



## **A paleoclimate database for the Indian subcontinent**

Atreyee Bhattacharya (1,4), Kushank Bajaj (2,3), Supriyo Chakraborty (2), Aditi Chatterjee (3,4), and Amir Bazaz (4)

(1) University of Colorado Boulder, Environmental Studies, United States (atreyee.bhattacharya@colorado.edu), (4) Indian Institute for Human Settlements, Bangalore, India, (2) Indian Institute of Tropical Meteorology, Pune, India, (3) ETH, Zurich

Science has entered the big-data era, where large volumes of data are mined to gain insight into complex problems. The field of paleoclimate is no different. Paleoclimatology attempts to characterize the causes and patterns of past changes in the climate system; as such, the discipline relies on observations across a myriad of spatial and temporal scales. In this line of work, independent scientists collect proxy data in natural archives, representing small geographic domains. Over time now, collectively, paleoclimate investigators have created records representing long temporal coverage (centuries to hundred of thousands of years) of local observations from several locations, globally. However, investigation of large scales of climate variability, which require assimilating disparate data sets, have limited means of using paleoclimate data effectively, simply because of the lack of appropriately curated portals of these local paleoclimate observations. In short, in absence of a systematic method of curating, access, re-use, utility, and application of paleoclimate data in decision-making becomes far more restricted than it should be.

Since the early 2000s, to share data, paleoclimate investigators have been making their data available to the community, through online databases such as the <http://www.ncdc.noaa.gov/paleo/wdc-paleo.html> and <http://www.pangaea.de/>. Existing portals are, however, based in the United States and Europe. In this work, conducted over the past year and a half, we have made an effort to create a paleoclimate database-of publications and related datasets-specific to the Indian sub-continent, so that investigators and planners, interested in the issues of the region are able to access regional data efficiently. Although the scope of the database is regional, the vision for creating the database is global. For example, the goal of the Indian paleoclimate database are

- (a) Making regional paleoclimate data easily accessible to the community
- (b) Creating a forum for investigators to share the data
- (c) Allowing easy integration of paleoclimate data with existing international portals
- (d) Identifying knowledge gaps and creating awareness for funding work to fill the knowledge gaps
- (e) Developing statistical tools to fill data gaps
- (f) Standardizing the database with emerging language and format
- (g) Creating climate products that would be of use to the scientists as well as environmental planners
- (h) Developing computation codes to visualize the data

The paleoclimate database currently comprises of information retrieved from peer-reviewed publications and institutional annual reports. Our initial review of the existing data suggests that there exists major gaps in coverage in north-eastern, eastern, interiors of western regions, and central parts of India. Except in the western, southern and south-central parts of the country, where a combination of tree ring, stalactite and lake-based reconstructions provide critical information, paleo data on timescales of human interest- sub decadal, decadal, multi decadal and centennial- are minimal. We have also identified that due to the lack of data coverage, especially of proxy observations in natural archives that are able to provide information relevant to timescales of human interest (such as tree rings, corals, lacustrine archives and glaciers), key components of climate dynamics remain difficult to constrain.