Gentry's Forest Transect Dataset (Alwin H. Gentry)

Source of data

- Phillips & Miller (2002)
- http://www.wlbcenter.org/gentry_data.htm
- http://www.mobot.org/MOBOT/research/gentry/transect.shtml

Description of the dataset

Two datasets are provided here. The dataset `gentry197` is composed of forest inventories in 197 localities distributed around the globe (from 40.7°S to 60.6°N and from 127.8°W to 166.8°E); at each locality, 10 transects 0.01 h (2 m x 50 m) each were sampled, and all woody individuals (including lianas) with DBH (diameter at breast height) larger than 2.5 cm were measured and counted. The dataset `gentry.ea` is a subset of the previous one, with five localities from East and South-East Asia.



Al Gentry examined leaves during Conservation International's Rapid Assessment Program expedition to Ecuador in January 1991. Photo from NY Times.

Alwyn Howard Gentry (1945-1993) was an American botanist and taxonomist, with a wide range of activities. He developed the sampling design for a quick inventory of diversity in species-rich tropical forests, later applied also to other regions. In relatively homogeneous forest stand, he placed ten contiguous transects of the size 2×50 m (100m^2), with the long and narrow shape allowing fast census. Within each transect, he measured DBH of each individual of woody species and lianas above a certain threshold[1)]. The original data thus contains information about DBH of each individual in each transect. In the example dataset provided here, I extract only the number of individuals of given species in each transect, but not DBH of each individual.

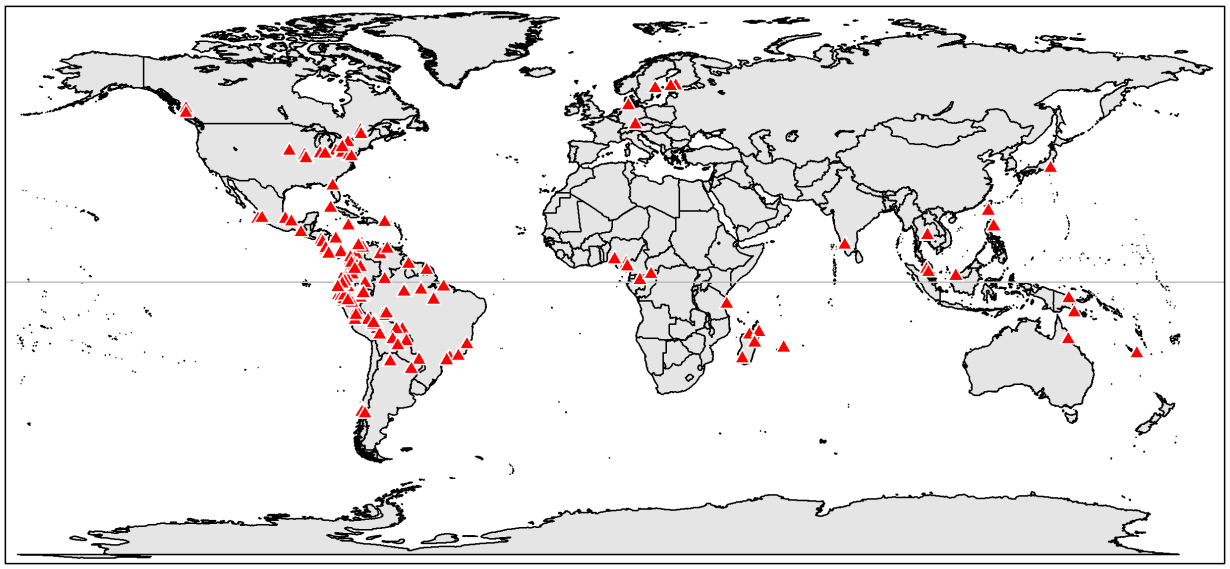
After Gentry's tragic death (in a flight accident on the way to Ecuador), the dataset was made available on the website mentioned above and also published in printed version (Phillips & Miller 2002).

Data from 225 forest plots are available on http://www.wlbcenter.org/gentry_data.htm. I prepared them for use in R (check this script if you want to know how). From the total of 225 plots, I removed those made by a different method (or containing less than ten transects), arriving at the number of 197 plots used for further analysis. Data contain many errors, which call for a manual treatment - I haven't attempted to do it for the purpose of this exercise [2)]. The dataset of 197 localities prepared for use in R is available below. The subset of five localities from East and South-East Asia is also available.

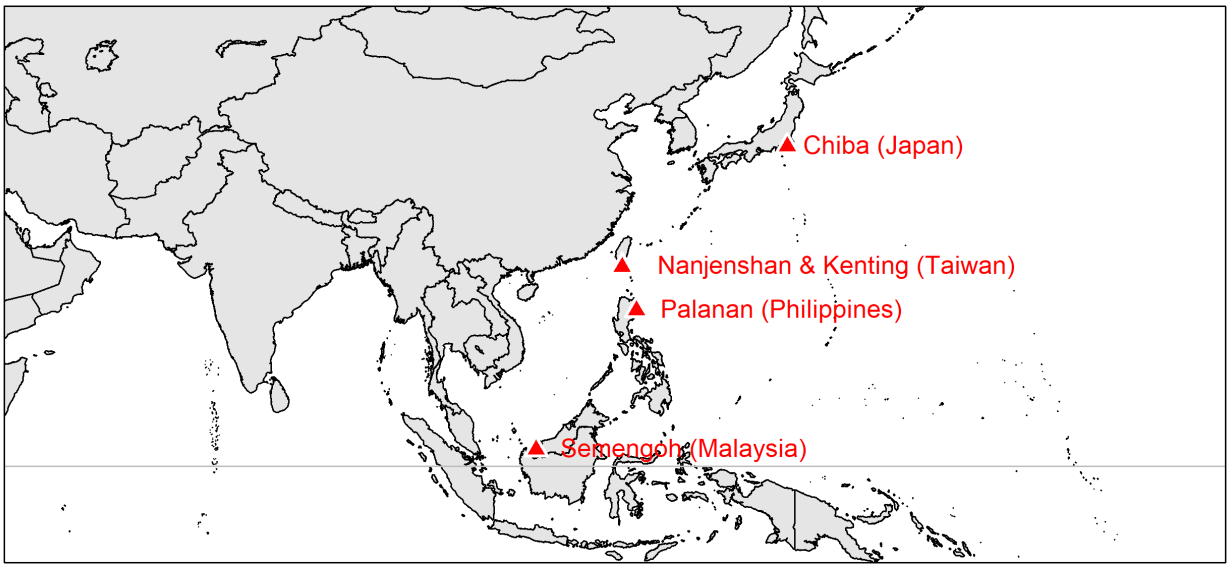
Locality

Highest concentration of localities are in South and North America, however, Gentry has other localities scattered around the globe (link for coordinates is here).

Distribution of 197 localities around the World (used in this dataset):



Distribution of five localities in East and South-east Asia (note that the two localities in Taiwan, Kending and Nanjenshan, are so close to each other that on the map they appear as one symbol):



Environmental variables

Original data contain, along to coordinates, also altitude and average year precipitation - see table available in Phillips & Miller (2002).

This table have been manually retyped into electronic form, available here as data frame gentry.coord.txt, with the following variables:

Name of variable	Description
Lat	Latitude [°] (-40.72°-60.67°)[3]
Long	Longitude [°] (-127.83°-166.83°)
Elev	Elevation [m a.s.l.] (10-2940 m asl, some values are missing)[4]
Precip	Average year precipitation [mm]] (400-9000 mm, some values are missing)

Data for download

File name	File type	Description
gentry197.RData	R object (list)	The list of data frames, each data frame contains species × sample matrix, values are the numbers of individuals of given species in the given transect. Data from 197 localities.
gentry.coord.txt	tab-delimited txt format	Table with with latitude, longitude, elevation and precipitation. Data from 197 localities.
gentry.ea.RData	R object (list)	The list of data frames, each data frame contains species × sample matrix, values are the numbers of individuals of given species in the given transect. Data from 197 localities. Data from five plots from East and South-East Asia (Japan, Taiwan (2x), Philipines, and Malaysia)
gentry.coord.ea.txt	tab-delimited txt format	Table with with latitude, longitude, elevation and precipitation. Data from five plots from East and South-East Asia (Japan, Taiwan (2x), Philipines, and Malaysia)

Script for direct import of data into R

Upload all 197 plots:

```
# load the object gentry197 (list of 197 elements)
load (url
('https://raw.githubusercontent.com/zdealveindy/anadat-r/master/data/gentry197.RDat
a'))

# load data frame with plot coordinates
gentry.coord <- read.delim
('https://raw.githubusercontent.com/zdealveindy/anadat-r/master/data/gentry.coord.t
xt', row.names = 1)
```

Upload the subset of five plots in East and South-East Asia:

```
# load the object gentry.ea (list of 5 elements, c('chiba', 'nanjensh', 'kenting',
'palanan', 'semengoh'))
load (url
('https://raw.githubusercontent.com/zdealveindy/anadat-r/master/data/gentry.ea.RDat
a'))

# load data frame with plot coordinates
gentry.coord.ea <- read.delim
('https://raw.githubusercontent.com/zdealveindy/anadat-r/master/data/gentry.coord.e
a.txt', row.names = 1)
```

The script loads an R object `gentry197.RData` in case of the whole dataset and `gentry.ea.RData` in case of asian subset into R; it will appear in the working space as a list `gentry197` with 197 components or `gentry.ea` with five components, respectively. Each component represents data from one locality (composed of 10 transects). The preview of the data frame from the first locality in `gentry197` looks like this (each row = transects, only 5 columns = species) :

```
gentry197[[1]][,1:5]
```

```
   ANACARDIACEAE M1 M1 ANNONACEAE Monanthotaxis M1 ANNONACEAE Polyalthia henriceii
APOCYNACEAE Landolphia M1 APOCYNACEAE Oncinotis M1
1                     0                           0                                1
2                         1
2                     0                           0                                0
2                         0
3                     0                           0                                0
0                         0
4                     0                           0                                0
1                         0
5                     1                           0                                0
0                         0
6                     0                           0                                0
1                         0
7                     0                           0                                0
1                         0
8                     0                           0                                0
3                         0
9                     0                           3                                0
0                         0
10                    0                           0                                0
1                         0
```

References

- Phillips, O.L. & Miller, J.S. 2002. //Global patterns of plant diversity: Alwyn H. Gentry's Forest Transect Data Set. //St. Louis, Missouri, Missouri Botanical Garden Press. (link to Amazon.com)

1)
See this description by Boyle (1996) for more details.

2)
e.g. file SANSEBAS.XLS contains one empty column inserted among `Lines` and `Family`, which needs to be removed from *.xls file prior to application of the script. There are several such problems, and you need to follow error messages if trying to extract data by yourself.

3)
For the locality `sierraro` (Cuba), the latitude reported in the original table was fixed - it should be 22.83 N, not 22.83 S.

4)
If the original table reported a range of the values, here is reported the mean elevation.

From:
https://anadat-r.davidzeleny.net/ - **Analysis of community ecology data in R**

Permanent link:
https://anadat-r.davidzeleny.net/doku.php/en:data:gentry

Last update: **2021/12/24 16:23**