



Systematically biological prioritizing remediation sites based on datasets of biological investigations and heavy metals in soil

Wei-Chih Lin, Yu-Pin Lin, and Johnathen Anthony

Department of Bioenvironmental Systems Engineering, National Taiwan University, Taipei, Taiwan (b97602046@ntu.edu.tw)

Heavy metal pollution has adverse effects on not only the focal invertebrate species of this study, such as reduction in pupa weight and increased larval mortality, but also on the higher trophic level organisms which feed on them, either directly or indirectly, through the process of biomagnification. Despite this, few studies regarding remediation prioritization take species distribution or biological conservation priorities into consideration. This study develops a novel approach for delineating sites which are both contaminated by any of 5 readily bioaccumulated heavy metal soil contaminants and are of high ecological importance for the highly mobile, low trophic level focal species. The conservation priority of each site was based on the projected distributions of 6 moth species simulated via the presence-only maximum entropy species distribution model followed by the subsequent application of a systematic conservation tool. In order to increase the number of available samples, we also integrated crowd-sourced data with professionally-collected data via a novel optimization procedure based on a simulated annealing algorithm. This integration procedure was important since while crowd-sourced data can drastically increase the number of data samples available to ecologists, still the quality or reliability of crowd-sourced data can be called into question, adding yet another source of uncertainty in projecting species distributions. The optimization method screens crowd-sourced data in terms of the environmental variables which correspond to professionally-collected data. The sample distribution data was derived from two different sources, including the EnjoyMoths project in Taiwan (crowd-sourced data) and the Global Biodiversity Information Facility (GBIF) field data (professional data). The distributions of heavy metal concentrations were generated via 1000 iterations of a geostatistical co-simulation approach. The uncertainties in distributions of the heavy metals were then quantified based on the overall consistency between realizations. Finally, Information-Gap Decision Theory (IGDT) was applied to rank the remediation priorities of contaminated sites in terms of both spatial consensus of multiple heavy metal realizations and the priority of specific conservation areas. Our results show that the crowd-sourced optimization algorithm developed in this study is effective at selecting suitable data from crowd-sourced data. By using this technique the available sample data increased to a total number of 96, 162, 72, 62, 69 and 62 or, that is, 2.6, 1.6, 2.5, 1.6, 1.2 and 1.8 times that originally available through the GBIF professionally-assembled database. Additionally, for all species considered the performance of models, in terms of test-AUC values, based on the combination of both data sources exceeded those models which were based on a single data source. Furthermore, the additional optimization-selected data lowered the overall variability, and therefore uncertainty, of model outputs. Based on the projected species distributions, our results revealed that around 30% of high species hotspot areas were also identified as contaminated. The decision-making tool, IGDT, successfully yielded remediation plans in terms of specific ecological value requirements, false positive tolerance rates of contaminated areas, and expected decision robustness. The proposed approach can be applied both to identify high conservation priority sites contaminated by heavy metals, based on the combination of screened crowd-sourced and professionally-collected data, and in making robust remediation decisions.