

数据差值分析及应用

张耀南

中科院 寒区旱区环境与工程研究所

摘 要

数据是科学，模型是对数据科学的补充。驱动模型需要连续的时间序列数据和空间覆盖数据。在有限的地面观测点的少量数据的支持下，开展模型研究需要利用多种数学方法将有限的点上数据扩展的面上，将有限的时间序列数据拓展成连续的时间序列数据，形成驱动模型的数据集，进而寻找数据中隐含的科学意义。插值分析在空间数据、时间序列数据的制备方面具有重要的作用。目前常用的插值分析方法有常用于气象要素空间插值的方法有多变量自回归模型，广义线性模型，反距离权重法（Inverse Distance Weighing）、多项式插值法（Interpolating Polynomials）、克里格法（Kriging）、样条插值法（Spline）等。这些方法在气象研究、水文研究中发挥着重要的作用。本文重点基于小波分析和神经网络两种数据分析方法，分别对黄河径流的变化趋势，以及黑河流域平均气温进行了差值分析。依据插值数据的推演分析，黄河源区年径流量具有 8a、15a、22a 和 36a 左右的变化周期，其中 8a、36a 左右的周期变化最为显著。2007 年后流量将呈增加的趋势；1928 年、1982 年和 1985 年是径流变化趋势重要的转变点。在小波分解的基础上，基于 BP 神经网络模型构建了黄河源区年径流量的长期动态预报模型，利用该模型对未来 10 年的流量变化进行了预测，并对其预报结果进行了分析。同时利用西部气象台站的观测数据，对神经网络模型进行了训练、测试。资料以来的近 40 年余年间，河源区气温与全球变暖有着较好的对应关系，该区不同区域的平均气温均呈不同程度的振荡上升态势，河源区径流的变化与该区域降水量、气温的变化之间存在较显著的非线性关系。由于气温的持续上升与主要产流区降水的持续减少，该区域径流量呈持续减少的趋势。利用优化了的神经网络

络模型对青海省、甘肃省、黑河流域的温度分布进行了网格插值。测试结果表明神经网络模型在空间数据插值应用中可以获得比较理想的数据。

关键词 时空数据、空间插值、小波分析、神经网络

张耀南 男

职 称：中科院 寒区旱区环境与工程研究所 研究员

职 务：中科院 寒区旱区环境与工程研究所「计算机网络室」主任

研究领域：基于数据、模型和计算的 e-science

个人简介：1996年毕业于兰州大学近代物理系原子核物理专业，获学士学位。1993年毕业于中科院研究生院，获野外测试方面的硕士学位。2005年毕业于中科院研究生院，获资源环境数据平台方向的博士学位。自2000年来，先后主持了寒区旱区特色数据库、大型计算与海量存储环境建设、特殊环境与灾害数据平台台建设、黑河流域交叉集成研究空间建模与模拟环境建设、支持黑河流域综合集成研究的 e-science 环境建设等 23 项。建立了青藏铁路、黑河集成的空间建模与模拟环境。建立了数据编码、分类、共享、应用、系统元数据等数据管理技术，全面定义了数据平台的结构和功能，提出利用「数据抽象」概念及 3 层逻辑数据库技术建立技术，利用「理论—数据（实验或观测）—计算」螺旋推进研究模式开展了 SSIB、SME 等应用研究。初步建立了地学研究的 e-science 体系结构和环境。重点致力于基于数据、模型、计算、可视化环境和生态 e-science 的研究。发表了 40 多篇文章，开发了 8 注注册软件。

联络电话：86-931-4967598

联络邮箱：yaonan@lzb.ac.cn