海洋环境监测数据库创新设计与建设*

杨　翼[1,2]，路文海[2]，付瑞全[2]，向先全[2]，刘　捷[2]

（1. 天津科技大学　天津　300222；2. 国家海洋信息中心　天津　300171）

摘　　要：国家、海区、省（自治区、直辖市）、市、县等五级海洋环境监测机构每年在执行多种海洋环境监测任务时，将产生海量的用于反映全海域环境质量、生态系统、环境状况等类别的监测数据，针对这多源异构的监测数据，文章在总结前人研究经验的基础上，对海洋环境业务监测工作进行梳理、数据进行整理分析，创新设计构建了能满足各种监测数据存储管理的数据库结构，开发了监测数据库管理系统，从而实现全国海洋环境监测数据统一高效管理，以及监测数据的共享及应用服务奠定基础。

关　键　词：海洋环境监测；监测数据；数据库结构；数据库管理系统

1　引言

海洋环境监测是海洋环境保护的“耳目”和“尺子”，是一切海洋保护工作的基础，其根本任务是掌握海洋环境污染的状况和长期变化趋势，评价海洋环境质量，为制定国家海洋环境保护政策提供基础依据[1-2]。在中央和沿海各级政府的支持下，现已形成了国家、海区、省（自治区、直辖市）、市、县五级海洋环境监测网络体系，在我国管辖的 300 万 km^2 海域开展岸基站、船舶、飞机、卫星、浮标、雷达等多种手段监测工作[3]。随着时间的推移，由海洋环境监测网所产生的各类监测数据类型不断丰富，数量也呈几何级增长和更新，海洋环境监测数据库建设迫在眉睫。

为加强监测数据存储管理与服务共享，有关单位在各自的研究或应用领域积极开展了监测数据库设计与构建尝试[4-5]。本文在前人研究的基础上，针对监测数据管理和应用需求，对全国海洋环境业务化监测工作进行梳理，对所获取的数据进行整理和分析，设计通用性较强的监测数据库结构，开发监测数据库管理系统，来对多源异构监测数据进行统一、标准化管理，向应用系统提供标准统一、高效的数据支持，从而促进监测数据的共享与应用服务。

2　监测数据来源

2.1　监测任务

为全面掌握我国管辖海域的环境状况及变化趋势，科学、客观地评价海洋环境质量，实现对环境质量与趋势清楚、对主要的污染源清楚、对潜在环境风险清楚，我国每年都会开展任务多样的海洋环境监测工作，以实现不同的海洋环境监测目标。

以《2012 年海洋监测工作任务》[6]为例，实施的监测工作共包括环境状况、环境风险、环境监管与公益服务等四大类 25 项监测任务（表 1）。其中，环境状况是针对全海域环境质量而开展的监测，旨在把握我国海域环境质量状况及演变趋势；环境风险是针对灾害多发区域及突发性污染事故区域而开展的监测，旨在识别我国海洋环境风险源；环境监管是针对敏感海域或重要海洋功能区而开展的监测，旨在分析把握敏感海域或重要功能区环境变化与影响及入海污染源状况；公益服务是针对与公众健康密切相关的海域而开展的监测，旨在保障公众的健康安全。

* **基金项目**：我国海洋环境监测评价体系优化与综合服务平台开发（201005014）；疏浚物和污水污泥倾废监测、管理技术研究与应用示范（201105010）.

表 1　2012 年海洋环境监测任务分类

任务类别	任务序号	监测任务	监测要素	监测范围
环境状况	1	海洋生物多样	水文气象；海水水质；沉积物质量；沉积物粒度；浮游植物；浮游动物；底栖生物；潮间带生物；鱼卵仔鱼；红树林群落；红树林栖息地；海草床群落；珊瑚礁底栖藻类；珊瑚礁群落；珊瑚礁鱼类	管辖海域全覆盖
	2	海水	水文气象；海水水质；水动力	管辖海域全覆盖
	3	海洋沉积物	沉积物质量；沉积物粒度	管辖海域全覆盖
	4	二氧化碳	海水水质；二氧化碳分压；气溶胶	管辖海域全覆盖
	5	海洋大气	干沉降；湿沉降	13 个海洋大气监测站
环境风险	6	溢油应急	水文气象；海水水质；沉积物质量；生物质量；浮游植物；浮游动物；底栖生物；潮间带生物；鱼卵仔鱼	事件周边海域
	7	蓬莱 19－3 溢油事故	水文气象；海水水质；沉积物质量；沉积物粒度；生物质量；浮游植物；浮游动物；底栖生物；鱼卵仔鱼	事件影响范围海域
	8	大连油污染事件	水文气象；海水水质；沉积物质量；沉积物粒度；生物质量；浮游植物；浮游动物；底栖生物；鱼卵仔鱼；潮间带生物	事件影响范围海域
	9	危险化学品污染	水文气象；海水水质；沉积物质量；生物质量；浮游植物；浮游动物；底栖生物；潮间带生物；鱼卵仔鱼	事件影响范围海域
	10	海洋放射性污染	水质放射性；沉积物放射性；生物放射性；大气放射性	10 个核电站与 12 个重点海域
	11	赤潮（绿潮）	水文气象；海水水质；沉积物质量；沉积物粒度；赤潮生物；赤潮贝毒；生物质量；增养殖区调查；赤潮跟踪	21 个赤潮监控区及应急区域
	12	海水入侵	海水入侵	29 个重点入侵区域
	13	土壤盐渍化	土壤盐渍化	21 个重点土壤盐渍化区域
	14	重点海岸侵蚀	海岸侵蚀	6 个重点侵蚀岸段
	15	长江口贫氧区	海水水质；水文气象	长江口区域
环境监管	16	海洋倾倒区	海水水质；水文气象；沉积物质量；沉积物粒度；生物质量；浮游植物；浮游动物；底栖生物；鱼卵仔鱼；水深	正在使用的海洋倾倒区
	17	海洋油气勘探开发区	海水水质；沉积物质量；生物质量；浮游植物；浮游动物；底栖生物	39 个油气田群
	18	海洋工程建设项目	水文气象；海水水质；沉积物质量；沉积物粒度；浮游植物；浮游动物；底栖生物；鱼卵仔鱼；潮间带生物；水动力	34 个重点建设工程
	19	海洋保护区	海水水质；沉积物质量；沉积物粒度；浮游植物；浮游动物；底栖生物；潮间带生物；保护区调查；保护对象；珊瑚礁群落；红树林群落	38 个国家级海洋保护区
	20	陆源入海排污口及邻近海域	排污口水质；海水水质；沉积物质量；生物质量；底栖生物	438 个入海排污口
	21	入海江河	江河水质	76 条入海江河
	22	海洋垃圾	海面垃圾；海滩垃圾；海底垃圾	36 处重点海域
公益服务	23	海水浴场	水文气象；海水水质；景观要素；沙滩要素	31 个海水浴场
	24	滨海旅游度假区	景观要素；水文气象；海水水质；沙滩要素	17 个滨海旅游度假区
	25	海水增养殖区	海水水质；沉积物质量；生物质量；浮游植物；底栖生物；赤潮贝毒；增养殖区调查	54 个海水增养殖区

2.2　监测要素

从表 1 可以看出，各监测任务由于承担的监测目标不同，所涉及的监测要素也就各不相同。从监测数据的应用角度考虑，可以将监测要素分为以下 3 类。

（1）环境质量要素，有评价标准，可以评价环境状况等级的要素，如海水水质、沉积物质量、生物质量等。

（2）生态系统要素，用于分析评价生物种类、分布、数量及变化状况，生态系统结构组成的要素，如浮游植物、浮游动物、珊瑚礁群落等。

（3）环境状况要素，用于反映海洋生态环境构成或人为活动影响或其他相关的要素，即除一、二类以外的监测要素，如水文气象、水深、二氧化碳分压等。各监测要素由不同的监测参数组成，同一监测要素的不同监测参数值就组成了用于分析评价的监测要素数据，所以监测参数值是组成监测数据的最小单元。

3　监测数据库创新设计

3.1　监测数据库设计原则

监测数据库设计采用“面向数据”的数据库设计方法与 E－R 建模技术，除坚持分类存储、建立关联、降低冗余、清除潜在的不一致性等基本原则外[7]，结合监测数据自身的特点，还将考虑以下原则。

（1）数据的共享性原则。兼顾监测数据的应用分析与共享利用需求，可方便不同应用系统对数据调用。

（2）数据的独立性原则。数据的组织与析取主要依据监测数据性质和被利用价值所决定，不受单项应用所左右。

（3）数据结构的可扩充性原则。在应用需求发生变化时，数据结构应能适应新数据类和新数据项的扩充。

（4）数据关联性原则。对由于各监测要素组成的监测数据，必须实现各要素数据及基本信息数据、元数据之间充分利用关联，才能降低数据库冗余、清除不一致性。

（5）数据之间相互约束原则。在不少数据小类之间，均有一些数据项（字段），分属不同的表，描述的对象却可能相同，数据之间有约束关系。

3.2　监测数据库总体设计

监测数据库主要针对通过海洋环境业务化监测所获得的数据来设计。数据库能够以标准的格式对多源、异构、海量的海洋环境监测进行录入和存储，并且能够有效地检索数据，快速响应来自客户端的数据共享请求。监测数据库的总体设计框架如图 1 所示，包括基本信息库、监测要素库和元数据库 3 部分。基本信息库用于反映监测数据的来源地；监测要素库用于存储监测数据，涉及的监测要素有近 40 项；元数据库用于反映监测数据的提供方及质控过程信息。三者通过关键字段相互关联，确保了监测数据信息的紧密性和可约束性。

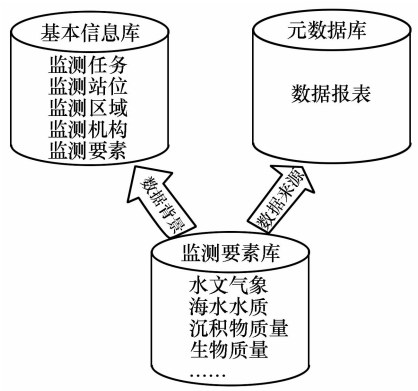

图 1　海洋环境监测数据库总体设计框架

3.3　监测数据库结构设计

数据库分析是数据库结构设计的关键所在，也是数据库程序开发的基础。数据库分析包括两部分：一是概念模型设计，即 E－R 分析；二是逻辑结构设计，即表与字段的分析[8]。在数据库分析的基础上，可能通过数据库管理软件，如 SQL Server、Oracle、Sybase、Access 等，来构建数据库物理结构，满足数据存储需求，实现数据库管理。

（1）数据库概念设计。进行 E－R 分析首先要确定数据库中的各种实体，并分析它们的属性和它们之间的关系，然后画它们的 E－R 图。在监测数据中，有监测站位、监测任务、监测区域等基本信息实体，有海水水质、水文气象、

沉积物质量、浮游植物、浮游动物等近 40 项监测要素实体（表 1），有跟踪数据来源的报表信息，即元数据实体。通过分析各实体之间的关系可以看出，监测要素实体是用于存储监测数据的实体，也是整个数据库建设的主体，通过与基本信息实体、元数据实体关联，可反映数据的空间位置及来源。利用业界排名第一的数据建模和应用设计工具 Sybase 公司的 PowerDesigner[9] 软件开展数据库 E－R 分析。在进行实体关系图设计时，要着重考虑的是实体的属性和主键等因素，考虑同一监测要素要满足多个监测任务的管理需求。

（2）数据库逻辑结构设计。数据库逻辑结构设计即表与字段分析，是建立在数据库 E－R 图的基础上，通过该 E－R 图可以确定数据库中所有的表及其字段。表与字段分析后就可以建立数据库模型。同样利用 PowerDesigner 自带的工具产生表和字段并建立数据库模型，可减轻建库工作量并提高设计质量。各表的字段参数名和字段类型、长度、是否空值以及表的主键、关联字段应在此阶段予以确认。需要说明的是，由于同一监测要素可分配给多个监测任务执行，但每个监测任务中该监测要素的监测参数不一定相同，所以监测要素表字段需要基于监测任务所确定的监测参数，采用去重追加的方法来确定。

（3）监测要素表结构优化分析。监测要素表的构建，可以有两种方式：一是将参数名作为字段，即站位编号＋任务编号＋区域编码＋监测日期＋参数名 1＋参数名 2＋…＋n 的方式来构建的二维表；二是将参数名作为记录，即站位编号＋任务编号＋区域编码＋监测日期＋参数名＋参数值＋参数单位的方式来构建二维表。

比较这两种表结构，可以看出方式二不需要考虑监测要素所涉及的参数，也不需要将各参数单位统一，便于数据存储；但建库的目的不只是存储数据，还要考虑数据的应用，所以方式二在应用方面存在的弊端包括：① 数据参数名与单位需要标准化处理；② 数据分析评价时，需要格式转换，即换成方式一的形式才能直观表达出监测或评价结果，海量数据的格式转换，效率一般极低；③方式二会导致存储的数据记录成倍增长，是方式一的 n 倍，也极易产生数据冗余[10]，相对于方式一来说，参数名与参数单位冗余，假设监测参数有 n 个，2 个冗余，数据冗余度 $r=2/(n+4)$。一般情况下，$n\leqslant 10$，即 $r>14\%$，这种结构冗余度比较大。

方式一在数据入库前，必需先做数据的标准化处理与质控，以保障数据质量；应用时，不需要格式转换，数据记录数和数据冗余也明显优于方式二，所以监测要素表应采用方式一来构建。

（4）数据库物理结构构建。逻辑结构设计是面向用户的，而物理结构设计是面向计算机的。数据库在物理设备上的存储结构和存取方法等就称为数据库的物理结构。本文采用 Microsoft SQL Server 2008 数据库管理系统对监测数据库进行建设。SQL Server 是微软的企业级数据库，它功能强大，图形化管理工具使得在 SQL Server 中管理服务器变得非常容易，可以支持字符型、整型、浮点型等常规数据类型以及影像、音频、视频等二进制大对象（BLOB）[11]。利用数据库管理软件的“新建数据库”和“新建表”两个命令按钮，创建监测数据库表，并设置数据应用分析所需要的视图、存储过程以及触发器、函数等。由于监测数据产生的单位多，涉及的数据量大，为避免监测要素表中监测数据的重复录入，需要对各要素表进行关键字段约束，以水质为例，约束字段包括任务编码、区域编码、站位编号、监测时期、层号。

4 监测数据库管理系统

一般的数据库管理系统可以通过两种方式来访问数据库：通过数据库软件客户端和自主开发的客户端[12]。本文监测数据库管理系统采用 C♯结合 Developer Express[13] 出品的 .NET 窗体构成控件开发客服端。C♯提供了功能强大的数据库管理功能，可快捷、方便，灵活地开发数据库管理系统，而 Developer Express 出品的 .NET 窗体控件，可以给用户提供操作便捷、数据表单显示清晰的工具，满足用户对数据的批量导入、删除、编辑、修改、查询、导出、统计、空间范围查询等功能需求。

数据库建设的目的除满足监测数据的存储管理外，还得让数据看得见、摸得着，实现共享服务。本文利用已构建的监测数据库，在“海洋环境保护管理综合信息系统”中开发了 3 种 WEB 方式的监测数据查询，即任务查询、要素查询、组织单位查询，实现了查询结果显示、及与空间定位的互联互动。

5　结束语

为了满足对全国海洋环境监测数据进行统一高效的管理的应用需求，本文基于 Microsoft SQL Server 2008 数据库管理软件，并针对海洋环境数据量庞大，传统数据存储结构访问效率低的问题，对海洋环境监测数据库进行了创新设计和构建，降低了数据存储冗余，提高了查询性能。目前，监测数据库及共享服务平台已经建成，并通过“海洋环境监测数据传输网”投入业务化运行，为各级海洋行政管理部门、科研人员提供了监测数据共享服务，为各专题应用系统提供基础数据，促进了海洋生态环境保护工作的开展，提高了监测数据的使用价值。

参考文献

[1]　许丽娜，王孝强．我国海洋环境监测现状及发展对策[J]．海洋环境科，2002，22(1)：63－68.

[2]　马春生，潘红，周洪英，等．发展海洋环境监测的意义和作用[J]．科技创新导报，2010(2)：123－124.

[3]　国家海洋局．全国海洋环境监测与评价业务体系“十二五”发展规划纲要[Z]．2011.

[4]　苏天赟，刘海行，刘保华，等．黄海大海洋生态系统数据库设计[J]．海洋科学进展，2008，26(2)：243－250.

[5]　顾军，龚建新，焦念志，等．胶州湾海洋生态环境 GIS 数据库的构筑[J]．海洋科学，2002，26(1)：15－16.

[6]　国家海洋局．2012 年海洋环境监测工作任务[Z]．2012.

[7]　陈锁忠，常本春，黄家柱，等．水资源管理信息系统[M]．北京：科学出版社，2006.

[8]　张奇，李伟松，卫建卫，等．Visual C＃数据库项目案例导航[M]．北京：清华大学出版社，2005.

[9]　张裔智，冯春辉．基于 UML 的仓库管理数据库系统设计[J]．计算机科学，2008，35(6)：295－297.

[10]　刘艳艳，崔松雪，陈勇，等．一种新的海量海洋环境数据库组织结构的研究与实现[J]．中国海洋大学学报，2008，38(4)：667－670.

[11]　赵斌．SQL Server 2008 应用开发案例解析[M]．北京：科学出版社，2009.

[12]　李禾珍，邹峥嵘．惠州市环境基础数据库的设计与实现[J]．计算机与数字工程，2009，37(11)：173－176.

[13]　Tools for Visual Stuidio (2009－04－17) －100% Native. NET technologies[EB/OL]．(2009－04－17)[2013－12－01]．http://www.devexpress.com/Products/NET/Controls/WinForms/Grid/，2009－4－17.