Editorial

An open access database for Himalayan environmental management

Environmental management in the Himalayas requires creation of a database with collaborative public sharing of data

Sahotra Sarkar

From arid high terrain in the northwest, through the world's highest peaks in the mid-regions, to tropical wet forests in the southeast, the Himalayan region includes some of the most biologically diverse habitats on Earth as well as homes to a bewildering variety of cultures. During the last half-century it has increasingly become clear that though some parts of the region are ecologically fragile, while others remain relatively intact, all are facing the uncertain effects of climate change. Recent research has established, though not beyond controversy, that the popular model of catastrophic environmental degradation in the Himalayas due to overpopulation and deforestation has little empirical support: understanding the dynamics of environmental change requires attention to socio-political oppression, resource appropriation by outsiders, and other social factors (Ives 2006). Consequently, Himalayan environmental planning must be based on region-wide data on socio-economic structures besides the customary data on physical and biological features, including the geographical distribution of all data. Because ethically appropriate planning must accommodate the cultural diversity of the region, besides its biological and physical diversity, the data sets used must necessarily be both extensive and also have a sufficiently high resolution to support planning at the local level at which most diversity is manifested. A vast amount of data must be collected and analyzed.

The purpose of this Editorial is to call for the collaborative creation of an open access database to help meet that challenge. The goal should be to provide free and reliable data sharing between researchers and planners throughout the region (and elsewhere) and to create a database that meets the strictest international standards for functionality. For biodiversity data, several such transnational databases exist (for instance, those maintained by the World Conservation Monitoring [http://www.unep-wcmc.org/] and NatureServe [http://www.natureserve.org/]) though they all provide sporadic coverage and are remarkably poor in Himalayan data.

Creating a dedicated Himalayan database will break new ground. But such a project has to be implemented carefully. Specifically, the database should conform to the following six principles:

- The database must be open access with public sharing of all data. We are living in an age of increased recognition that academic and intellectual exchange is not well served by proprietary attitudes towards data and software—witness the success of projects such as R, OpenBugs, and GenBank, to name just a few of the many successful open access initiatives. Increasingly journals are requiring software to be open access as a condition for publication. Throughout the world, the trend is towards the non-proprietary creation of intellectual products.
- The effort to create the database must be truly collaborative, drawing in researchers from all

> Reliable analyses of environmental processes and dynamics on a regional scale require the availability and easy accessibility of relevant data from across the region. To understand the mechanisms behind environmental processes and dynamics so as to make accurate predictions of the future, we need adequate information on the physical, biological and socio-economic features of a region. In the Himalayan region we need to collect, collate, and provide public access to such data on a regional scale...

the Himalayan regions. There have been many transnational efforts in the Himalayas in recent years (Wikramanayake et al. 2001, CEPF 2005, WWF and ICIMOD 2005) but coverage has not been uniform over the entire region. For instance, there has been less effort directed to Sikkim or Arunachal Pradesh than to many other parts of the Indian Himalayas and to Nepal. As the database gets developed, these biases must be corrected to ensure homogeneous coverage of all regions.

- All data must be geo-referenced and have their provenance recorded. This is the most important requirement that the database must satisfy. Almost all environmental planning today is spatial. Recording locational co-ordinates with GPS units has become technologically trivial and inexpensive and there is no reason not to record such data systematically in every project. Unfortunately there is at present very little high resolution geo-referenced biodiversity data for the Himalayas and this is a major impediment to scientific research. Almost none of the existing records in museums and other databases provide precise longitude-latitude co-ordinates. With respect to socio-economic data, all that often exists are summaries at the level of political units such as sub-divisions, districts, and provinces. This situation must be changed. It is equally important to record the method of collection of the data so that future researchers can independently assess the reliability of data sets and their appropriateness for their specific research needs.

- The database should be made available at several different mirror sites both within the Himalayan region and outside, wherever there is significant research and policy interest in the region. Several universities, research institutes and academies within the region, especially in China, India, and Nepal, as well as the International Centre for Integrated Mountain Development (ICIMOD) are natural host sites and should be brought into the discussion as the database is planned.

- The process of creating the database will require significant commitment to the information technology infrastructure of the host sites within the Himalayan region. Developing this infrastructure should be accompanied by relevant technology transfer and capacity building, especially the transfer of knowhow to young researchers from the Himalayan region. Collaborators from Northern institutions must make an explicit commitment to the training of local personnel, and this commitment must be monitored.

- Finally, care must be taken to ensure the proper design of the database so that eventual growth in size does not destroy functionality. The software must be stable. Database queries should receive fast answers. There must be variable ways of searching, for instance, based on taxonomic or geographical specifications. Data should be downloadable in a variety of formats. Contributors should be able to upload data and edit their own information easily following a straightforward protocol. There are many other such criteria and these

> ...Analyses based on such data will both advance our understanding of regional environmental challenges and also let us assess the geopolitically convenient explanations of environmental problems that the countries and institutions in the region have long advocated. It is high time we create a collaborative open-access database that can receive public input and access to data while meeting international standard of functionality.

should be discussed and correctly implemented from the outset.

An example of a collaborative open access database is the Latin American Biodiversity Database (http://www.consnet.org/biodiversity/) which does not include socio-economic data. However, it is a helpful pointer to where to take efforts to initiate a database project for the Himalayas.

The expertise and technological and financial resources required to create such a database are not particularly daunting. However, for success, co-operation will be necessary from regional and international museums, herbaria, and other repositories of information from the region such as the Kew Gardens (http://www.kew.org/) and the California Academy of Sciences (http://www.calacademy.org/). Much of the traditional data stored in these repositories may be useless for most planning purposes because they are not georeferenced. However, they may still be valuable for historical analyses. Most importantly, the bulk of the data must come from individual researchers who must be brought into collaboration in all aspects of the project—including its design and management—so that everyone remains committed to the goals of data sharing and open access. There should be no doubt that such a project would be valuable and that it can be implemented—if there is the will.

Sahotra Sarkar is a professor of biology (Section of Integrative Biology) and philosophy (Department of Philosophy) at University of Texas at Austin, USA.

References

CEPF [Critical Ecosystem Partnership Fund] 2005. *Ecosystem profile: Indo-Burman hotspot, Eastern Himalayan region*. Washington: WWF US-Asian Program. 97 p

Ives JD. 2006. *Himalayan perceptions: Environmental change and the well-being of mountain peoples*, 2nd ed. Lalitpur (Nepal): Himalayan Association for the Advancement of Science. 284 p

Wikramanayake E, Dinerstein E, Loucks C, Olson D, Morrison J, Lamoreux J, McKnight M and P Hedao. 2001. *Terrestrial ecoregions of the Indo-Pacific: A conservation assessment*. Washington: Island Press

WWF and ICIMOD. 2001. Ecoregion-based conservation in the Eastern Himalaya: identifying important areas for biodiversity conservation. Katmandu: WWF Nepal. 178 p