

doi:10.3969/j.issn.1003-2029.2020.05.012

高分辨率中国近海风能资源数据库开发设计

兰志刚¹, 岳磊², 孙洋洲¹, 吴勇虎¹, 郭雪飞¹, 于汀¹

(1. 中海石油(中国)有限公司北京研究中心, 北京 100028; 2. 北京天机数测数据科技有限公司, 北京 100020)

摘要: 构建中国近海风能资源数据库可以为中国近海风电场规划、前期选址、风能资源评估提供强有力的数据支持。本文对国家气候中心基于中尺度气象模式所获得的中国近海高时空分辨率的海上风能资源数据做了基本分析和评价, 论证了利用该数据作为中国近海风能资源数据库数据源的有效性, 给出了中国近海风能资源数据库的设计原则, 采用当前主流的 Spark 技术体系, 结合地理信息系统和网络技术, 给出了该数据库的详细技术架构, 构建了基于地理信息系统的高分辨率中国近海风能资源数据库, 并介绍数据库的技术特点以及主要功能。研究结果表明, 本数据库数据源可靠、架构清晰合理, 可以便捷地实现风能资源数据的检索、查询、统计计算和数据展示及有效存储管理, 为开展中国近海风能资源的管理与评价提供了有效的技术平台。

关键词: 海上风能资源; 数据库; 技术架构

中图分类号: TP311.1; TK89 **文献标识码:** A **文章编号:** 1003-2029(2020)05-0079-07

我国近海风能资源十分丰富。2004 年“我国近海海洋综合调查与评价”专项的研究结果表明, 我国近海 10 m 高度、50 m 等深线以浅海域风能资源的理论装机量可达 8.83×10^8 kW。在海上风电开发决策过程中, 风能资源是海上风电选场址和风机选型的先决条件, 更是直接影响风电场发电量和经济性的关键因素, 决定着海上风电项目的成败。由于海上测风数据匮乏, 当开展风电场资源评价时, 常苦于缺乏有效的风能资源数据而无法着手工作。近年来, 利用中尺度气象模式对海上风场进行模拟计算, 从而获得大范围内较高空间分辨率的风场分布结果, 已成为海上风能资源评估的重要技术手段^[1]。美国马萨诸塞州大学曾运用中尺度动力诊断模式—中尺度大气模拟系统(Mesoscale Atmospheric Simulation System, MASS)对新英格兰州南部海上风能进行评估^[2],

获得了很好的效果; 袁春红等^[3]运用中尺度气象模式(Mesoscale Model5, MM5)对海上风场进行模拟计算, 得到了 8×8 km 分辨率下的风能资源分布结果; 穆海振等^[4]运用空气污染模型(The Air Pollution Model, TAPM)模式对上海沿海风能资源进行评估, 得到了 3 km 分辨率的风能资源分布结果; 龚强等^[5]运用 MM5 模式模拟, 得到了辽宁省沿海 10 km 分辨率的风能资源分布图; 周荣卫等^[6]运用中国气象局开发的风能资源数值模拟评估系统对我国近海风能资源进行数值模拟研究, 得到了我国近海 20 年平均的高分辨率风能资源分布。

如能利用高精度、高分辨率风场数据, 开发构建一套方便高效的中国近海风能资源数据库, 无疑会对开展海上风电场开发前期研究带来极大的便利, 特别是对于开展风能资源的初期快速评

收稿日期: 2020-06-11

基金项目: 中国海洋石油集团公司综合科研项目(CNOOC-KJ QZJC ZY 2019-01); 中海油研究总院基础前瞻及新能源技术探索研究项目(2019-ZYZL-XNY-01); 中海石油(中国)有限公司科技项目(YXKY-SHANGHAI-2020-01)

作者简介: 兰志刚(1963-), 男, 博士, 教授级高级工程师, 主要从事海洋工程及新能源研究。E-mail: lanzhg@cnooc.com.cn

价和风电场的宏观选址可以提供强有力的数据支撑。国家气候中心自 2007 年起,在科技部国家高技术研究发展计划(863 计划)和发改委、财政部全国风能资源详查和评价项目的支持下,开发了风能资源数值模拟评估系统,相继完成了 3 套高时空分辨率的风能资源数据集,为在应用层面构建风能资源数据库奠定了坚实的基础。本项目以国家气候中心 2016 年开发的数据集为基本数据元素,采用当前主流的 Spark 技术体系,依据大数据分析思路,开展高分辨率中国近海风能资源数据库开发设计,构建高分辨率中国近海风能资源数据库。

1 数据源及数据检验

2016 年底,国家气候中心联合清华大学、国家超算无锡中心,利用 2 400 站的地面气象观测资料、169 站的探空气象观测资料、卫星遥感资料以及中尺度数值模式,在对大量实测测风数据进行同化、校验和优化的基础上,历时 13 个月完成了最新中国风能资源大数据—国家气候中心风能资源数据(NCC 3 km),构建了 1995—2016 年全国陆地和近海(离岸 100 km 之内)的逐小时风场时间序列数据,最终得到了高精度风能资源数据集。数据生成过程见图 1。

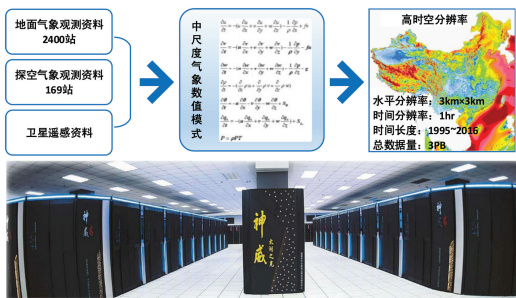


图 1 国家气候中心 3 km 风能资源数据制作流程

模拟得到的主要气象要素包括风速,风向,空气密度,相对湿度等。在此基础上,经统计分析给出了年平均风速、平均风功率密度、风向频率、风速频率、风能方向频率和风速韦布尔(Weibull)分布参数等风能资源参数。该数据的水平分辨率为 3 km,垂直分层超过 80 层,其中 200 m 以下垂直间隔精细至 10 m,涵盖我国陆地和海域以及“一带一路”主要国家和地区,覆盖

面积超过 $2\,500\times 10^4\text{ km}^2$,是国内首套高时空分辨率长年代风能资源数据。图 2 是此数据与美国国家航空航天局天气模式再分析资料(MERRA)以及中国各省实测资料的对比。

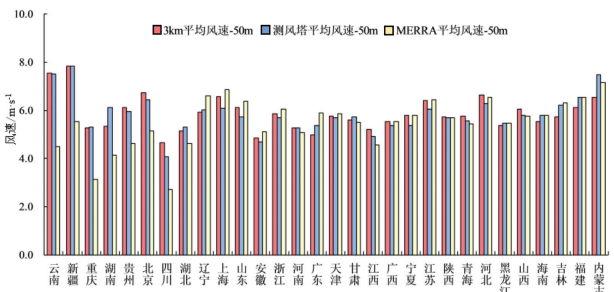


图 2 不同地区 NCC3km 数据和 MERRA 数据的平均风速与实测风速的对比

由图 2 可以看出,NCC 3 km 平均风速与测风塔的实测风速更接近,而 MERRA 风速在云南、贵州、四川等复杂地形省区则存在较大偏差。表 1 是 100 m 高度的模拟年平均风速和已有的 60 个相同高度测风塔的实测年平均风速之间的误差统计。

表 1 模拟年平均风速和实测年平均风速之间的误差统计

相对偏差	测风塔个数	占比 / %	累计占比 / %
< 5%	28	46.7	46.7
5%~10%	20	33.3	80
10%~15%	10	16.7	96.7
> 15%	2	3.3	100

从表 1 中可以看出,在 60 个站位的数据对比中,相对误差小于 5 % 的站位为 28 个,占比为 46.7 %;相对误差在 5 % ~ 10 % 之间的站位为 20 个,占比为 33.3 %,相对误差在 10 % ~ 15 % 之间的站位为 10 个,占比为 16.7 %;相对误差大于 15 % 的站位为 2 个,占比为 3.3 %,可以认定数值模拟的精度较好、数据可信。

2 数据库设计原则

鉴于风能资源数据信息量大,且与地理信息数据密切相关,风能资源数据库必须集数据储存、处理、展示和服务为一体,以地理信息系统和网络技术为基础,实现对数据的存储和数据系统的管理。为此,风能资源数据库的设计应遵循以下原则:

(1) 采用基于 WebGIS 的空间数据管理架构,实现风能资源的地理空间数据管理、空间查询、空间分析以及空间展示功能;

(2) 可以实现风能资源相关数据的检索、查询和数据图谱展示,方便风能资源数据的查询和应用;

(3) 系统具有标准化和规范化的框架设计,提供开发应用程序接口(Application Programming Interface, API),以便于数据使用者方便的接入本系统。而且配置合理、易于调整、监视及控制,以保证系统良好运作;

(4) 平台软件应具有较好的可扩展性,为新增业务、新增资料的灵活、快速扩充和改造提供方便;

(5) 为方便应用,用户和服务器可以分布在不同的地点和不同的计算机平台上,以便能不同的人员,可从不同的地点,以不同的接入方式访问和操作数据。

3 数据库架构

基于以上原则,中国近海风能资源数据库平台架构选择基于 Spark^[7] 技术体系。数据采集、存储、分析各环节的技术架构如图 3 所示。

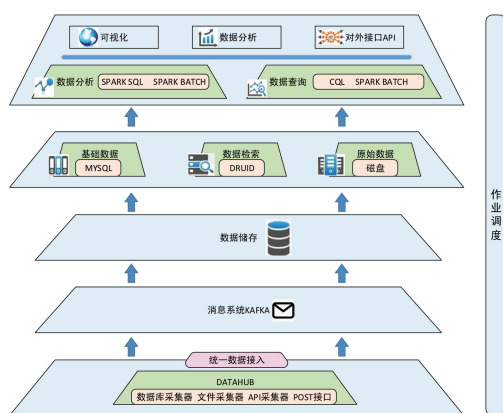


图3 中国近海风能资源数据库系统架构图

系统采用 Kafka^[8] 消息系统对采集任务实现任务调度管理,根据数据的用途采用不同的存储和分析查询方式,基于 Druid 实现应用数据的存储。Druid^[9] 是一项开源、分布式、列存储、实时的数据存储技术,可以为交互式应用提供低延迟的数据导入和查询,方便在大数据集之上做实时统计

分析。在分析查询方面,由于涉及 1995—2016 年共计 22 a 的历史数据,采用基于 Spark SQL 和 Spark Batch 的组件技术实现分析查询。Spark 具有高速、通用和可融合性优势,可以用于批处理、交互式查询、实时流的处理和图计算,不管是在基于内存方面还是在基于磁盘方面,运算速度均很快,可更高效地处理数据流,保证实时查询的效率。数据采集及分析完成后,通过 API 接口方式提供数据可视化。

Spark 架构组件的主要功能如下:

(1) 聚合数据处理:基于 Druid 技术,实现了大数据的实时查询和高容错、高性能分析,提供了交互方式访问数据的能力,优化了存储格式。

(2) 实时查询组件:基于 Spark SQL 技术,使用了自身的语法解析器、优化器和执行器,不仅支持 Hive 数据的查询,同时可实现多种数据源的数据查询。

(3) 批量事件处理:基于 Spark 流技术,将流式计算分解成一系列短小的批处理作业,提高了效率。

(4) 分布式协调服务:基于 Zookeeper^[10] 技术,实现了分布式应用程序协调服务。

数据库整体部署方案如图 4 所示。

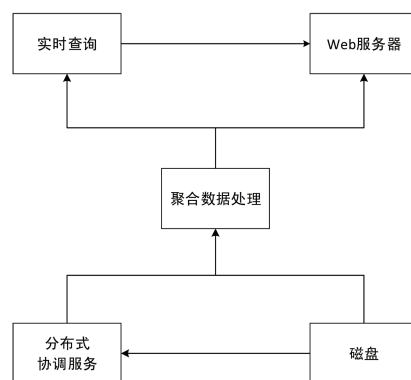


图4 数据库部署架构

3.1 交互架构

数据库交互为浏览器/服务器架构模式(Browser/Server, B/S),即浏览器请求、服务器响应的工作模式。B/S 架构技术是目前国内外运用最为广泛的体系结构。它可以支持较丰富的服务器端语言,并且集成了功能完备的富客户端技术,使得软件表现层的展现更为灵活,更趋近于用户的使用习惯,同时可充分发挥程序结构时效性、

高效性和并发性的长处。其 3 层体系架构的分布式设计,可以很好地满足系统的可扩展性要求。B/S 架构模式的优势是统一了客户端,将系统功能实现的核心部分集中到服务器上,用户工作界面通过浏览器来实现,能够确保不同人员、从不同地点、以不同接入方式(比如局域网、广域网、因特网或企业内部网等)访问和操作共同的数据。

3.2 WebGIS 架构

本数据库的 GIS 架构选用互联网地理信息系统 WebGIS^[11]。WebGIS 是建立在互联网基础之上的、具有 B/S 体系结构的、交互式、分布式动态地理信息系统,拥有图 5 所示的展示层、地图服务层和数据层 3 层结构^[12],能够实现 Internet 环境下的空间信息访问、管理和发布。利用 WebGIS 网络上的任意用户均可通过该引擎获得所需的地理信息并对其进行检索和分析,以实现空间数据的共享和互操作。WebGIS 支持插件扩展,插件涵盖地图应用的各个方面,包括地图服务、数据提供、数据格式、地理编码、路线和路线搜索、地图控件和交互等,具有良好的可扩展性。WebGIS 架构在地理信息的空间分布式获取,地理信息的空间查询、检索和联机处理,空间模型的分析服务以及互联网上资源的共享等方面具有明显技术优势。

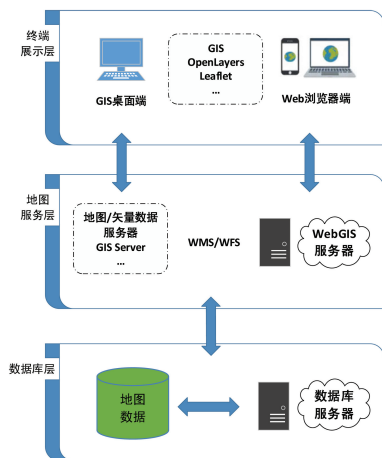


图 5 WebGIS 的三层结构示意图

3.3 后台管理系统架构

本系统的后台管理采用模型—视图—控制器 (Model-View-Controller, MVC) 框架,利用 Velocity^[13] 模板引擎实现前台 Web 端的展示,然后通过程序控制生成静态页面展示给用户。技术路线图如图 6 所示。

其中,视图层面向用户操作,实现风能资源信息的展示、点查询、区域查询、交互分析及页面监控管理。主要技术内容包括:5 代超文本标记语言 (Hyper Text Marked Language 5, HTML5)、3 代层叠样式表 (Cascading Style Sheets 3, CSS3)、jQuery、LeafLet^[14] 和 Velocity。运用 HTML5 可以在移动设备上支持多媒体,更方便于用户与文档的交互;CSS3 是 CSS 技术的模块化升级版本,它提升了系统的视觉效果,增加了系统的可访问性;jQuery 是一个快捷的 JavaScript 框架。它提供了一种简便的 JavaScript 设计模式,进一步优化了 HTML 文档操作、事件处理、动画设计和 Ajax 交互;LeafLet 是用来在页面中创建并操作地图的开源 JavaScript 库,具有开发在线地图的大部分功能,能够在所有主要桌面和移动平台上高效运作,并支持浏览器访问和插件扩展。

业务应用层用于实现系统核心功能。它基于自定义数据查询语言 (DQL),兼容多种数据存储方式,从而实现风能资源点查询、区域查询、历史对比分析以及交互分析算法。

基础层是框架组成的核心部分,是整个框架的基础,主要技术内容包括:Vertx 和 GeoTools 等。Vertx^[15] 是一个基于 Java 虚拟机 JVM 的轻量级、高性能的应用平台,非常适用于最新的移动端后台、互联网、企业应用架构;GeoTools^[16] 是一个构建在国际标准之上的 Java 类库,它提供了很多的标准类和方法来处理空间数据,是开源空间数据处理的主要工具,很多 Web 服务、命令行工具和桌面程序都可以由 Geotools 来实现。

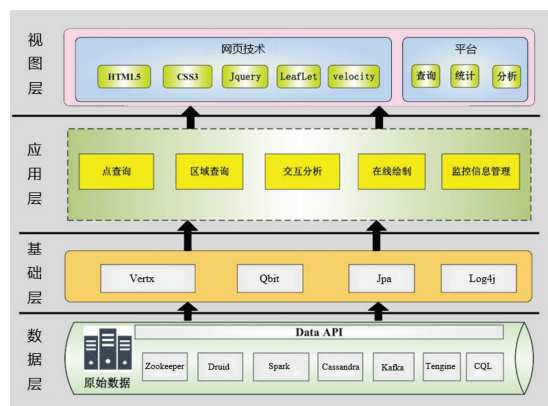


图 6 技术框架图

数据层负责对数据库中的数据进行添加、删除、修改和查询等操作, 并将数据传递给上层的业务层进行处理。本数据库数据层采用 Spark 技术体系实现对风能资源的调度和存储, 它可根据大数据分析需求的变化, 灵活高效地进行分析和运算, 并透明地与原数据交换平台实现数据分析应用整合。数据层管理系统是一种自定义的作业调度系统, 通过采集器 (DataHub) 对多种数据源进行采集和数据整合。数据采集及分析完成后, 通过 API 接口方式提供数据可视化。

4 数据库功能及应用

开发完成的高分辨率中国近海风能资源数据库从操作上看, 其功能主要划分为 4 部分: 点查询、空间分布、数据导出以及用户授权。功能系统结构划分如图 7 所示。

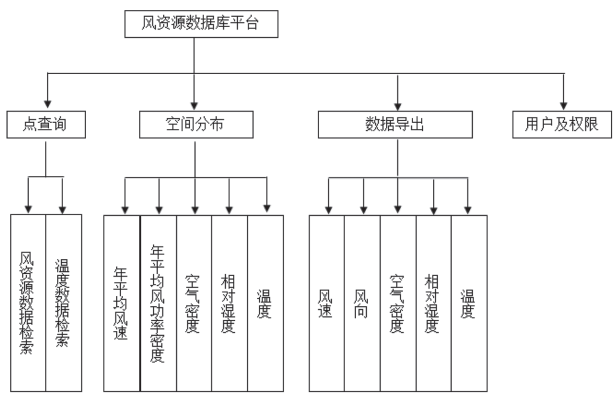


图 7 系统功能结构

数据库可以实现中国近海 100 km 范围内, 任意经纬度的风能资源参数的检索、查询及结果展示, 给出 1995—2016 年逐年小时风速、风向、空气密度、相对湿度、温度, 以及多年平均的 Weibull 参数、风玫瑰、风功率密度。上述数据可以以点查询、空间分布展示、数据导出 3 种方式进行操作。点查询结果分为风能参数信息和时间变化信息两类, 点击鼠标选择地图模型内任意的坐标点, 可查看该点的各类风能参数信息以及风能资源时间变化, 给出风能资源日平均变化、月平均变化和年平均变化等趋势曲线, 也可在整个区域的二维空间上给出风能资源参数的空间分布

展示。图 8 和图 9 是通过某台互联网计算机登录至风资源数据库服务器后, 查询风资源数据得到的结果。图 8 上图是点查询得到的某点位风资源基本参数, 下图是该点位风速年际变化。图 9 上图是点查询得到的中国近海 100 km 内海区的风速分布, 下图是中国近海 100 km 内海区的年平均风功率密度分布。

另外, 通过导出数据功能, 可以将对应的风能资源数据以特定的格式进行下载和导出, 并快速输出各种图表。所有数据能够以 Excel 格式下载到本地磁盘。



图 8 风能资源参数点查询示例

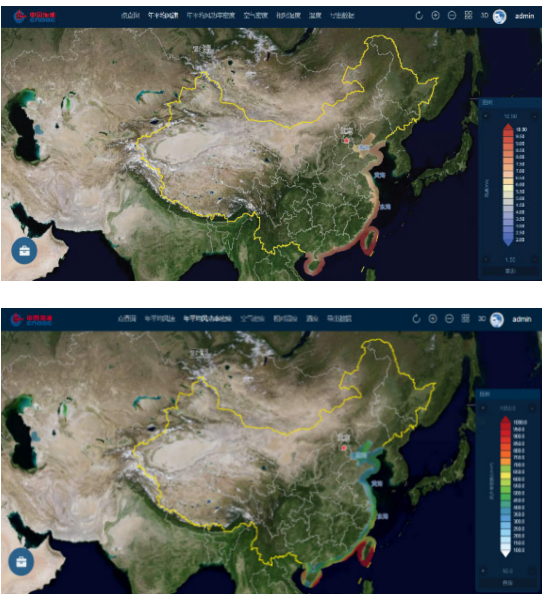


图 9 风能资源参数的二维空间图谱展示

5 结论

利用中尺度气象模式获得大范围、较高空间分辨率的海上风能资源数据, 可以为用户开展风电场规划、前期选址、风能资源评估提供可靠的数据支持, 这已成为海上风能资源评估的重要技术手段。本文规划设计的数据库以国家气候中心2016年开发的风能资源大数据为基础。该数据具有时间序列长、空间分辨率高、数据精度高等特点。本文依据风能资源数据所具有的空间地理信息特征以及信息量大的特点, 结合实际应用需求, 确立了中国近海风能资源数据库的设计原则和技术架构。数据库采用当前主流的 Spark 技术体系, 并有效结合了 WebGIS 技术, 架构清晰合理, 功能强大便捷。其中 Druid 技术的应用, 使系统可以对空间风能资源这类基于时序的大数据进行聚合查询, 实现查询的快速和灵活。B/S 和 WebGIS 架

构可以确保不同的人员、从不同的地点、以不同的接入方式在 GIS 地图上访问和操作具有地理信息特征的风能资源数据, 实现地理信息的空间分布式获取、空间查询、检索和联机处理。后台管理采用 MVC 框架, 增加了系统的视觉效果和可访问性, 并进一步实现了数据分析应用的整合功能。标准化和规范化并带有 API 接口的框架设计, 使数据使用者可以方便地接入本系统。依据上述技术构建的高分辨率中国近海风能资源数据库, 从后期的使用效果来看, 功能完全符合设计要求, 可以方便实现风能资源数据的检索、查询、统计计算和空间分布展示及有效存储管理, 从而为开展风能资源的初步快速评价和风电场的宏观选址提供强有力的数据支撑。

致谢: 感谢国家气候中心和北京天机数测数据科技有限公司为本文风能资源数据库提供技术和资料支持。

参考文献:

- [1] 何晓风, 朱蓉, 周荣卫. 风能资源的气候学数值模拟方法研究 [C]// 工业空气动力学研究进展. 长沙: 中南大学出版社, 2009:124-130.
- [2] Manwell J F, Rogers A L, McGowan J G, et al. An offshore wind resource assessment study for New England [J]. Renewable Energy, 2002, 27(2): 175-187.
- [3] 袁春红, 薛桁, 杨振斌. 近海区域风速数值模拟试验分析 [J]. 太阳能学报, 2004, 25(6): 740-743.
- [4] 穆海振, 徐家良, 杨永辉. 数值模拟在上海海上风能资源评估中的应用 [J]. 高原气象, 2008, 27(S1): 196-202.
- [5] 龚强, 袁国恩, 张云秋, 等. MM5 模式在风能资源普查中的应用试验 [J]. 资源科学, 2006, 28(1): 145-150.
- [6] 周荣卫, 何晓风, 朱蓉, 等. 中国近海风能资源开发潜力数值模拟 [J]. 资源科学, 2010, 32(8): 1434-1443.
- [7] 高彦杰. Spark 大数据处理: 技术、应用与性能优化 [M]. 北京: 机械工业出版社, 2014.
- [8] Apache Software Foundation. Kafka 2.4 Documentation [EB/OL]. [2020-04-15]. <https://kafka.apache.org/24/documentation.html>.
- [9] Apache Software Foundation. Introduction to Apache Druid [EB/OL]. [2020-04-18]. <http://druid.apache.org/docs/latest/design>.
- [10] Apache Software Foundation. ZooKeeper 3.5 Documentation [EB/OL]. [2020-05-15]. <https://zookeeper.apache.org>.
- [11] 刘南, 刘仁义. Web GIS 原理及其应用: 主要 Web GIS 平台开发实例 [M]. 北京: 科学出版社, 2007.
- [12] 张绍平, 史杏荣, 王培康. 三层结构的组件式 WEBGIS 的设计与实现 [J]. 计算机工程与应用, 2002, 38(6): 113-115.
- [13] Apache Software Foundation. The Apache velocity project [EB/OL]. [2020-06-14]. <https://zookeeper.apache.org>.
- [14] Vladimir Agafonkin. Leaflet: an open-source JavaScript library for mobile-friendly interactive maps [EB/OL]. [2020-06-14]. <https://leafletjs.com>.
- [15] Eclipse Foundation. A gentle guide to asynchronous programming with Eclipse Vert.x for Java developers [EB/OL]. [2020-06-13]. <https://vertx.io/docs/guide-for-java-devs>.
- [16] Open Source Geospatial Foundation. GeoTools 24-SNAPSHOT user guide: GeoTools [EB/OL]. [2020-06-18]. <http://docs.geotools.org/latest/userguide/geotools.html>.

Development and Design of a High Resolution China Offshore Wind Energy Resource Database

LAN Zhi-gang¹, YUE Lei², SUN Yang-zhou¹, WU Yong-hu¹, GUO Xue-fei¹, YU Ting¹

1. *CNOOC Research Institute, Beijing 100028, China;*

2. *Beijing Tianji digital measurement data Technology Co., Ltd., Beijing 100020, China*

Abstract: The China offshore wind energy resource database can provide strong support for China offshore wind power plant planning, site selection and early-stage wind energy resource evaluation. In this paper, the data of offshore wind energy resources with high spatial resolution obtained by the National Climate Center are analyzed and evaluated based on the mesoscale meteorological model. The validity of using the data as the source of the database is demonstrated, and the design principle of the database is provided. The detailed technological architecture of the database is described based on the current mainstream Spark technology combined with geographic information system and network technology. China's offshore wind energy resource database with high resolution is finally constructed based on GIS. The technical characteristics and main functions of the database are introduced in this paper. Results show that the data source of the database is reliable and the technological architecture is clear and reasonable, which can facilitate the search, query, statistical calculation, data display and effective storage management of wind energy resources. The database can serve as an effective technical platform for the management and evaluation of offshore wind energy resources in China.

Key words: offshore wind resources; database; technological architecture