Machine-learning based spatiotemporal heterogeneity analysis for population behavior mining of incentive-based waste segregation

Peng Jiang[1,*], Shanni Liu[1], Jieyu Zhou[2], Yee Van Fan[3]

[1]Department of Industrial Engineering and Management, Sichuan University, Chengdu, 610064, China.
[2]NUS Environmental Research Institute, National University of Singapore, 117411, Singapore.
[3]Environmental Change Institute (ECI), University of Oxford, Oxford, OX1 3QY, United Kingdom
* Corresponding authors: pengjiang@scu.edu.cn

Abstract: Effective household waste segregation and recycling are crucial to achieving urban sustainable development. The incentive-based segregation guaranteed by digital technologies has been demonstrated as a practical pathway to promote waste segregation behavior in emerging megacities. Due to the imbalance of resource inputs and management policies in different regions in megacities, spatial and temporal heterogeneity of population participation behavior exists broadly. However, beyond questionnaire and interview-based studies, there is limited literature quantifying the heterogeneity and its spatiotemporal dynamics using big urban data and Internet of Things (IoT) technology. With the support of spatiotemporal Bayesian-based machine learning modeling, we analyze incentive-based Shanghai Green Account data that covers more than 7.6 M households and 22 M population, to conduct quantification analysis via spatiotemporal data calibration, data mapping, and disparity analysis of public engagement across regions in the megacity Shanghai. Drawing on calibrated data from 16 administrative regions, we systematically evaluate and visualize spatiotemporal patterns of incentive-based waste segregation behavior, identifying significant disparities in population engagement across time and space. These patterns reveal clear hotspots and temporal peaks, which align closely with variations in socioeconomic development, infrastructure provision, policy visibility, and demographic composition.We further construct a structural equation model to elucidate the underlying mechanisms driving this spatiotemporal heterogeneity, quantifying the influence of economic development levels, public health incidents, demographic composition, infrastructure indices, and publicity intensity. This study employs large-scale natural population experiments that feed into data mining modeling, enriching methodological approaches to population participation analysis with multimodal urban big data in megacities, while providing theoretical and empirical insights to optimize incentive-based waste segregation management and advance sustainable smart city strategies.

Spatiotemporal heterogeneity; Emerging megacity; Machine learning; Sustainable urban development
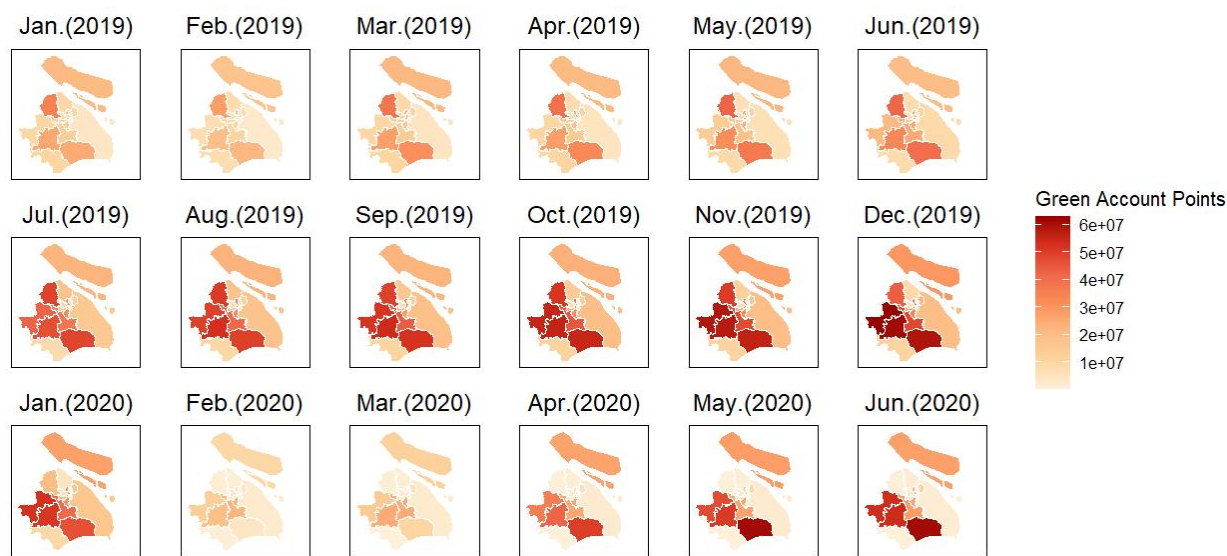

Fig. 1. The spatiotemporal dynamics of newly added accumulating points across 16 regions of Shanghai.