EDITORIAL

Lack of social data hinders integrative urban ecology research: obstacles and opportunities

**Ganlin Huang · Liyuan Chen · Morgan Grove ·
Weiqi Zhou · Steward Pickett**

Landscape sustainability science (LSS) focuses on linkages and interactions among landscape pattern, biodiversity, ecosystem services and human wellbeing (Wu 2013). As an interdisciplinary and transdisciplinary research field, LSS increasingly calls for cross-disciplinary approaches, especially those integrating natural and social sciences (Wu 2021, 2022). Urban landscapes are social-ecological systems, and provide ideal places to develop cross-disciplinary approaches for integrated research across natural and social sciences (Pickett et al. 2016; Zhou et al. 2021). Urban ecology (UE) studies have grown rapidly over the past 25 years to understand the impact of urbanization and promote residents' well-being. A recent review shows that the number of peer-reviewed publications on urban ecology has grown 15% annually in the past two decades (Chen and Huang 2020). Urban ecology pays attention to people in urban change,

along with the expansion of urbanized land cover and change of its spatial form (Zhou et al. 2022). Indeed, the two major characteristics of urban ecology: spatial heterogeneity and the coupled social-ecological nature of the city often come in side by side in UE studies (Pickett et al. 2001). These studies view cities as coupled social-ecological systems and therefore require social data for analysis. Social data are critical for understanding how are people impacted by urbanization processes? Who receives the benefits and who bears the cost during the urbanization process? How do urban dwellers' responses, decisions, or behaviors feed back to the system? Such questions are highly relevant to the core research agenda for LSS.

Answering these questions requires incorporating social dimension into UE studies and views cities as coupled social-ecological systems, with interactions between people and nature at multiple scales (Pickett et al. 2011). Our search for peer-reviewed publications with keywords of "city or urban" and "ecology"

G. Huang (✉) · L. Chen
School of Natural Resources, Faculty of Geographical Science, Beijing Normal University, Beijing 100875, China
e-mail: ghuang@bnu.edu.cn

G. Huang · L. Chen
State Key Laboratory of Earth Surface Processes and Resource Ecology (ESPRE), Beijing Normal University, Beijing 100875, China

M. Grove
Baltimore Field Station, Northern Research Station, USDA Forest Service, Baltimore, MD 21228, USA

W. Zhou
State Key Laboratory of Urban and Regional Ecology, Research Center for Eco-Environmental Science, Chinese Academy of Sciences, Beijing 100085, China

S. Pickett
Cary Institute of Ecosystem Studies, Millbrook, NY 12545, USA

indicated that 39% studies considered social factors (Fig. 1; also see Supplementary Material—Appendix 1 and Table S1 in Supplementary Material—Appendix 2). However, this number varies significantly across countries and regions. For example, 46% of the UE studies conducted in the US included a social aspect. In contrast, less than 30% of China's UE studies considered social factors (Fig. 1). We argue that the (un)availability of social data is a major reason for such an unbalanced situation, which is indispensable for integrative UE research.

## Importance of social data to urban ecology

We define social data as information about people. Social data describe people from a variety of aspects, such as demographic characteristics, preference or perception about certain issues (e.g., trees in parks or global warming), their willingness to take action (e.g., exercise or take public transit), and their behavior (Chen and Huang 2020). Social data enable researchers to capture social differentiation among people when studying coupled, urban socio-ecological systems (Pickett et al. 1997; Grimm et al. 2000;

Redman et al. 2004; Singh et al. 2010; Collins et al. 2011; Miller et al. 2014).

Understanding social differentiation among urban residents is crucial in UE studies for at least two reasons. First, social differentiation provides clues for differences in how people make decisions affecting the environment. Second, social differentiation tells how vulnerable people are in the face of environmental challenges such as heatwaves, water scarcity, or hurricanes. Therefore, it is essential for urban ecologists to be able to describe social differentiation among residents and associate these groups with a variety of research topics such as the provision of urban greenspace (Wolch et al. 2014), heat vulnerability (Huang et al. 2011), and health consequences from air pollution (Huang et al. 2019).

Urban greenspace (UGS) refers to vegetation and open space in cities, which provides a variety of ecosystem services to residents, including places for recreation, exercise, restoration, and heat mitigation. We use UGS as an example to show how incorporating social data can expand the range of research topics, and provide some sample research questions that need social data to address (Fig. 2). Notice that only the first question focusing on features of UGS and its impact on biodiversity can be answered without social data. In addition to its spatial patterns and

Fig. 1 Publications of urban ecological studies with/without a social dimension. Globally 39% published UE studies considered social factors. This proportion is higher in US (46%), and lower in China (29%). Studies of UE in US and China take 14% and 15% of the global total respectively
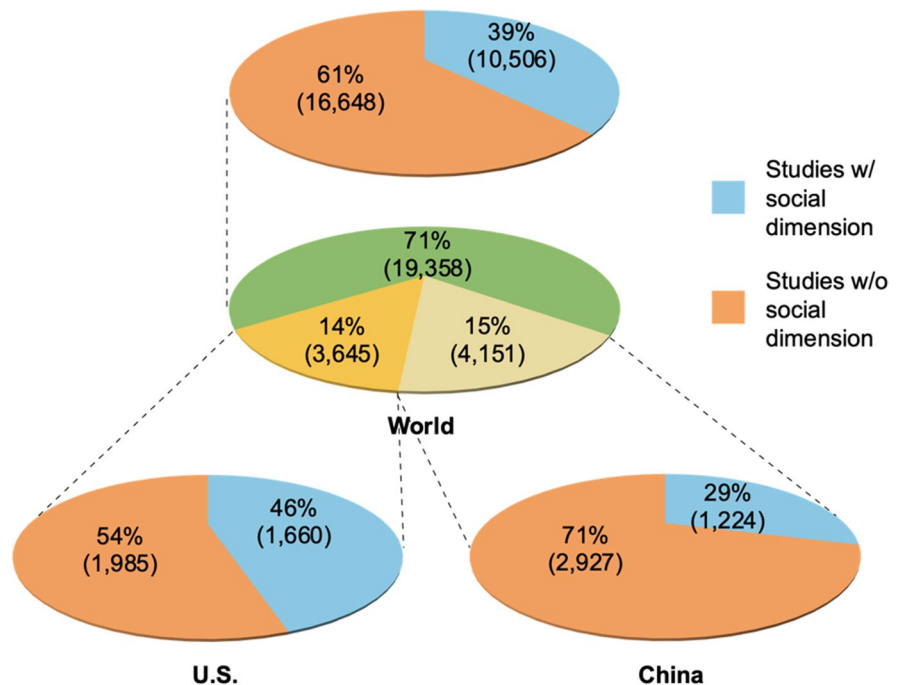
Fig. 2 Examples of research questions addressed by different types of data
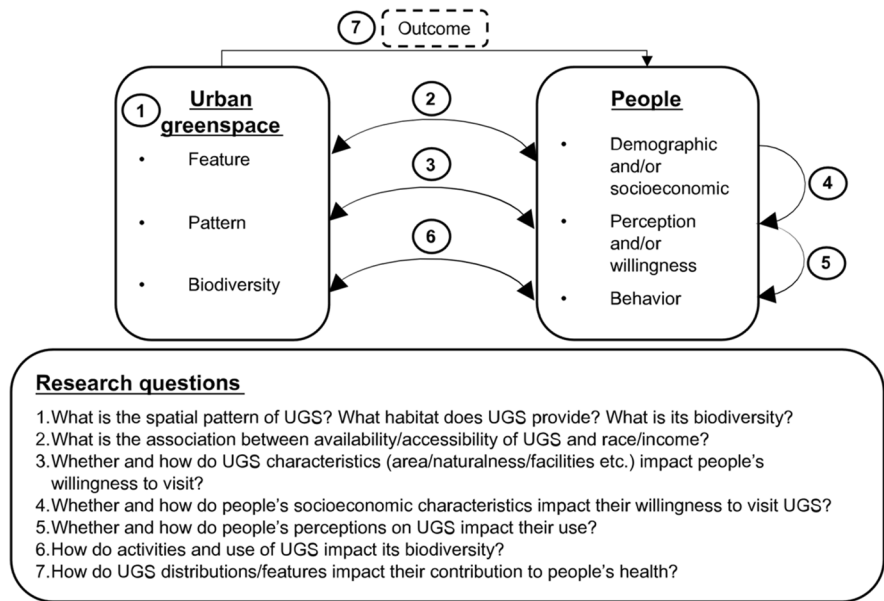
biodiversity, we would also want to know accessibility of UGS to different social groups (Rigolon 2016), factors affect people's perceptions of a greenspace (Pietrzyk-Kaszynska et al. 2017), their willingness to visit (Tu et al. 2020), impacts on biodiversity from both visitors and UGS management (Chang and Lee 2016), and finally the health benefit of using greenspace (Liu et al. 2017). Answering these questions requires social data. With such information, researchers can go beyond the biophysical features or patterns of urban greenspace to a more complete picture of how it functions and its interactions in the coupled social-ecological system.

Different types of social data in urban ecological studies

Social data used in UE studies can vary substantially from case to case. We summarize commonly used social data into three categories. The first type describes people's demographic-socioeconomic characteristics. This type of data includes the number of people age, gender, educational attainment, income, marital status, employment, race, etc. The second type focuses on how people think, often described as their perception about certain issues or their willingness to take certain actions. Examples include visitors' perceptions of biodiversity in parks (Muratet

et al. 2015), and peoples' willingness to pay for new parks (Andrews et al. 2017). The third type is related to behaviors. Usually measured are those behaviors with direct environmental consequence, such as how often residents fertilize their lawns and how much fertilizers they use (Zhou et al. 2008).

These social data mostly come from three sources: statistics compiled and published by government, commercial datasets, and investigations or surveys conducted by researchers (Chen and Huang 2020). Availability varies among different sources as well as in different places. Next, we explain each source in more detail, using examples from US and China for comparation.

Government statistical data

Government statistical data consists mostly of demography and socioeconomics, with some information on people's behaviors. For example, the USA Census asks about how people commute to work. The decennial census conducted by the USA federal government is collected at relatively fine scales organized in block group of roughly 400 households on average. The census enumerates the number of residents, their gender, age, race and ethnicity, ownership versus rental of the dwelling, and relationships of the members of the household. More aggregated, community-based statistics on, for example, income,

poverty, employment, marital or partnership status, housing value, national origin, business statistics, and other detailed information is collected every three to five years through the American Community Survey (data.census.gov),. These data are readily available without charge from the Census Bureau website (http://data.census.gov).

The situation for government statistical data in China is quite different. Coarse scale, limited information, and inconsistent availability all pose obstacles for researchers to use government statistical data in China. First, census data are usually collected by the lowest administrative unit, which can encompass an entire town or be the equivalent of a district in a large city (Zhang et al. 2020; National Bureau of Statistics 2021). Second, census data include such things as age, gender, educational attainment, residential registry situation (known as *hukou* in Chinese), and types of housing. Unfortunately, household income, which is commonly used to describe people's economic situation, is not collected at the finest level of census. This is likely because census data was designed in 1950's when income was not an important indicator to differentiate people's social status in China (Xie and Zhou 2014). Instead, indicators such as *hukou* registry and number of bedrooms captured more differences among people than income at that time. Currently, however, these indicators often fail to capture the relevant differences among people as urban and consumption-based lifestyles become more prevalent (Wu 2019).

Finally, data availability varies across cities and regions. While some cities (such as Beijing) may have census data readily downloadable for free, in many cases, data are only directly accessible from a county or district office (Li and Liu 2016). Therefore, if researchers investigate the relationship between educational attainment and greenspace use in a city, they need to contact (or sometimes visit) the Department of Statistic in each district of that city for education data. The number of districts in a city varies from several to over a dozen.

Commercial data

Private sources of data relevant for UE studies go beyond the basics of demographics, education, and income to include information on consumption and lifestyle (Grove et al. 2014). In the US, information on consumption patterns by census block or postal zones are collected by retailers, and this information is aggregated into lifestyle clusters (e.g., Claritas© or Tapestry©). In addition, private firms organize and curate census data in ways convenient to users for a fee. In contrast, there are few commercially available data sources in China. Researchers may purchase some datasets from enterprises for analysis such as comments or posts from social media, place of interest, and mobile app use (Xiao et al. 2019). However, these datasets are often generated as a "by-product" from their main business, containing limited information and often need extra work to organize before they can be used (Zhang and Zhou 2018). A particularly serious gap in China is the absence of data that indexes consumption and lifestyle characteristics, given its rapid economic growth and urbanization process since the economic reform in 1978 (Chen et al. 2016).

Survey

Survey usually meets the needs of a study better than the statistics and commercial datasets in terms of the selection of indicators, but demands much more time, knowledge, labor, and financial resources. It is often the last choice for obtaining social data when relevant information is not available or accessible from government statistics or commercial datasets. When studying a large population, it can be challenging to have a representative sample size.

Limited social data hamper integrative urban ecological research

It is important to recognize the dependence of urban ecologists on social data to advance social-ecological understanding and applications. Lack of adequate, good quality, or accessible social data in certain countries or regions may further enlarge the existing gap in integrative urban studies and the ability to promote human well-being and urban resilience. For example, UE studies in China account for 15% of the world's output with a growth rate nearly twice the global rate measured by the number of peer-reviewed publications (China 27% vs. World 15%) (Supplementary Material—Appendix 3). Due to a lack of social data, however, the number of publications with a social

dimension in China has been growing at a slower pace (Fig. S1 in Supplementary Material—Appendix 3). If these trends continue, we will see more UE studies from China but less emphasis on the social dimensions and the ability to inform understanding and decision making.

We compared UE studies of Baltimore, US and Beijing, China for inclusion of social data (Table S2 in Supplementary Material—Appendix 2). Both cities have well-established, long term ecological study sites. We examined peer-reviewed publications between the year the research site was established (1997 for Baltimore and 2008 for Beijing) to 2019 and our search identified 18 and 23 publications conducted in Baltimore and Beijing respectively. We found that studies in Beijing relied heavily on individual surveys (e.g., direct observation, questionnaire survey, and interview) to obtain social data. Only one study in Beijing used an education indicator from a government statistical database. In contrast, government statistical data was much more commonly used in Baltimore, where seven studies used indicators from seven different government statistical databases. In addition, one study in Beijing also employed "big data" from social media in an attempt to acquire wide-ranging data at relatively low cost. Disadvantage in accessibility and quality of social data in other places (often developing countries) may also discourage researchers to conduct integrative urban studies, which is badly needed to understand their unique urbanization trajectories (McHale et al. 2015).

Call for actions

While data availability poses an obvious obstacle to integrate social data into UE studies, we urge researchers to incorporate social data in UE studies, especially in places that have been less studied. There are always opportunities to overcome the present challenges. First, we recommend that governments make existing statistical data more accessible to academic communities. If resource allows, we encourage governments to review and update census data to capture social differentiation. For example, considering the rapid economic growth in China, household income is an important indicator that shall be included in census information. Second, we suggest that researchers share the data from their own work,

regardless of its source: individual surveys or a collection of existing government datasets. Without data documentation and data-sharing, much of the precious resource (i.e., time, labor) is wasted in repetitive work of data collecting, obtaining, and cleaning. Creative ways are needed to give appropriate credits to encourage researchers to share datasets. Last, big data from social media and other platforms is becoming a new and promising source for social data. Big data may reveal information that has seldom been captured before, expose much larger scales, elicit much lower costs, and especially, may address places where traditional datasets from government and commercial sources are limited. The scale, coverage, access, content, and cost of big data vary substantially in different countries and regions, which is beyond the scope of this paper. It also has some limitations that have been recognized (Ilieva and McPhearson 2018). Nevertheless, it provides an important source to collect information about people. In addition, the rise of technology and social media also makes it easier for some traditional ways to generate social data. For example, apps in smartphones make it easier to automatically link citizen science observations to locations in contrast to uploading data to a designated webpage.

In conclusion, it is important to recognize the dependence of urban ecologists on social data to advance social-ecological understanding and applications. Future research could comparatively evaluate the necessary social-ecological data informatics systems to advance actionable science, knowledge, and applications for this 21st urban century. We urge practitioners and researchers to work together to improve data integration for UE research through interdisciplinary collaborations, data sharing, and the utilization of big data, especially in places that have been less studied. Such actions advance the social-ecological understanding of urban landscapes andfacilitate the development of cross-disciplinary approaches for integrated research across natural and social sciences to advance LSS.

References

Andrews B, Ferrini S, Bateman I (2017) Good parks - bad parks: the influence of perceptions of location on WTP and preference motives for urban parks. J Environ Econ Policy 6(2):204–224. https://doi.org/10.1080/21606544.2016.1268543

Chang HY, Lee YF (2016) Effects of area size, heterogeneity, isolation, and disturbances on urban park avifauna in a highly populated tropical city. Urban Ecosyst 19(1):257–274. https://doi.org/10.1007/s11252-015-0481-5

Chen L, Huang G (2020) Socioeconomic factors in global long-term urban ecological studies: indicators, sources and applications (in Chinese). Acta Ecol Sin 40(18):6678–6686

Chen M, Liu W, Lu D (2016) Challenges and the way forward in China's new-type urbanization. Land Use Pol 55:334–339. https://doi.org/10.1016/j.landusepol.2015.07.025

Collins SL, Carpenter SR, Swinton SM, Orenstein DE, Childers DL, Gragson TL, Grimm NB, Grove JM, Harlan SL, Kaye JP, Knapp AK, Kofinas GP, Magnuson JJ, McDowell WH, Melack JM, Ogden LA, Robertson GP, Smith MD, Whitmer AC (2011) An integrated conceptual framework for long-term social-ecological research. Front Ecol Environ 9(6):351–357. https://doi.org/10.1890/100068

Grimm NB, Grove JM, Pickett STA, Redman CL (2000) Integrated Approaches to Long-Term Studies of Urban Ecological Systems: Urban ecological systems present multiple challenges to ecologists—pervasive human impact and extreme heterogeneity of cities, and the need to integrate social and ecological approaches, concepts, and theory. Bioscience 50(7):571–584. https://doi.org/10.1641/0006-3568(2000)050[0571:IATLTO]2.0.CO;2

Grove JM, Locke DH, O'Neil-Dunne JPM (2014) An Ecology of prestige in New York City: examining the relationships among population density, socio-economic status, group identity, and residential canopy cover. Environ Manage 54(3):402–419. https://doi.org/10.1007/s00267-014-0310-2

Huang G, Zhou W, Cadenasso ML (2011) Is everyone hot in the city? Spatial pattern of land surface temperatures, land cover and neighborhood socioeconomic characteristics in Baltimore, MD. J Environ Manage 92(7):1753–1759. https://doi.org/10.1016/j.jenvman.2011.02.006

Huang G, Zhou W, Qian Y, Fisher B (2019) Breathing the same air? Socioeconomic disparities in PM2.5 exposure and the potential benefits from air filtration. Sci Total Environ 657:619–626. https://doi.org/10.1016/j.scitotenv.2018.11.428

Ilieva RT, McPhearson T (2018) Social-media data for urban sustainability. Nat Sustain 1(10):553–565. https://doi.org/10.1038/s41893-018-0153-6

Li H, Liu Y (2016) Neighborhood socioeconomic disadvantage and urban public green spaces availability: a localized modeling approach to inform land use policy. Land Use Pol 57:470–478. https://doi.org/10.1016/j.landusepol.2016.06.015

Liu H, Li F, Li J, Zhang Y (2017) The relationships between urban parks, residents' physical activity, and mental health benefits: A case study from Beijing, China. J Environ Manage 190:223–230. https://doi.org/10.1016/j.jenvman.2016.12.058

McHale MR, Pickett STA, Barbosa O, Bunn DN, Cadenasso ML, Childers DL, Gartin M, Hess GR, Iwaniec DM, McPhearson T, Peterson MN, Poole AK, Rivers L, Shutters ST, Zhou W (2015) The new global urban realm: complex, connected, diffuse, and diverse social-ecological systems. Sustainability 7(5):5211–5240. https://doi.org/10.3390/su7055211

Miller TR, Wiek A, Sarewitz D, Robinson J, Olsson L, Kriebel D, Loorbach D (2014) The future of sustainability science: a solutions-oriented research agenda. Sustain Sci 9(2):239–246. https://doi.org/10.1007/s11625-013-0224-6

Muratet A, Pellegrini P, Dufour AB, Arrif T, Chiron F (2015) Perception and knowledge of plant diversity among urban park users. Landsc Urban Plan 137:95–106. https://doi.org/10.1016/j.landurbplan.2015.01.003

National Bureau of Statistics (2021) Statistical report system for basic urban conditions. https://www.stats.gov.cn/tjsj/tjzd/gjtjzd/202103/t20210316_1814949.html. Accessed 25 May 2022

Pickett STA, Burch WR, Dalton SE, Foresman TW, Grove JM, Rowntree R (1997) A conceptual framework for the study of human ecosystems in urban areas. Urban Ecosyst 1(4):185–199. https://doi.org/10.1023/a:1018531712889

Pickett STA, Cadenasso ML, Grove JM, Nilon CH, Pouyat RV, Zipperer WC, Costanza R (2001) Urban ecological systems: linking terrestrial ecological, physical, and socioeconomic components of metropolitan areas. Annu Rev Ecol Syst 32(1):127–157. https://doi.org/10.1146/annurev.ecolsys.32.081501.114012

Pickett STA, Buckley GL, Kaushal SS, Williams Y (2011) Social-ecological science in the humane metropolis. Urban Ecosyst 14(3):319–339. https://doi.org/10.1007/s11252-011-0166-7

Pickett STA, Cadenasso ML, Childers DL, McDonnell MJ, Zhou W (2016) Evolution and future of urban ecological science: ecology in, of, and for the city. Ecosyst Health Sustain 2(7):e01229. https://doi.org/10.1002/ehs2.1229

Pietrzyk-Kaszynska A, Czepkiewicz M, Kronenberg J (2017) Eliciting non-monetary values of formal and informal urban green spaces using public participation GIS. Landsc

Urban Plan 160:85–95. https://doi.org/10.1016/j.landurbplan.2016.12.012

Redman CL, Grove JM, Kuby LH (2004) Integrating social science into the long-term ecological research (LTER) network: Social dimensions of ecological change and ecological dimensions of social change. Ecosystems 7(2):161–171. https://doi.org/10.1007/s10021-003-0215-z

Rigolon A (2016) A complex landscape of inequity in access to urban parks: a literature review. Landsc Urban Plan 153:160–169. https://doi.org/10.1016/j.landurbplan.2016.05.017

Singh SJ, Haberl H, Gaube V, Grünbühel CM, Lisivieveci P, Lutz J, Matthews R, Mirtl M, Vadineanu A, Wildenberg M (2010) Conceptualising Long-Term Socio-ecological Research (LTSER): integrating the social dimension. In: Müller F, Baessler C, Schubert H, Klotz S (eds) Long-term ecological research: between theory and application. Springer, Dordrecht, pp 377–398. https://doi.org/10.1007/978-90-481-8782-9_26

Tu X, Huang G, Wu J, Guo X (2020) How do travel distance and park size influence urban park visits? Urban For Urban Green 52:126689. https://doi.org/10.1016/j.ufug.2020.126689

Wolch JR, Byrne J, Newell JP (2014) Urban green space, public health, and environmental justice: The challenge of making cities "just green enough." Landsc Urban Plan 125:234–244. https://doi.org/10.1016/j.landurbplan.2014.01.017

Wu J (2013) Landscape sustainability science: ecosystem services and human well-being in changing landscapes. Landsc Ecol 28(6):999–1023. https://doi.org/10.1007/s10980-013-9894-9

Wu X (2019) Inequality and social stratification in postsocialist China. Annu Rev Sociol 45:363–382. https://doi.org/10.1146/annurev-soc-073018-022516

Wu J (2021) Landscape sustainability science (II): core questions and key approaches. Landsc Ecol 36(8):2453–2485. https://doi.org/10.1007/s10980-021-01245-3

Wu J (2022) A new frontier for landscape ecology and sustainabbility: introducing the world's first atlas of urban agglomerations. Landsc Ecol 37(7):1721–1728. https://doi.org/10.1007/s10980-022-01475-z

Xiao Y, Wang D, Fang J (2019) Exploring the disparities in park access through mobile phone data: evidence from Shanghai, China. Landsc Urban Plan 181:80–91. https://doi.org/10.1016/j.landurbplan.2018.09.013

Xie Y, Zhou X (2014) Income inequality in today's China. Proc Natl Acad Sci USA 111(19):6928–6933. https://doi.org/10.1073/pnas.1403158111

Zhang S, Zhou W (2018) Recreational visits to urban parks and factors affecting park visits: evidence from geotagged social media data. Landsc Urban Plan 180:27–35. https://doi.org/10.1016/j.landurbplan.2018.08.004

Zhang J, Yu Z, Cheng Y, Chen C, Wan Y, Zhao B, Vejre H (2020) Evaluating the disparities in urban green space provision in communities with diverse built environments: the case of a rapidly urbanizing Chinese city. Build Environ 183:107170. https://doi.org/10.1016/j.buildenv.2020.107170

Zhou W, Troy A, Grove JM (2008) Modeling residential lawn fertilization practices: integrating high resolution remote sensing with socioeconomic data. Environ Manage 41(5):742–752. https://doi.org/10.1007/s00267-007-9032-z

Zhou W, Pickett STA, McPhearson T (2021) Conceptual frameworks facilitate integration for transdisciplinary urban science. npj Urban Sustain 1(1):1–11

Zhou W, Yu W, Qian Y, Han L, Pickett STA, Wang J, Li WF, Ouyang Z (2022) Beyond city expansion: multi-scale environmental impacts of urban megaregion formation in China. Natl Sci Rev 9(1):nwab107. https://doi.org/10.1093/nsr/nwab107