Explaining and predicting US/China invasive species distributions using statistical and machine learning tools

Albright, Thomas P[1], Chen, Lijun[2]
[1]SAIC, USGS/EROS Data Center, Sioux Falls, SD 57198, USA; and [n1] University of Wisconsin, Madison, WI 53706, USA (e-mail: talbrig1@wisc.edu; tel: 608-265-6427)
[2]National Geomatics Center of China, Beijing, P.R. China

The U.S. and China are major sources of invasive organisms for each other due to floristic, climatic, and other geographic similarities coupled with their rapid transformation from historic isolation to present-day high levels of trade and travel. Because of this potential for biological invasion, we are investigating the distribution of Sino-American invasive plants, the environmental factors that influence these distributions, and the ability to predict them using statistical and machine-learning tools, including generalized linear models and the Genetic Algorithm for Rule-set Prediction. Early work has focused on three species native to China that are currently considered invasive in the US: Chinese tallow tree (*Triadica sebifera*), kudzu (*Pueraria montana* var. *lobata*), and tree-of-heaven (*Ailanthus altissima*). Models were generated from data derived by overlaying a sample of species locations on a suite of climate, topographic, land cover, and landscape metric raster data sets with cell sizes ranging from 1 km to 10 decimal minutes. While mean annual temperature and precipitation were commonly the most significant variables, we found that the relationships and factors varied according to species and scale of analysis. By developing separate models based on both native and invaded ranges, we could compare the significant factors and project distributions from one range to the other. Such projections yielded generally satisfactory results, suggesting an ability to predict potential invaded distributions based on native range distributions.

keywords: *Ailanthus altissima*, biogeography, China, distribution, GIS, invasive species, modeling, *Pueraria montana* var. *lobata*, *Triadica sebifera*, United States