



STRADITIZE: An open-source program for digitizing pollen diagrams and other types of stratigraphic data

Philipp S. Sommer, Basil A.S. Davis, and Manuel Chevalier

University of Lausanne, Institute of Earth Surface Dynamics (IDYST), Davis Group, Chavannes-près-Renens, Switzerland
(philipp.sommer@unil.ch)

In an age of digital data analysis, gaining access to data from the pre-digital era – or any data that is only available as a figure on a page – remains a problem and an under-utilized scientific resource. Whilst there are numerous programs available that allow the digitization of scientific data in a simple x - y graph format, we know of no semi-automated program that can deal with data plotted with multiple horizontal axes that share the same vertical axis, such as pollen diagrams and other stratigraphic figures that are common in the Earth sciences. STRADITIZE (Stratigraphic Diagram Digitizer) is a new open-source program that allows stratigraphic figures to be digitized in a single semi-automated operation. It is designed to detect multiple plots of variables analyzed along the same vertical axis, whether this is a sediment core or any similar depth/time series.

The program is written in python and supports mixtures of many different diagram types, such as bar plots, line plots, as well as shaded, stacked, and filled area plots. The package provides an extensively documented graphical user interface for a point-and-click handling of the semi-automatic process, but can also be scripted or used from the command line. Other features of STRADITIZE include text recognition to interpret the names of the different plotted variables, the automatic and semi-automatic recognition of picture artifacts, as well as an automatic measurement finder to exactly reproduce the data that has been used to create the diagram. Evaluation of the program has been undertaken comparing the digitization of published figures with the original digital data. This generally shows very good results, although this is inevitably reliant on the quality and resolution of the original figure.