

Abstract Submitted
for the 4CS20 Meeting of
The American Physical Society

Clustering geospatial data for machine learning modeling of ambient soundscapes¹ MITCHELL CUTLER, KATRINA PEDERSEN, MARK TRANSTRUM, KENT GEE, Brigham Young University, SHANE LYMPANY, MICHAEL JAMES, Blue Ridge Research and Consulting, LLC — Outdoor ambient acoustical environments may be predicted through supervised machine learning using geospatial features as inputs. Previous work used K-Means clustering applied to the geospatial features to identify distinct geographic regions. The clustering results help provide physical insights regarding which features are likely to play the largest roles in supervised learning models and which locations are impacted by different acoustic training data. However, these results may be sensitive to details of the geospatial data, such as how the data are scaled or the presence of similar redundant features. This work builds on previous results by constructing a reduced feature set by removing redundant geospatial features and by using a physically motivated scaling scheme. Clustering analysis applied to this dataset indicates that the contiguous United States can be naturally clustered into eight human-interpretable geographic regions. Hierarchical clustering is used to further subdivide these eight clusters into more fine-grained regions. One finding of interest is that no geospatial layer in the present soundscape model uniquely identifies rivers. These results will guide further geospatial layer development and acoustical data collection for more accurate soundscape models.

¹Work supported by a U.S. Army SBIR.

Mitchell Cutler
Brigham Young University

Date submitted: 25 Sep 2020

Electronic form version 1.4