



Pitfall of big databases

Zhangqiang You^a, Junhua Hu^{b,1}, Qing Wei^a, Chunwang Li^c, Xiaofei Deng^a, and Zhigang Jiang^{c,d,1}

Now, more and more studies use big databases, such as the World Database on Protected Areas (WDPA) (<https://www.protectedplanet.net/>), which is now a predominant source of information on world protected areas (PAs) (1–4). In PNAS, Elsen et al. (5) report that mountain ranges in Africa and Asia had the lowest elevational protection, based on analyses of the WDPA database. We suspect that this may not be true. China is a mountainous country in Asia (Fig. 1A) (5); we analyzed the elevational protection by China's National Nature Reserves (CNNR) and found that higher elevation regions are under a higher proportion of coverage by CNNR (Fig. 1B). The area of CNNR differed obviously among elevational gradients (Kruskal–Wallis test: $\chi^2 = 143.29$, $P < 0.01$), with the proportion of coverage becoming higher along the elevation (Fig. 1C). The proportion of CNNR coverage in regions with elevation of 4,000 m neared the Aichi Conservation Target 11, and this proportion reached ~30% when elevation was >4,000 m (Fig. 1C). However, figure 1B of Elsen et al. (5) displays that the heartland of Asia, mainly the Qinghai–Tibetan Plateau, has low coverage of I–IV PAs and does not meet the 17% target. This picture contradicts the fact that this region has the highest PA coverage owing to larger CNNRs established across the vast plateau with few or no human settlements (Fig. 1B).

Why is there such a discrepancy? The pitfall lies in the WDPA database itself. First, compared with current >400 CNNRs, only 112 PAs of China are listed in WDPA (Table 1). Second, in WDPA, only 1 (ID95418) and 18 CNNRs are listed as category IV and VI, respectively; most PAs in China are categorized as “not reported,” “not applicable,” or “underdesignated.” This consequence is more profound in figure 2 B–D of

Elsen et al. (5), with clear divergence in the heartland of Asia, considering different categories of PAs. Moreover, besides national nature reserves, many more provincial and county reserves were established in China (www.mep.gov.cn/gkml/hbb/bwj/201611/t20161102_366739.htm). This obviously differs from the data in WDPA (Table 1). Other Asian countries most likely encounter a similar problem. Such a pitfall must be noticed by both conservationists and managers, as the trend of using the WDPA database for authoritative information continues.

We thus strongly urge local authorities to review the data variability of these kinds of big databases, whose importance is beyond description, being used for more and more metaanalyses. The WDPA's manager should contact each country for reporting and categorizing their PAs and check the data to eliminate confusion in areas, because many PAs may have more than one title; for instance, a nature reserve also bears other type names, such as Geopark, forest park, scenic site, or Natural Heritage Site, and so forth. Only when the data are validated and reviewed periodically can we have sound data for further analyses and avoid the pitfall of big databases. We cannot rely on the users to validate the data, nor can we rely on the reviewers and editors of scientific journals to validate these data.

Acknowledgments

This study was supported by the National Natural Science Foundation of China (Grants 31572290 and 31770568), State Key Research and Development Project (Grant 2016YFC0503303), Science and Technology Service Network Initiative of the Chinese Academy of Sciences (Grant KFJ-STZ-ZDTP-013), and the Youth Innovation Promotion Association of the Chinese Academy of Sciences (Award 2015304).

^aEcological Security and Protection Key Laboratory of Sichuan Province, Mianyang Normal University, Mianyang 621000, China; ^bCAS Key Laboratory of Mountain Ecological Restoration and Bioresource Utilization and Ecological Restoration and Biodiversity Conservation Key Laboratory of Sichuan Province, Chengdu Institute of Biology, Chinese Academy of Sciences, Chengdu 610041, China; ^cInstitute of Zoology, Chinese Academy of Sciences, Beijing 100101, China; and ^dCollege of Life Sciences, University of Chinese Academy of Sciences, Beijing 100049, China

Author contributions: Z.J. designed research; Z.Y. and J.H. performed research; Z.Y. and J.H. analyzed data; and Z.Y., J.H., Q.W., C.L., X.D., and Z.J. wrote the paper.

The authors declare no conflict of interest.

Published under the [PNAS license](#).

¹To whom correspondence may be addressed. Email: hujh@cib.ac.cn or jiangzg@ioz.ac.cn.

Published online September 12, 2018.

