Comparison of spatial interpolation methods using a simulation experiment based on Australian seabed sediment data

Jin Li* and Andrew D. Heap
Marine & Coastal Environment
Petroleum and Marine Division
Geoscience Australia
GPO Box 378
Canberra ACT 2601

* Corresponding author:
Jin Li
GPO Box 378
Canberra ACT 2601
Australia
Email: jin.li@ga.gov.au

Spatial distribution data of environmental variables are increasingly required as geographic information systems (GIS) and modelling techniques become powerful tools in natural resource management and biological conservation. However, the spatial distribution data are usually not available and the data available are often collected from point sources. This is particularly true of seabed data for the world's oceans, especially the deep ocean. A typical example is Australia's marine region. Here, Geoscience Australia has to derive spatial distribution data of seabed sediment texture and composition for 8.9 million km^2 of Australia's marine region from about 14,000 sparsely and unevenly distributed samples. The need for these data comes from seabed habitat classifications and predictions of marine biodiversity as key information sources supporting ecosystem-based management. Spatial interpolation techniques provide essential tools to generate such spatial distribution data by estimating the values for the unknown locations using the point samples, but they are often data- or even variable- specific. The estimation of a spatial interpolator is usually affected by many factors including the nature of data and sample density. There are no consistent findings about how these factors affect the performance of spatial interpolators. Therefore, it is difficult to select an appropriate interpolator for a given input dataset. In this study, we aim to select appropriate spatial interpolation methods by comparing their respective performance using a simulation experiment based on Australian seabed sediment data in R. Three factors affecting the accuracy and precision of the interpolations are considered: the spatial interpolation method, spatial variation in data, and sample density. Stratification based on geomorphic features is also used to improve estimation. Bathymetry data are considered as secondary information in the experiment. Cross-validation is used to assess the performance of spatial interpolation methods. Results of this experiment provide suggestions and guidelines for improving the spatial interpolations of marine environmental data, which have application for using seabed mapping and habitat characterisations in achieving management and conservation goals.