Use of the Python programs

Python is a programming language and to use the programs written in this language You need the Python interpreter. This interpreter is installed on Linux systems (as Ubuntu, Debian, Red Hat and so on) by default, while on Windows (test done on Windows 10) you need to install Python downloading it from the site https://www.python.org/downloads/windows/. Our programs were written in Python 2.7, to use them with Python 3.x may require some minor changes.
As declared in Materials and Methods, You need also to install the following Python libraries: NumPy (http://www.numpy.org/), SciPy (http://www.scipy.org/), Matplotlib (http://matplotlib.org/), Scikit-learn (http://scikit-learn.org). You can find all these libraries on the site http://www.lfd.uci.edu/~gohlke/pythonlibs/#numpy. Install first PIP and then the other packages (pip install package). Lastly You need the Python Image Library -PIL-. You can find the exe file to install it on the site http://www.pythonware.com/products/pil.

In Windows 10 search on the search window:"CMD". Under Windows 7 the menu selection is Start ‣ Programs ‣ Accessories ‣ Command Prompt.
The terminal with the command line will open. Move with the cd command to the directory containing Your data.
In Windows write c:\Phyton27\Python Geoprof3.0.2.py -M=M -d=dati_colera_imageJ_pronti_senza_Pesi.csv -m=Snow-cholera-map.jpg -o=Y
In Linux simplywrite from the terminal: Python Geoprof3.0.2.py -M=M -d=dati_colera_imageJ_pronti_senza_Pesi.csv -m=Snow-cholera-map.jpg -o=Y

The map, whose name is mapname may be a jpg file containing the map of the investigated area where the infection cases were recorded, while to dati_colera_imageJ_pronti_senza_Pesi.csv You should substitute Your own data obtained with Neuronmorpho (You can also produce the coordinates on the map by hand, but it takes much longer, of course). The fourth column of the csv file (You can open and modify it with Excel or Libreoffice Calc) contains the weight given to each position of the map and corresponding to the number of infection cases found in a given location.

How to produce the file with the infection cases in csv format: imageJ and its plugin neuronmorpho

The data about the positions of the cases on the map were acquired with Neuronmorpho (http://www.southampton.ac.uk/~dales/morpho/), a plugin of ImageJ (National Institute of Health; http://rsb.info.nih.gov/ij/), that can read a map position with a mouse click, building a csv file containing the coordinates x,y point by point. Weights (number of infection cases in the same address) can be added manually in the csv file. You can open it with Excel or Libreoffice Calc and add a fourth column (the first is the numbering of the cases; the second column the X coordinate and the third the Y coordinate. containing an integer number corresponding to the weights. In case all the points have the same weight, that is there is only ONE infection case in each place, the user

can insert the same weight in each location. The data can be analyzed with the Python script Geoprof3.0.2.py.

Example of usage of the software:
The python scripts were written by the authors and can be retrieved from the site www.unifi.it/caryologia/PapiniPrograms.html.

write on the command line (linux):
python Geoprof3.0.2.py -t=plain -d=dati_colera_imageJ_pronti_senza_Pesi.csv -m=Snow-cholera-map.jpg -o=Y

Alternative in windows (tested on windows 7):
You can download the folder Geoprofwin.zip on Your desktop. Dezip it in Your desktop. Then Open the command line in Windows 7 (Start Menu-Programs-Accessories-Prompt/command prompt). Then from the command line write:
cd Desktop
cd Geoprofwin
Geoprof3.0.2.exe -m=Snow-cholera-map.jpg -M=M -d=dati_colera_imageJ_pronti_con_Pesi.csv -o=Y
This will start the analysis using the provided map (-m, in this example SnoSnow-cholera-map.jpg; the type of distance Manhattan, to be used in cities as in the example: -M=M [if not in a city better to use -M=E, the euclidean distance]; the csv file containing Your data -d=dati_colera_imageJ_pronti_con_Pesi.csv; and Your output: -o=Y, meaning Yes, with full output).

CSV files:
It is a simple format meaning “comma separated values” for spreadsheets as Excel or Libreoffice Calc.

Authors will be very happy to help users who may be interested to use the software also by analyzing the data!