User's Manual

The Matlab software package *lichen* was developed to conduct lichenometrical dating. As outlined in the main text, this modeling framework essentially comprises two steps: (1) inferring the growth curve using lichen size data obtained from rock surfaces of known age, and (2) determining the age of an undated rock surface using the modeled growth curve and the lichen size data thereof. An inverse modeling approached was adopted to infer the calibration curve. Specifically, a suite of model parameters generating lichen size distributions that best match the observational data was sought by minimizing the total Jensen-Shannon divergence.

Once the optimal model parameters were obtained, the Fokker-Planck equation associated with the stochastic differential equation that governs lichen growth was solved again to predict an optimal growth curve, which provides a probabilistic link between surface age and lichen populations. Given a growth curve, the probability density distribution of ages of a rock surface can be obtained by calculating the Jensen-Shannon similarity between the growth curve and the size data. The 68.2 % and 95.4% highest posterior density (HPD) regions as well as the median-probability age were also reported.

Running the package is quite straightforward. The only file that the users need to modify is *lichen_main.m*, which provides a user-friendly textual interface (Fig. 1) allowing data entry, specification of model parameters, and visualization of the results.

```
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%% Inverse Model of Lichen Growth Curves                                 %%
%% Copyright: Shiyong Yu                                                 %%
%% E-mail: shiyong.yu@gmail.com                                          %%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
clear
clc;
%% Entering the data for sites of known age (calibration data)
SizeUnit = 'mm'; %Unit of lichen size
StudyArea = 'Huashan'; %Study area
AgeScale = 'b2k'; % Year scale of the ages to be reported (b2k or CE)
delta = 5; % nearest ages to be rounded
```

Fig. 1 Textual interface of the package

In the Matlab environment, users need to open this file first, and then specify the unit of lichen size, the study area, which are only used for plotting the results. Users also need to specify the age scale for plotting and reporting the results. Two scales are supplied: b2k (before year 2000) and CE (common era). The latter is a universal scale, while the former is arbitrary – users may convert this scale to other scale by subtracting the difference between 2000 and the “zero year” (e.g., 1950 for radiocarbon ages and the year that the data were obtained for other types of calendar ages).

Then users need to enter the age and size data obtained from sites of known age, which will be used to estimate the model parameters. Due to the variable data size for each site, a tab delimited spread sheet was not provided; instead, a structure was used to store the data (Fig. 2).

```
Site(1).age = 1984;
Site(2).age = 1976;
Site(3).age = 1961;
Site(4).age = 1844;
Site(5).age = 1835;
Site(6).age = 1820;
Site(7).age = 1787;
Site(8).age = 1776;
Site(9).age = 1762;
Site(10).age = 1696;
Site(11).age = 1670;
Site(12).age = 1593;
Site(1).size = [20,23,25,20,24,30,26,26,17,19,23];
Site(2).size = [26,29,32,29,31,32,27,29,22,35];
Site(3).size = [33,36,41,26,44,38,31,33,35,34];
Site(4).size = [70,63,61,56,66,64,58,63];
Site(5).size = [64,66,67,68,65,65,76,61,71];
Site(6).size = [68,74,72,79,70,68,65,71,73,69];
Site(7).size = [82,80,78,75,80,85,82,80,83];
Site(8).size = [79,85,92,85,83,75,77,86,96,85,68];
Site(9).size = [84,87,81,75,86,81,91,94];
```

Fig. 2 Textual interface for data entry of rock surfaces of known age

For example, in this illustrative case, the structure variable *Site* holds the age and lichen sizes of 12 rock surfaces. Users also need to assign the structure variable *sample* with information obtained from the rock surfaces to be dated such as sample ID and lichen sizes (Fig. 3).

```
%% Entering the data for sites to be datad
sample(1).ID = 'Grotto No. 24';
sample(2).ID = 'Grotto No. 2';
sample(1).size = [92,101,82,90,85,87,91,84,86];
sample(2).size = [41,55,47,48,51,56,44];
```

Fig. 3 Textual interface for data entry of rock surfaces to be dated

Note that users may change the spatial and temporal resolution of the modeled growth curve by resetting the variables *dx* and *dt* (default values are 1 mm and 1 year, respectively) as well as the spatial and temporal domains of the modeled growth curve by resetting *Xmin*, *Tmin*, *Xmax*, and *Tmax*. The default values of the first two floating parameters are 0 mm and 0 year, respectively, while the latter two were given automatically.

The rest of the code implements the analytical framework, which should not be modified. For example, function *parest* is for estimating parameters of the growth curve by minimizing the cost function (*costfun.m*) using the Nelder-Mead simplex algorithm, and function *growthpde* is for generating the optimal growth curve by solving the Fokker-Planck equation, which was visualized using function *plot_curve*.

The next step is to calculate the age of the undated rock surfaces using lichen size data thereof and the modeled growth curve. The probability density function of the lichen size data will be built up first. Two methods were provided. One is the empirical method, which was implemented using function *size2pdf_e* through numerical differentiation of the empirical cumulative distribution function of the data; the other is the curve fitting method, which was implemented by fitting a generalized extreme distribution (GEV) function to the data using function *size2pdf_gev*. As illustrated in the case study, both methods yield nearly identical results.

The probability distribution of the age of the undated rock surface was obtained by calculating the similarity of the growth curve and the lichen populations, which was implemented using function *D_JS*. Moreover, both the 68.2% and 95.4% credible intervals of the age distribution were calculated using the Bayesian highest posterior density (HPD) method, which was implemented using function *pdf2hpd*. The results were reported in terms of a structure variable *AGE* and also visualized using function *plot_ages*.

The code was tested as working in MATLAB R2019b and likely working in earlier versions too. Should you have any questions, please contact Shiyong Yu (E-mail: syu@jsnu.edu.cn or shiyong.yu@gmail.com).