

# 4 自然 NATURE

# 中国科学报

由于研究人员使用了不完整的数据进行统计分析,那么结论必然会出现错误。在蒋志刚看来,这是值得所有自然资源保护主义者和管理者正视的问题。

## 研究失实,祸起数据

■本报记者 胡珉琦

随着大数据时代的来临,数据密集型的科研方式已经越来越成为保护生物学研究的一个重要方向,这也意味着科研过程的成败与大数据的质量和完整性息息相关。而对研究者来说,一旦忽视对数据缺陷的辨别,就可能落入大数据的“陷阱”。最近,中国的科研人员在《美国科学院院刊》上发表文章,提出了这一值得被正视的问题。

### 一项“失实”的研究

今年6月,美国加州大学伯克利分校环境科学、政策和管理系 Elsen 等研究人员在《美国科学院院刊》上发表了一项研究成果,他们根据对世界保护地数据库(WDPA)的数据,分析了全球六大洲 1010 条山脉中 44155 个保护区的高程分布,从而量化海拔梯度的全球保护。由于跨海拔梯度保护区强化了对生物多样性和生态系统发育多样性的保护,并且促进了气候变化条件下的物种范围迁移,所以这是研究者关注的问题。

早在 2010 年 10 月,《生物多样性公约》各缔约方通过了《生物多样性战略计划》(2011—2020 年),该战略计划确定了 2020 年全球生物多样性目标(也称爱知目标)。其中一条目标要求,到 2020 年,缔约国至少要保护 17% 的陆地面积和 10% 的海域面积。

这项研究认为,他们所观察到的海拔保护远低于理想状态。就平均海拔保护力度而言,非洲和亚洲的山脉最低,欧洲和南美洲的山脉中等,北美洲和大洋洲的山脉最高。

尤其是一项数据统计结果显示,亚洲的高海拔地带,主要是青藏高原, I-IV 类自然保护地(IUCN 提出的保护地分类标准指南将全球保护地分为六类,其中 I-IV 类可以说属于较严格的保护地类型)的覆盖率很低,远

达不到 17% 的目标。

但在 2016 年,中科院动物所研究员蒋志刚团队在《Biological Conservation》在线发表的有关中国生物多样性保护的文章曾经提到,过去 40 年里,中国自然保护区的数量有显著增加,截至目前,共有 11 个类型,保护区占地面积覆盖了全国国土面积的 17% 以上和海洋面积的 3.5%。

中国是亚洲典型的多山国家,于是,针对国际同行的这项研究,蒋志刚与几位国内同行再次就中国国家自然保护区(CNNR)对不同海拔地区的覆盖率进行了统计分析,得到的结果也的确与他们产生了很大的差异。

中国不仅在青藏高原上已经建立了许多大面积的保护区,在海拔超过 4000 米的地区达到甚至超过了 17% 的“爱知目标”,且海拔越高, CNNR 覆盖的面积比例越大。

蒋志刚等撰文指出, Elsen 等人的研究结论明显与中国的实际情况不符。而造成这一分析结果偏差的原因就在“数据”。

### 成也数据,败也数据

中国的研究人员在世界保护地数据库中发现,列入该系统的中国自然保护区总共 112 个,其中只有一个保护区被归为 IUCN 保护地分类中的 IV 类,12 个被归为不那么严格的保护类别 VI 类,剩余的则被视为未报告或不可分类。

也就是说, Elsen 等人的这项研究所统计的,与中国相关的较为严格的自然保护区数据只有唯一的一个。

而事实上,截至目前,中国国家级自然保护区的数量就达到了 428 个,且属于 IUCN 保护地分类中最严格的类型。如加上国内各个地方级别的自然保护区,总数达到了 2700 多个。

由于研究人员使用了不完整的数据进行统计分析,那么结论必然会出现错误,这就是蒋

志刚所说的“数据陷阱”。但在在他看来,这是值得所有自然保护者和管理者正视的问题。

世界自然保护地数据库(WDPA)是目前全球海洋和陆地自然保护地最大的数据库,收录了全球超过 20 万个自然保护地的信息,它的重要性不言而喻。它是世界保护地的主要数据来源,并不意味着这些数据就是完全可靠的。

尽管, WDPA 对数据提供者制定了严格的门槛,且根据其使用手册的说明,制定了定期更新、数据验证、质量检查的原则和机制,但作为一个覆盖全球免费开放的数据库,数据提供者并没有强制性义务,而是自愿提交的,那么至少在数据完整性上,必然会受到客观条件的影响。

除了数据不全,蒋志刚还提到,一个大数据库往往会在分类问题上, WDPA 使用的是 IUCN 保护地分类标准,但中国自然保护区分类与其并不一致。

他在不久前发表的《论保护地分类与以国家公园为主体的中国保护地建设》一文中指出, IUCN 保护地分类系统定义模糊,缺乏量化标准,保护地的属性也常常存在非唯一属性的问题,比如可可西里自然遗产地,既是自然遗产地,也是严格保护区和荒野地,物种与生境保护区,还是地质景观、国家公园。因此,保护地面积可能重复统计,数据整合可能出现问题。

他认为,世界保护地数据库涉及其他亚洲国家的数据库时很可能也存在这些问题。而且,这也是相同类型数据库普遍存在的问题。

不过,对研究人员而言,数据库并不是唯一可利用的数据资源。蒋志刚想要强调的是,尤其在宏观问题研究中,研究者在使用单一数据库资源时必须足够审慎,如果无法辨识数据完整性,宁愿不

发表论文。

“这关系到科学家的职业操守。”蒋志刚说。

### 科学的生命力在质疑与证伪

“我们已经进入了一个利用大数据识别大格局、发现大规律的年代。尤其在保护生物学中,应用大数据是时代潮流。”

蒋志刚解释,大数据可以帮助科学家节省大量野外收集的时间和资源,将研究水平提高到以往研究者所不能开展的空间尺度与时间跨度。同时,大数据对微观的保护生物学的研究也同样重要,比如来自 Genbank 的数据,对于了解、比较濒危物种的种群历史和演化有重要意义。

也正是由于大数据的重要性,数据库自身在提升数据质量和完整性方面需要作出更多的努力。“比如,数据库管理者应该更主动地与每个提供数据报告的国家建立联系,针对大数据数据库数据来源复杂、数据质量良莠不齐等问题,管理者和维护者在执行层面就应该更认真地完成定期审查、数据查错和验证过程,尽可能提升数据的真实性与准确性,消除混乱,弥补缺陷。”他告诉《中国科学报》记者,“科学的生命力正是在于质疑与证伪。”

只有如此,研究者才会有一个良好的数据库开展进一步的分析,发现真实的规律,避免陷入大数据的陷阱。蒋志刚坦言,“这不能依靠大数据的用户来验证数据,也不能依赖于科学期刊的审稿人和编辑来验证”。

相关文章信息:  
<https://doi.org/10.1073/pnas.1813323115>  
<https://doi.org/10.1073/pnas.1720141115>  
<https://doi.org/10.1016/j.biocon.2016.05.005>  
DOI:10.17520/biods.2018168

### 绿色视野

随着技术的不断进步,将青藏高原的水引入新疆的设想已经不再遥不可及,越来越多的人开始关注其可行性。在不久前召开的西部调水工程研讨会上,专家们再次系统介绍相关方案,成为社会关注焦点。

## 藏水何以解疆之渴

■ 苏京来

城市文明多由水兴起,又因水繁荣。历史悠久的城市一般都依偎着一条主要河流,先有河,后有城,城市的历史便沉淀在河道之旁。河流文明孕育了人类早期文明,并不断发展成熟,最终形成了地域文化,比如源自黄河与长江的中华文明,在发展中不断融合其他文明,从而形成了现在多元、包容的中华文化。

我国河流资源总量比较丰富,有 1500 多条河流的流域面积在 1000 平方公里以上,但大多注入海洋或流出国境,仅有小部分被使用或流入内陆湖泊。长江、黄河、黑龙江、珠江、辽河、海河、淮河等河流向东流入太平洋;西藏的雅鲁藏布江流出国境后注入印度洋;新疆的额尔齐斯河流出国境后注入北冰洋。最近发布的《2016 年中国水资源公报》显示,2016 年全国水资源总量为 32466.4 亿立方米,含地表水资源量 31273.9 亿立方米,其中从国境外流入我国境内的水量为 179.9 亿立方米,从我国流出国境的水量 6083.6 亿立方米,流入界河的水量 1124.6 亿立方米,全国入海水域总量 20825.5 亿立方米,而 2016 年全国用水总量为 6040.2 亿立方米。

虽然我国拥有丰富的水资源,但因为区域分布的不均衡,产生了面积巨大的缺水地区和数量巨大的缺水人口,也直接导致我国区域发展严重不平衡。我国东部地区和中部地区地势平坦,河流众多,虽说人均水资源量远低于世界人均水平,但也可以基本满足区域生态、工农业及生产生活的用水需求。而我国西部地区的水资源分布有着巨大的南北差异,西南地区是许多大江大河的发源地,水量丰沛,水系发达,人均水资源为全国人均水平的 14 倍;而广袤的西北地区地势平坦,光照充足,非常缺水,很多地方常年干旱。其中新疆作为我国最大的省级行政区,有四分之一的土地是沙漠,曾兴盛一时的“城廓之国”楼兰现在也已经销声匿迹,只留下了荒凉的古城废墟。

清朝时期,善于治水的林则徐为解决新疆地区的缺水问题,曾大规模兴建“坎儿井”,取得了很好的效果,后来又提出了将青藏高原的水引入新疆的设想,希望能彻底解决新疆的缺水问题,不过以清朝当时的技术水平,无法将这一设想实现。时过境迁,随着技术的不断进步,人类已经可以“上九天揽月,下五洋捉鳖”,将青藏高原的水引入新疆的设想,已经不是那么遥不可及了,越来越多的人开始关注其可行性。仁者见仁,智者见智,但要想实现这一设想,必须要先解决三个问题。

一是要突破工程建设方面的技术瓶颈。青藏高原地势条件非常复杂,处处高山深谷,千峰万壑,想要跨越一路的大山大河,隧道穿山、架桥穿河是常用办法。在高海拔地区进行超长隧道、超高水坝的建设,在技术上有一定难度,而大量扬水、高比例取水等问题也需要详细论证。此外,青藏高原及其附近有多个地震带,容易发生地质灾害,在保证施工安全的同时,还需要妥善考虑地震等地质灾害对工程产生的影响。

二是要考虑对生态环境的影响。青藏高原是世界海拔最高的高原,具有重要的生态价值,也是生态环境极为脆弱的地区,特别是号称“世界生物基因库”和“天然高山花园”的三江并流区,不仅是世界上生物物种最丰富的地区之一,更是中国境内面积最大的世界自然遗产地。要把青藏高原的水引到新疆,无论是取水还是输水都需要进行大规模施工建设,对动植物生存、地质构造、区域气候等都会产生重大影响,甚至引发难以预测的灾难。

三是要慎重处理国际关系。青藏高原上的雅鲁藏布江、怒江、澜沧江都是著名的国际河流,雅鲁藏布江流出国境后,流经印度后改称布拉马普特拉河,流经孟加拉后又改称杰木纳河,是印度、孟加拉国内仅次于恒河的第二大河;怒江流入缅甸后改称萨尔温江,是缅甸最长的河流,其下游构成了缅甸和泰国之间约 130 公里的国界线;澜沧江是世界第七长河,东南亚第一长河,流出中国国境后改称湄公河,流经老挝、缅甸、泰国、柬埔寨、越南等国家。在这些国际河流进行大规模调水,必须遵守国际河流利用约定的《赫尔辛基规则》,否则容易引发国际纠纷。

水利工程历来与经济民生密不可分,美国从 1919 年开始研究跨流域调水,用 60 年时间完成了中央河谷工程,一举解决了南北淡水资源差异性,我国历史上的京杭大运河、都江堰、郑国渠等大型水利工程也都产生了积极影响。将青藏高原的水引入新疆的设想,或许能在一定程度上解决我国水资源分布不平衡的问题,值得进一步论证,但这样的大型水利工程,牵扯了方方面面的问题,必须谋而后动,量力而行。

### 前沿观察

#### 花儿为啥这么艳

绚丽多姿的花朵装点着大自然,使大自然变得光彩夺目。那么,是什么让花朵绽放出了浓妆?日本的一项研究揭开了其中的奥秘,是存在于其中能够促进色素合成的蛋白质,使花朵开得更加艳丽。

众所周知,百花之所以能够怒放,色彩斑斓,与其中被称为花色苷的色素有关,研究人员通过对牵牛花遗传基因的寻查,发现了被命名为 EFP(黄酮类化合物促生因子)的新的蛋白质,该蛋白质能够提高 3 倍左右花色苷的生产效率,促进花色苷的合成,从而使花朵变得艳丽。

已知花色苷具有很好的抗氧化作用,还被用于健康食品,而且颜色越浓,含量越多。因此,调整花色苷的含量,经过改良,即可开发出具有新的附加值的鲜花和果实的品种。

#### 新蜂种缘何带巨针

日前,芬兰图尔库大学的研究小组发表了一项研究成果,告知他们发现了带有巨大针的新种蜜蜂。从照片看上去,这根由普通蜜蜂躯体上伸出的螫针,足有蜂体的一半,着实让人有点恐怖。

该大学从事生物多样性研究的学者表示,发现的新种蜜蜂是一种寄生蜜蜂,是在秘鲁收集的诸多样本中发现的,宿主为蜘蛛。这根针不仅具有通常所说的螫刺作用,而且还用于孵化时作为对不动宿主蜘蛛的毁杀。用针把蜘蛛包围在蜘蛛网上,然后在体内产卵。产卵时,既能向宿主蜘蛛注入毒液,亦可作为“砒针”将蜘蛛网合上。

至于新种蜜蜂所带有的巨针,作为假说,据推测是用于插入树木表面的洞里,从而达到宿主蜘蛛,起到“叉子”的作用。有这种行为的寄生蜂虽然很多,而这新蜂种却以其针的粗大放射着异彩。长很普遍,但尖端异常粗壮则十分罕见,若按一般人类大小来计算,蜂针应该占身长的 50% 以上。(宁蔚夏编译)



带有巨大针的新种蜜蜂

### 奇趣天下

## 鲸类也有绝经期

自然界中绝大部分动物终其一生都不会停止生育,它们生殖器官的活性几乎与寿命一样长久。我们人类是个例外。最近,科学家新发现两种鲸类——白鲸和独角鲸也要经历绝经期。这将我们已知有绝经期的物种数量增加到了 5 个。除人类外,其他 4 种动物都是齿鲸,包括白鲸、独角鲸、虎鲸和短须鲸。

这项研究由来自英国艾克斯特大学、约克大学以及鲸鱼研究中心的研究人员进行。他们也在找寻鲸类动物进化出绝经期的原因。

“从进化的角度来说,一个物种如果会经历绝经期,它需要一个停止生育的理由,同时也需要一个绝经后继续活下去的理由。”论文第一作者、埃克塞特大学的 Sam Ellis 说。

经过 40 多年的详细研究,科学家对虎鲸的绝经期已经作了很好的记录。

“虎鲸停经的原因是其雄性和雌性后代都和它们的母亲生活在一起,所以作为一头雌性虎鲸,它所在的群体中会有越来越多的子孙。这种日益增长的亲缘关系意味着如果它一直生育,就会和自己的直系后代争夺食物等资源。”Ellis 解释说,“在停经后,虎鲸还需要继续生存的理由是,年长的雌性对其后代有很大的好处。例如,它们更了解要到哪里寻找食物,这有助于群体生存。”

此次的新研究使用了来自 16 种死亡鲸类的数据库,研究人员在较老的白鲸和独角鲸中发现了休眠的卵巢。

根据这些发现,研究人员预测,这两种鲸类的社会结构就像虎鲸一样,这意味着雌性白鲸和独角鲸在年老时也会生活在越来越多的近亲中。对人类祖先的研究表明,我们人类的情况

### 视觉瞬间



## 新疆首个世界地质公园开园

10月10日,新疆可可托海世界地质公园正式开园。它位于阿尔泰山南麓,额尔齐斯河源,由额尔齐斯大峡谷、三号矿、萨依布拉克、可可苏里以及卡拉先格勒五大景区组成,拥有峡谷河源、沼泽湿地、寒极湖泊等多种自然景色,是新疆第一家世界地质公园。图为三号矿景区。(新华社记者张骁龙摄)