578 吨/年, 两种研究方法的不同之处在于, 本文在计算多水期污染指数 P 以确定最不利区域时, 将所有的水期, 即 1996 年至 2002 年的水质监测数据都纳入了考虑, 而原来的研究方法只考虑了 1999 年至 2002 年的水质监测数据, 所以确定的最不利区域有所不同; 另一方面, 两种研究方法采用的空间插值方法不同, 本研究对研究区域做了一定的简化, 从而导致计算结果有了一定的差距。

详细的计算表格见表 5-2。

5.6 小结

在本章内容中，结合分形理论、分形曲面插值算法和 MATLAB 软件，提出了依据环境监测数据的近海有机物浓度估值和近海水环境容量计算的新方法。首先，在第二章分析近海有机物空间分布具有连续多重分形分布特征的基础上，依据分形理论建立近海有机物浓度估值的分形曲面插值算法模型；然后，借助于 MATLAB 的编程和作图功能编写分形曲面插值算法，对近海有机物空间分布进行插值，从而得到整个研究海域的有机物空间估值；最后，根据质量守恒原理，计算在最不利状态下整个研究海域的水环境容量。具体来讲，本文有以下主要创新点：

(1) 首次采用分形曲面插值算法对近海有机物空间分布进行插值，成功实现了对于整个研究海域的有机物浓度的空间估值和近海水环境容量的计算，扩展了分形理论的应用领域，也充分挖掘了监测数据的有用信息。

(2) 本文提出的分形曲面插值算法有效的克服了其它插值方法不具备刻画近海有机物空间变异特征的不足，随着分形曲面插值算法的不断完善和计算机技术的发展，该方法将成为一种全新而有效的计算近海水环境容量方法的基础，该方法完全依据环境监测数据，避开了复杂的模型计算，有非常显著的理论意义和实用价值。

表 5-2 近海水环境容量计算表

年	月	整个研究海域					最不利区域						
		平均 浓度 mc	容纳量 T	陆源排 放量 S	月均陆 源排放 量 Spm	残差量 O	月均残 差量 Opm	平均 浓度 mc _w	容纳量 T _w	变化量 ΔT_w	月均变 化量 ΔT_{wpm}	月均残 差分配 量 O _{wpm}	月均陆 源引起 变化量 S _{wpm}
2000	5	0.967	17947.18	0.00			1.068	413.038					
2000	8	1.363	25295.75	2335.11	778.37	-5013.46	-1671.15	1.792	693.038	279.99	93.33	34.82	58.51
2000	10	1.095	20326.51	975.31	487.65	5944.55	2972.27	1.328	513.591	-179.45	-89.72	-61.92	-27.80
2001	5	0.880	16341.65	5168.35	738.34	9153.21	1307.6	0.968	374.364	-139.23	-19.89	-27.24	7.35
2001	8	0.897	16650.41	2664.26	888.09	2355.5	785.170	1.424	550.718	176.35	58.78	-16.36	75.14
2001	10	1.167	21668.27	776.41	388.21	-4241.44	-2120.72	2.040	778.949	238.23	119.12	44.18	74.94
2002	5	0.917	17022.95	3934.84	562.12	8580.16	1225.74	1.456	563.093	-225.86	-32.26	-25.54	-6.72
2002	8	0.791	14687.34	2169.01	723.00	4504.61	1501.54	1.248	482.651	-80.44	-26.81	-31.28	4.47
2002	10	0.711	13191.52	856.95	428.47	2352.77	1176.39	1.328	513.591	30.94	15.47	-24.51	39.98

第六章 结束语

近海水环境容量的计算是进行近海水污染总量控制的依据和前提,并由此确定陆域和海域可控的污染物的最大排放量。而近海水质变化规律的分析、预测和模拟是进行水环境容量计算的基础,以往在研究近海水污染的规律时,多以机理性研究为主,即以二维或三维水动力学方程为基础,配合物理、化学、生物等模型,并将这些模型组合在一起,力图模拟出各种污染物在近岸海域的时空变化规律,并在此基础上通过考虑研究海域的环境标准,确定水环境容量和相应的总量控制方案。本文通过对复杂系统理论的研究及相关软件的应用,提出了一套完整的完全基于近海水环境监测数据的近海水质分析、预测、评价、有机物浓度估值和近海水环境容量计算的新方法。

6.1 研究结果及结论

本文首次尝试采用复杂系统理论中的分形理论和非线性映射分析用于近海水环境的研究,所做主要工作及研究成果为:

(1) 针对近海水质 COD 和 IN 时间序列数据具有监测频率低、时间跨度大、监测数据前后关系很不明显,使研究的单个水环境监测点的单个水质指标的变化规律变得难以把握等问题。本文采用分形理论中的重标极差分析方法(R/S)对 COD 和 IN 时间序列进行分析,通过计算不同水环境监测点 COD 和 IN 时间序列的 Hurst 指数 H , 计算结果显示其 Hurst 指数 H 均在 0.80 左右,表明近海水质时间序列的变化具有明显的分形特征。该部分内容原理简单,实用性强,为近海水质变化规律的分析 and 预测提供了理论和决策依据,具有较高的理论意义和实用价值。

(2) 针对我国近海水污染物空间分布特征难于辨析的问题,而普通统计方法用于描述和刻划近海水污染物空间分布规律时,并不考虑样品的空间分布和统计特征随空间度量尺度的变化性,不具备刻划异常值的功能。本文通过对多重分形理论的研究,提出了一种新的近海水污染物空间分布特征的分析方法。以天津市近岸海域 20 个水环境监测点为例,随机选择两个不同水期的 COD 监测数据进行分析,采用多重分形矩方法计算 COD 空间分布的多重分形维数谱函数 $f(\alpha)$ 。

计算结果显示其 $\alpha(q)-f(\alpha)$ 曲线均为连续不对称的上凸曲线, 表明近海水污染物在空间上的分布并不是简单的正态或对数正态分布, 而是连续多重分形分布, 近海水污染物高浓度部分的奇异性较低浓度部分大, 平水期的奇异性较丰水期大。应用表明, 该分析结果与天津市近岸海域实际情况相一致, 这些成果的取得为辨析近海水污染物的空间分布特征, 并进一步采用基于多重分形理论的算法模型计算近海水环境容量提供了理论依据。

(3) 针对我国近海水水质污染呈现的年周期性和自相似性特点, 本文通过对分形插值算法进行研究, 在分析近海水水质 COD 时间序列具有分形特征的基础上, 建立了近海水环境质量的分形预测模型。首先, 根据分形拼贴定理, 由基于仿射变换的分形插值方法求取各历史时间阶段水质变化的迭代函数系, 然后根据近海水水质变化的年周期性, 对上述求得的迭代函数系加权求和, 得到预测年份水质变化的统计意义上的迭代函数系, 建立分形预测模型, 最后应用随机迭代算法求得预测年份水质变化曲线的吸引子, 对近海水水质进行预测, 预测结果显示平均预测误差为 24.4%。应用表明, 该方法预测精度较高、实用性强, 能够为近海水环境管理提供决策支持。

(4) 本文通过对非线性映射理论及相关技术进行深入学习, 借助 DPS 数据处理系统, 建立了近海水水质综合评价的新模型, 成功实现了近海监测数据的有效评价和高维水质信息的二维显示。首先, 通过非线性映射理论对高维水质监测数据进行非线性变换, 得到二维映射值并将其描绘在平面坐标上; 其次, 根据映射点之间的位置关系进行分类, 从而实现对近海水质的准确评价。应用表明, 该方法可操作性强、不受监测数据类型的限制、可视化程度高, 能准确的评价高维监测数据反映的水质信息, 可为近海水域污染状况的辨析和采取相应的措施提供决策支持。

(5) 本文通过对分形曲面插值算法的研究, 在分析近海水污染物空间分布具有连续多重分形分布特征的基础上, 成功的实现了近海有机物的浓度估值, 并在此基础上提出了一种完全依据环境监测数据的近海水环境容量计算的新方法。首先, 在分析近海有机物空间分布具有多重分形特征的基础上, 根据分形曲面插值算法模型, 在 MATLAB 上编制相关程序对近海有机物浓度进行空间插值, 从而得到整个研究海域的有机物空间分布情况。然后, 依据质量守恒原理, 立足于最不利区域, 求得了天津市近岸海域在最不利条件下的水环境容量, 即陆源最大允许排放量。通过实例分析表明, 该方法避免了复杂的模型计算, 实用性强, 有利于在近海水环境管理中的试用。

6.2 存在的不足和建议

本文通过对分形理论的研究和应用,首次将分形理论引入到近海水环境的研究中,尽管分形理论已广泛用于其它领域并取得了显著的成果,这些成果的取得为本研究提供了依据和借鉴意义,本文尽管达到了预期的目标,实现了依据环境监测数据进行近海水质特征的分析、预测、评价、有机物浓度的估值和水环境容量的计算等多个成果。但是,本文毕竟为初次尝试,还存在一定的不足,在今后各种条件逐渐成熟的基础上,可以对论文中提出的理论和方法进行进一步完善:

(1) 由于本文所用到的水环境监测数据具有监测频率低、空间跨度大等特点,单个水环境监测点的水质时间序列不是严格的等间隔时间序列,这给水质的准确预测带来了很大困难;由于水环境监测点空间跨度大,所监测的水质信息在反映整个研究海域水污染物的空间变异特征和用于有机物浓度估值方面也存在一定缺陷。如果今后收集到更多的基础数据,可以就近海水污染的控制和管理开展更加广泛的研究;

(2) 分形理论用于近海水质的分析和预测等方面研究时,本文只是以天津市近岸海域为例,尽管取得了显著的效果,但是还需要对其它海域进行类似的研究以确定近岸海域水质变化分形特征的普遍性和分形理论用于近海研究的适用性。

(3) 分形曲面插值算法用于近海有机物浓度估值和近海水环境容量的计算研究时,只针对的是矩形区域,而实际的近岸海域的水环境监测点布置很难达到这个要求,所以本文对研究海域做了一定的简化,只做了尝试性的应用。随着计算机技术的发展和分形曲面插值算法的不断完善,分形插值的优势将更加明显,这也需要科技工作者的不断努力和 innovation。

(4) 本文提出的分形方法具有大小尺度规律可以互用的特点,可以为机理性研究进行数据补充。机理性研究与本文提出的基于环境监测数据的宏观研究具有各自的优缺点,应该互相借鉴两种研究的经验和成果,充分发挥各自的优点,突破在近海水环境研究中的一些瓶颈。

(5) 分形理论是一门新兴的数学分支,同时与计算机编程技术结合紧密,所以用于实际领域的研究时,需要学习者具有很高的数学理论水平和计算机编程技术。本人在研究中也遇到了很大的困难和障碍,所以后续研究者需要充分发挥团队协作精神。

参考文献

- [1] HELCOM. Third Periodic Assessment on the State of the Marine Environment of the Baltic Sea, 1989-1993. Background Document, Baltic Sea Environment Proceedings No. 64B, Helsinki Commission, Baltic Marine Environment Protection Commission, 1996, 252.
- [2] C. Sheppard (Ed.) . Regional Chapters: Europe, The Americas and West Africa. Seas at the Millennium: an Environmental Evaluation, Pergamon, Amsterdam, Oxford, 2000, 1: 934.
- [3] C. Sheppard (Ed.) . Regional Chapters: The Indian Ocean to the Pacific. Seas at the Millennium: an Environmental Evaluation, Pergamon, Amsterdam, Oxford, 2000, 2: 920.
- [4] Shahidul Islam, Masaru Tanaka. Impacts of pollution on coastal and marine ecosystems including coastal and marine fisheries and approach for management: a review and synthesis, Marine Pollution Bulletin, 2004, 48: 624-649.
- [5] 国家海洋局, 2003 年中国海洋环境质量公报, 人民网 <http://www.People.com.cn>, 2004.
- [6] 天津市海洋局, 2003 年天津海洋质量公报, 国家海洋局北海分局网站 <http://www.ncsb.gov.cn>, 2004.
- [7] I. D. James. Modelling pollution dispersion, the ecosystem and water quality in coastal waters: a review [J]. Environmental Modelling & Software, 2002, 17 (4) : 363-385.
- [8] Nina Gjerde Winther, Geir Evensen. A hybrid coordinate ocean model of shelf sea simulation [J]. Ocean Modelling, 2006, 13: 221-237.
- [9] Minna Kuusisto, Jorma Koponen, Juha Sarkkula. Modelled phytoplankton dynamics in the Gulf of Finland [J]. Environmental Modelling & Software, 1998, 13 (5-6) : 461-470.
- [10] Qiuwen Chen, Arthur E. Mynett. Modelling algal blooms in the Dutch coastal waters by integrated numerical and fuzzy cellular automata approaches [J]. Ecological Modelling, 2006, 99: 73-81.

- [11] Dong S. Ko, Paul J. Martin, Clark D. Rowley, Ruth H. Preller. A real-time coastal ocean prediction experiment for MREA [J]. *Journal of Marine Systems*, 2007, 2: 22-34.
- [12] Michael D.krom, Hava Hornung, Yuval Cohen. Determination of the environmental capacity of Haifa Bay with respect to the input of mercury [J]. *Marine Pollution Bulletin*, 1990, 21 (7) : 349-354.
- [13] 孙涛, 陶建华. 波浪作用下渤海湾近岸海域污染物的输移扩散规律[J]. *海洋与湖沼*, 2004, 35 (2): 110-119.
- [14] 沈永明, 郑永红, 吴修广. 近岸海域污染物迁移转化的三维水质动力学模型[J]. *自然科学进展*, 2004, 14 (6): 694-699.
- [15] 牛志广, 张宏伟, 辛志伟. 基于 log-logistic 概率分布的近海水质组合预测方法研究[J]. *系统工程理论与实践*, 2006, 5: 111-116.
- [16] 娄安刚, 王学昌, 吴德星等. 胶州湾大沽河口邻近海域海水水质预测[J]. *海洋环境科学*, 2002, 21 (1): 52-56.
- [17] 栗苏文, 李红艳, 夏建新. 基于 Delft3D 模型的大鹏湾水环境容量分析[J]. *环境科学研究*, 2005, 18 (5): 91-95.
- [18] 郭森, 韩保新, 杨静等. 纳污海域水环境热量计算与总量分配研究[J]. *环境科学与技术*, 2006, 29 (增刊): 19-22.
- [19] 程声通, 陈疏龄. *环境系统分析*[M]. 北京: 高等教育出版社, 1994.
- [20] 郭劲松, 王红, 龙腾锐. 水资源水质评价方法分析与进展[J]. *重庆环境科学*, 1999, 21 (6): 1-3.
- [21] 谢宏斌. 南湖富营养化的人工神经网络评价[J]. *广西科学院学报*, 1999, 15 (1): 29-32.
- [22] 王晓鹏, 朱明奎. 河流水质评价的多级模糊模式识别模型及在湟水流域的应用[J]. *沈阳化工学院学报*, 2001, 15 (2): 147-150.
- [23] 罗定贵, 王学军, 后立胜. 模糊综合评价—聚类复合模型在地下水水质评价与分区中的应用[J]. *地理与地理信息科学*, 2003, 19 (6): 77-79, 88.
- [24] 罗定贵, 郭青, 王学军. 地表水质评价的径向基神经网络模型设计[J]. *地理与地理信息科学*, 2003, 19 (5): 77-81.
- [25] 牛志广, 张宏伟. 灰色关联分析用于近海水污染区域性特征研究[J]. *天津大学学报*, 2006, 39 (增刊): 311-315.
- [26] 牛志广. 近岸海域水环境容量的研究[D]. 天津: 天津大学, 2004.
- [27] 郝柏林. *混沌与分形—郝柏林科普文集*[M]. 上海: 上海科技出版社, 2004.
- [28] Mandelbrot B B. *The fractal geometry of nature* [M]. San Francisco, 1982: 9-17.

- [29] Mandelbrot B B. Fractal: Form, Chance and Dimension [M]. San Francisco, 1977: 15-23.
- [30] 王美荣, 金志琳. 分形理论及其应用[J]. 菏泽师范专科学校学报, 2004, 26 (4): 51-56.
- [31] 谢和平, 张永平, 宋晓秋等. 分形几何—数学基础与应用[M]. 重庆: 重庆大学出版社, 1991.
- [32] Hurst H E. Long-term storage of reservoirs [J]. Trans. Amer. Civil Eng, 1951, 116: 770-808.
- [33] 倪化勇, 刘希林. 自然灾害发生时间序列的分形特征及 R/S 分析[J]. 自然灾害学报, 2005, 14 (6): 37-41.
- [34] 范英, 魏一鸣. 基于 R/S 分析的中国股票市场分形特征研究[J]. 系统工程, 2004, 22 (11): 46-51.
- [35] 马军海. 复杂非线性系统的重构技术[M]. 天津: 天津大学出版社, 2004. 75-76.
- [36] Nawrocki D. R/S analysis and long term dependence in stock market indices [J]. Managerial Finance, 1995, 21 (7) : 78-92.
- [37] Cheng Q. Multifractality and spatial statistics [J]. Computer & Geosciences, 1999, 25 (9) : 949-962.
- [38] Cheng Q, Agterberg F P. Multifractal modeling and spatial point processes [J]. Math Geol, 1995, 27: 831-845.
- [39] 天津市环境保护局. 天津市碧水工程研究报告[R]. 天津: 天津市环境保护局, 2003.
- [40] 谢淑云, 鲍征宇. 多重分形与地球化学元素的分布规律[J]. 地质地球学, 2003, 31 (3) : 97-102.
- [41] 文战久, 刘洪臣, 高星. 地球物理和地球化学异常的多重分形分析与分解 [J]. 地球物理学进展, 2007, 22 (3) : 972-978.
- [42] Evertsz C J G, Mandelbrot B B. Multifractal measures (Appendix B) [A]. In: Peitgen H-O, Jurgens H, Saupe D. Chaos and fractals [C]. New York: Springer Verlag, 1992: 922-953.
- [43] 姜志强. R^2 空间上分布数据的多重分形维数谱计算和分形特征提取[J]. 小型微型计算机系统, 2004, 25 (10) : 1844-1846.
- [44] 陶涛, 刘遂庆. 基于分形理论的需水量预测方法[J]. 同济大学学报 (自然科学版), 2004, 32 (12): 1647-1650.
- [45] 李水根. 分形[M]. 北京: 高等教育出版社, 2004.

- [46] Xie H, Sun H. The study on bivariate fractal interpolation functions and creation of fractal interpolation surface [J]. *Fractals*, 1997, 5 (4) : 625-634.
- [47] David S M, Moson H H. Using iterated function systems to model discrete sequences[J]. *IEEE Trans on Signal Processing*, 1999, 40 (7) : 1724-1734.
- [48] 薛万磊. 电力负荷分形预测方法研究[D]. 哈尔滨: 哈尔滨工业大学, 2004.
- [49] 孙博文. 分形算法与程序设计[M]. 北京: 科学出版社, 2004.
- [50] 廖广兰, 史铁林, 刘世元. 基于非线性映射的特征提取技术研究[J]. *华中科技大学学报 (自然科学版)*, 2004, 32 (12): 54-56.
- [51] 杨志勇, 李秋胜, 张礼平等. 用非线性映射方法对多因素影响下结构阻尼的研究[J]. *华中科技大学学报 (城市科学版)*, 2003, 20 (2): 24-27.
- [52] 徐志节, 杨杰, 王猛. 利用非线性降维方法预测膜蛋白类型[J]. *上海交通大学学报*, 2005, 39 (2): 279-283.
- [53] 岳云华, 冉清红. 基于非线性映射的西部区域经济发展协调性的分形研究[J]. *绵阳师范学院学报*, 2006, 25 (2): 69-73.
- [54] 唐启义, 冯明光. DPS 数据处理系统[M]. 北京: 科学出版社, 2007: 701-703.
- [55] 迟国泰, 刘轶芳, 余方平. 基于非线性映射分析的期货逼仓风险判定模型及其应用[J]. *系统工程理论与实践*, 2006, 5: 1-11.
- [56] Ebiefung, Aniekan. Nonlinear mappings associated with the generalized linear complementarity problem [J]. *Mathematical Programming*, 1995, 69: 255-268.
- [57] Lichtenberg A J, Ujihara A. Application of nonlinear mapping theory to commodity price fluctuations [J]. *Journal of Economic Dynamics & Control*; 1989, 13 (2) : 225-247.
- [58] Barnsley MF. Fractal functions and interpolation [J]. *Constructive Approximation*, 1986, 2 (2) : 303-329.
- [59] 王梦, 金文标. 基于函数迭代系统的3-D分形插值算法[J]. *计算机应用*, 2006, 26 (11) : 2701-2703.
- [60] 孙洪泉. 分形插值曲面的MATLAB程序[J]. *苏州科技学院学报 (工程技术版)*, 2006, 19 (4) : 18-21.
- [61] 张艳. 基于OpenGL的分形山三维模拟生成技术[J]. *计算机技术与发展*, 2007, 17 (8) : 243-246.

附录1 近海水质预测主要程序

```
CZhangDoc* pDoc = GetDocument();
    ASSERT_VALID(pDoc);
    // TODO: add draw code for native data here
    double a[NUM]={0.2,0.2,0.2,0.2,0.2};
    double b[NUM]={0,0,0,0,0};
    double c[NUM]={0.1736,-0.2551,-0.0649,-0.0972,0.2435};
    double d[NUM]={0.2,0.2,0.2,0.2,0.2};
    double e[NUM]={0,0.2,0.4,0.6,0.8};
    double g[NUM]={0.2474,0.4972,0.3182,0.3294,0.3084};
    double p[NUM]={0.2,0.2,0.2,0.2,0.2};
    srand(time(NULL));
    double x,y,newx,newy,rn,total;
    int dx,dy,i;
    CPen * pOldPen ;
    CPen BluePen(PS_SOLID,1,RGB(0,0,255));
    CPen GreenPen(PS_SOLID,1,RGB(100,100,100));
    pOldPen = pDC->SelectObject(&GreenPen);
    char* str = (char*)malloc(sizeof(char)*10);
    x=y=newx=newy=rn=total = 0 ;
    double *ave = (double*)malloc(sizeof(double)*AVENUM);
    double *count = (double*)malloc(sizeof(double)*AVENUM);
    double *sum = (double*)malloc(sizeof(double)*AVENUM);
    FILE* fp = fopen("date2.txt","w");
    for(int m = 0 ; m < AVENUM;m++)
    {
        ave[m] = 0.0;
        count[m] =0.0;
        sum[m] = 0.0;
```

```
}
for(i=0;i<YBASE/BLOCK;i++)
{
    pDC->MoveTo(XBASE,i*BLOCK);
    pDC->LineTo(XMAX,i*BLOCK);
    sprintf(str,"%0.1f", (YBASE/BLOCK-i)/(float)(ZOOM/BLOCK));
    pDC->TextOut(XBASE-20,i*BLOCK-BLOCK/2,str);
}
for(i=YBASE/BLOCK;i<YMAX/BLOCK;i++)
{
    pDC->MoveTo(XBASE,i*BLOCK);
    pDC->LineTo(XMAX,i*BLOCK);
    sprintf(str,"%0.1f", (YBASE/BLOCK-i)/(float)(ZOOM/BLOCK));
    pDC->TextOut(XBASE-20,i*BLOCK-BLOCK/2,str);
}
for(i=0;i<(XMAX-XBASE)/BLOCK;i++)
{
    pDC->MoveTo(XBASE+i*BLOCK,0);
    pDC->LineTo(XBASE+i*BLOCK,YMAX);
    sprintf(str,"%0.1f", i/(float)(ZOOM/BLOCK));
    pDC->TextOut(XBASE+i*BLOCK-BLOCK/2,YMAX+(i%2*10),str);
}
pDC->SelectObject(&BluePen);
pDC->MoveTo(XBASE,0);
pDC->LineTo(XBASE,YMAX);
pDC->MoveTo(0,YBASE);
pDC->LineTo(XMAX,YBASE);
for( i=0;i<100000;i++)
{
    int k=0;
    m=rand()/32767.0;
    total=p[0];
    while(total < m)
    {
```

```
        k++;
        total+=p[k];
    }
    newx=a[k]*x+b[k]*y+e[k];
    newy=c[k]*x+d[k]*y+g[k];
    x=newx;
    y=newy;
    if(i>200)
    {
        dx=(int)(x*ZOOM+XBASE);
        dy=(int)(YBASE-y*ZOOM);
        pDC->SetPixel(dx,dy,RGB(255,0,0));
        if((x <=0.0001)||((x>=0.1764 && x <=0.1766)||((x>=0.2940 &&
x<=0.2942)||((x>=0.7058 && x<= 0.706)||((x>=0.8823 && x<= 0.8825)||((x>=0.9999) )
        {
            fprintf(fp,"x[%d] = %f , y[%d]=%f \n",i,x,i,y);
        }
        if(x<0.00005)
        {
            count[0]++;
            sum[0]+=y;
        }else if(x>=0.19995 && x <=0.20005)
        {
            count[1]++;
            sum[1]+=y;
        }else if(x>=0.39995 && x<=0.40005)
        {
            count[2]++;
            sum[2]+=y;
        }else if(x>=0.59995 && x<= 0.60005)
        {
            count[3]++;
            sum[3]+=y;
        }else if(x>=0.79995 && x<= 0.80005)
```

```
        {
            count[4]++;
            sum[4]+=y;
        }else if( x>=0.99995 )
        {
            count[5]++;
            sum[5]+=y;
        }
    }
}
for(m = 0 ; m < AVENUM ; m++)
{
    ave[m] = sum[m] / count[m] ;
    fprintf(fp, "sum[%d]=%f,count[%d]=%d,ave[%d]=%f\n",m,sum[m],m,(int)count[
m],m,ave[m]);
}
pDC->SelectObject(pOldPen);
fclose(fp);
free(str);
free(ave);
free(sum);
free(count);
```

附录2 分形曲面插值算法的主要MATLAB程序

```
z=[0.72 0.92 2.44 0.80 0.96;0.96 0.72 1.24 0.56 0.8;1.08 0.56 1.28 1.12 0.60;0.72
0.32 0.88 1.0 0.40];
[m,n]=size(z);
x=0:100:(n-1)*100;
y=0:100:(m-1)*100;
subplot(2,1,1);
mazz=max(max(z))*3;
dip=40;dir=340;
meshz(z);
view(dir,dip);
axis([0 n-1 0 m-1 0 mazz]);
nn=(n-1)*(n-1);
mm=(m-1)*(m-1);
x1=x(n)-x(1);
y1=y(m)-y(1);
a=(x(2:n)-x(1:n-1))/x1;
b=(x(n)*x(1:n-1)-x(1)*x(2:n))/x1;
c=(y(2:m)-y(1:m-1))/y1;
d=(y(m)*y(1:m-1)-y(1)*y(2:m))/y1;
cz=z(1,1)+z(m,n)-z(1,n)-z(m,1);
cm=x(1)*y(1)+x(n)*y(m)-x(n)*y(1)-x(1)*y(m);
bz1=z(1,1)-z(1,n);bz2=x(1)*y(1)-x(n)*y(1);
bm=x(1)-x(n);
dz1=z(1,1)-z(m,1);dz2=x(1)*y(1)-x(1)*y(m);
dm=y(1)-y(m);
dn=ones(m-1,n-1);
dn=dn*af;
cc=(z(1:m-1,1:n-1)-z(1:m-1,2:n)-z(2:m,1:n-1)+z(2:m,2:n)-dn*cz)/cm;
```

```

bb=(z(1:m-1,1:n-1)-z(1:m-1,2:n)-dn*bz1-cc*bz2)/bm;
dd=(z(1:m-1,1:n-1)-z(2:m,1:n-1)-dn*dz1-cc*dz2)/dm;
kk=(z(2:m,2:n)-bb*x(n)-dd*y(m)-dn*z(m,n)-cc*x(n)*y(m));
for j=1:m-1;
    for j0=1:m;
        yv=c(j)*y(j0)+d(j);
        jj=(j-1)*(m-1)+j0;
        for i=1:n-1;
            for i0=1:n;
                ii=(i-1)*(n-1)+i0;
                xv=a(i)*x(i0)+b(i);
                zt=bb(j,i)*x(i0)+dd(j,i)*y(j0);
                zz(jj,ii)=zt+cc(j,i)*x(i0)*y(j0)+dn(j,i)*z(j0,i0)+kk(j,i);
            end;
        end;
    end;
end;
mm=(m-1)*(m-1)+1;
nn=(n-1)*(n-1)+1;
subplot(2,1,2);
meshz(zz);
view(dir,dip);
axis([0 nn-1 0 mm-1 0 mazz]);
axis off;
fwd=fopen('sp','w');
for j=1:mm;
    for i=1:nn;
        fprintf(fwd,'%8.4f,zz(j,i));
    end;
    fprintf(fwd,'\n');
end;
fclose(fwd);

```

发表论文和科研情况说明

发表的论文：

- [1] 分形理论用于天津市近海水质预测方法的研究(水科学进展 已收录)；
- [2] 基于相空间重构理论的天津市近海水质混沌特性研究(海洋科学进展 已收录)；
- [3] 基于分形理论的城市日用水量预测方法研究(已投稿至天津大学学报)；
- [4] 基于多重分形理论的近海水污染物空间分布特性研究(已投稿至系统工程学报)；
- [5] 基于重标极差分析方法的近海水水质数据分形特性研究(已投稿至系统工程理论与实践)；
- [6] 基于分形理论的近海有机物空间变异特性和浓度估值的研究(已投稿至天津大学学报)。

参与的科研项目：

- [1] 参与了天津市科技发展计划项目“生态校园建设标准、模式及智能化管理信息系统研究”的研究工作；
- [2] 参与了国家自然科学基金资助项目“城市供水管网漏损及优化维护的研究”的研究工作；
- [3] 参与了天津市自然科学基金项目“天津市近海水环境容量计算与污染物总量控制的研究”的研究工作。

致 谢

时光荏苒，转眼间硕士阶段的学习即将结束，博士阶段的学习也将开始。在这二年的时间里，除了我个人的努力，还得到了很多人的帮助，才使我顺利完成了硕士学位论文的研究工作。在论文成文之际，谨向每一位关心我和帮助我的人衷心地道一声：谢谢！

首先，将最真挚的谢意献给我的导师张宏伟教授，本论文的工作是在您悉心指导下完成的，您在论文的写作过程中给予了热情的帮助和指导，您严谨而又开放的治学思想，认真踏实的工作态度将成为我今后学习和工作的榜样。

同时，最要感谢的是牛志广副教授对我论文研究工作的指导，您不仅教授我科研和写作的基本方法，在学习和生活方面也教给我很多做人的道理，在此向牛志广老师表示衷心的感谢。

在研究过程中，理论部分得到了柴立和老师、冯繁一同学的指导，程序调试过程中得到了张刚同学、南开大学的谭泽超同学的帮助，在此表示衷心的感谢。

课题组赵新华教授、田一梅副教授、刘志强副教授、刘洪波副教授对于我的科研工作和论文都提出了许多的宝贵意见，在此表示衷心的感谢。

在实验室工作及撰写论文期间，王晨婉、张蕊、陈健敏、裴同英、邓特刚、李海等同学对我论文的研究工作给予了热情帮助，在此向他们表达我的感激之情。

此外，本论文的研究过程中得到了天津市自然科学基金项目（07JCYBJC07200）的资助和支持。

最后，我要特别的感谢我的父母、两位姐姐及关心帮助我的所有亲友们，是你们用最朴实的爱、理解和支持才得以使我顺利的完成硕士学业。