

# Geocomputation & Modelling for Vector-borne Disease

**Every Tuesday** starting 11<sup>th</sup> October

**4:00 - 6:45 pm Abuja, Nigeria**

8:00 - 10:45 am Seattle, WA, USA

5:00 - 7:45 pm Rome, Italy

**Every Thursday** ending 1<sup>st</sup> December

**4:00 - 6:45 pm Abuja, Nigeria**

8:00 - 10:45 am Seattle, WA, USA

5:00 - 7:45 pm Rome, Italy

**Giuseppe Amatulli, Instructor**

**Erin Stearns, Instructor**

**Tushar Sethi, Program Manager**



# Learning objectives

- Habitat suitability model built with open source tools
  1. Scalable
  2. Adaptable
- Your job? Learn the use of:
  - Open source tools
  - Data processing
  - GIS and remote sensing applications
  - **Machine Learning**
- Varying skill levels: We are here for each one of you – there are no silly questions or answers



# Qualifications, certificates & remuneration

- ❖ Learning for now and the future
- ❖ Applied skills
- ❖ Remuneration – The Carter Center
- ❖ Workshop & certificates

# Introductions (<3 min each)

- Name and where you come from....
- What are your personal interests and background?
- Final project / PhD thesis objectives / keywords?
- What data are you going to analyse?
- Not sure about your plan yet... no problem
- Do you have any experience with Linux OS or open source software?
- Do you currently use any programming language?
- What are your expectations of this course?

# Learning objectives – elaborated

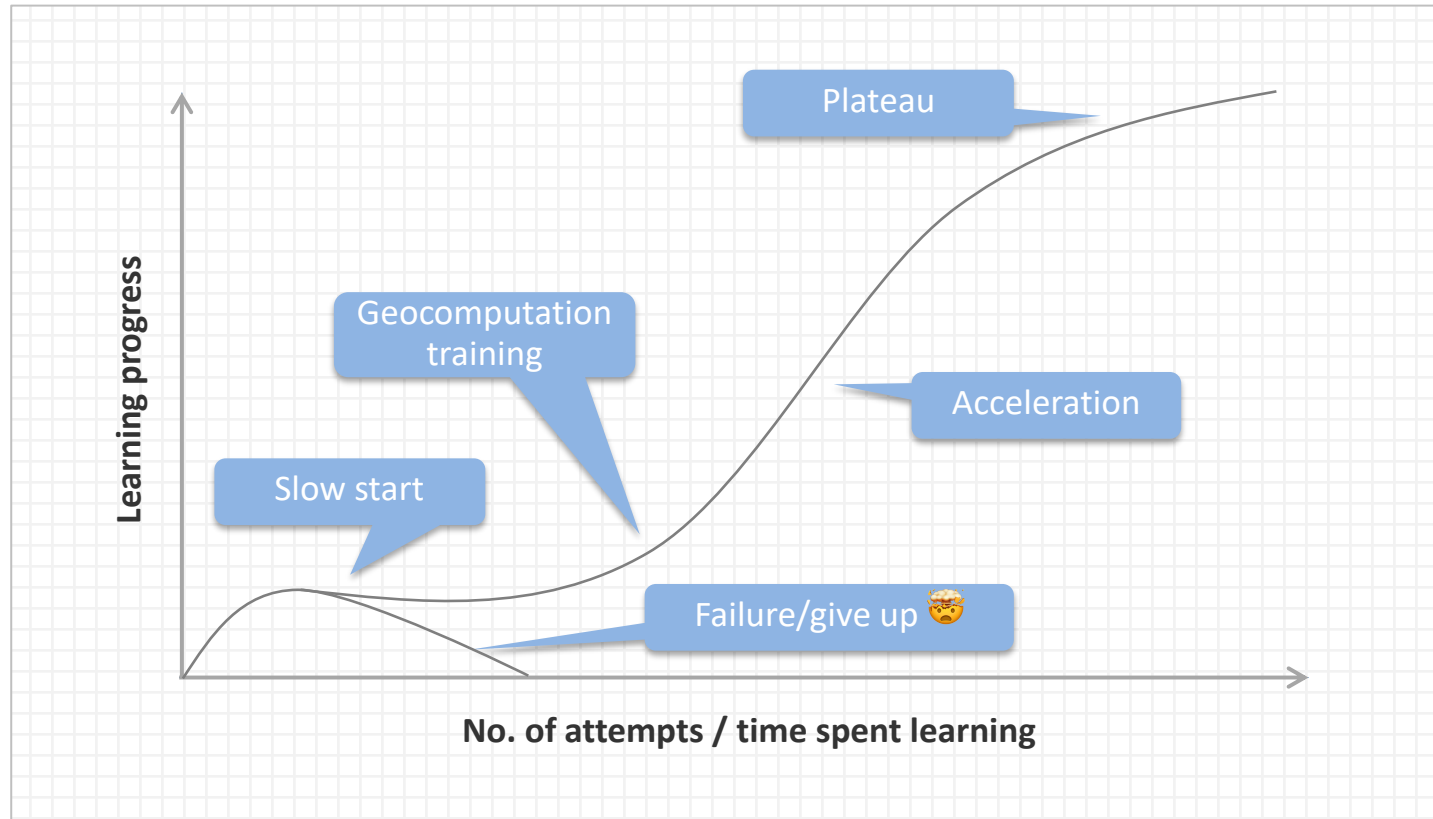
With continuous practice through the lectures, you will become familiar with new command lines and cover numerous topics, including:

- Learning open source tools for GIS and RS applications
- Acquiring command line utilities for spatial/temporal data under Linux OS
- Acquiring command line utilities and SDM theoretical foundation
- Developing data processing skills
- Independent learning, critical thinking, problem solving

Upon completion of the course, you should be able to:

- Apply the process of science, by conducting, analyzing, and interpreting findings related to GIS & RS project in the framework of SDM
- Use quantitative reasoning for statistical/spatial analysis
- Convey your understanding of environmental phenomena

# Learning curve



# Scientific knowledge

- ❖ Spatio-temporal analysis
- ❖ Spatio-temporal data integration
- ❖ Spatio-temporal modelling
- ❖ Geostatistics
- ❖ Species Distribution Modelling (Machine Learning)



# Tools

- Grass & QGIS:** Geographic Information Systems
- R:** GIS, statistic, modeling, text manipulation
- LINUX Bash:** Shell programming
- AWK:** Processing text-based data
- GDAL/OGR/pktools:** Geotools library for the manipulation of geospatial data



# Coding knowledge

## Covered on the course

- AWK, GDAL, PKTOOLS, R
- Parallel processing in bash and python environment
- Species Distribution modeling (SDM): theoretical foundation and application in vector-borne disease modelling
- Supervised regression/classification
- Image processing / raster processing / large data-set processing – efficiently!
- Environmental applications (hydrology, forestry)

# Syllabus: Key links

**Recorded lectures** (#2 to #7)

**Live lectures** (#8 to #16)

**On-site workshop**

## Class webpage

[http://spatial-ecology.net/docs/build/html/COURSESAROUNDTHEWORLD/course\\_geocomp\\_modelling\\_10-11\\_2022.html](http://spatial-ecology.net/docs/build/html/COURSESAROUNDTHEWORLD/course_geocomp_modelling_10-11_2022.html)

## Handling scripts and data via GitHub

[https://github.com/selvaje/SE\\_data](https://github.com/selvaje/SE_data)

## Recorded video lectures

[http://spatial-ecology.net/docs/build/html/COURSESAROUNDTHEWORLD/course\\_geocomp\\_modelling\\_10-11\\_2022.html](http://spatial-ecology.net/docs/build/html/COURSESAROUNDTHEWORLD/course_geocomp_modelling_10-11_2022.html)

## Community support for trouble shooting

<https://spatial-ecology.cloud.mattermost.com>



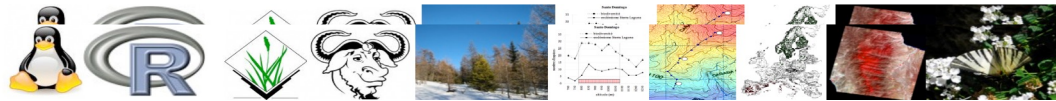
# Why use Linux/OpenSource?

**Security:** extremely stable and reliable, no viruses, interoperable: Unix, Windows, Mac, Android etc.

**Applications:** thousands of free programs, programming languages, server services

**Versatility:** minimum HW requirements, extremely portable, very fast performance

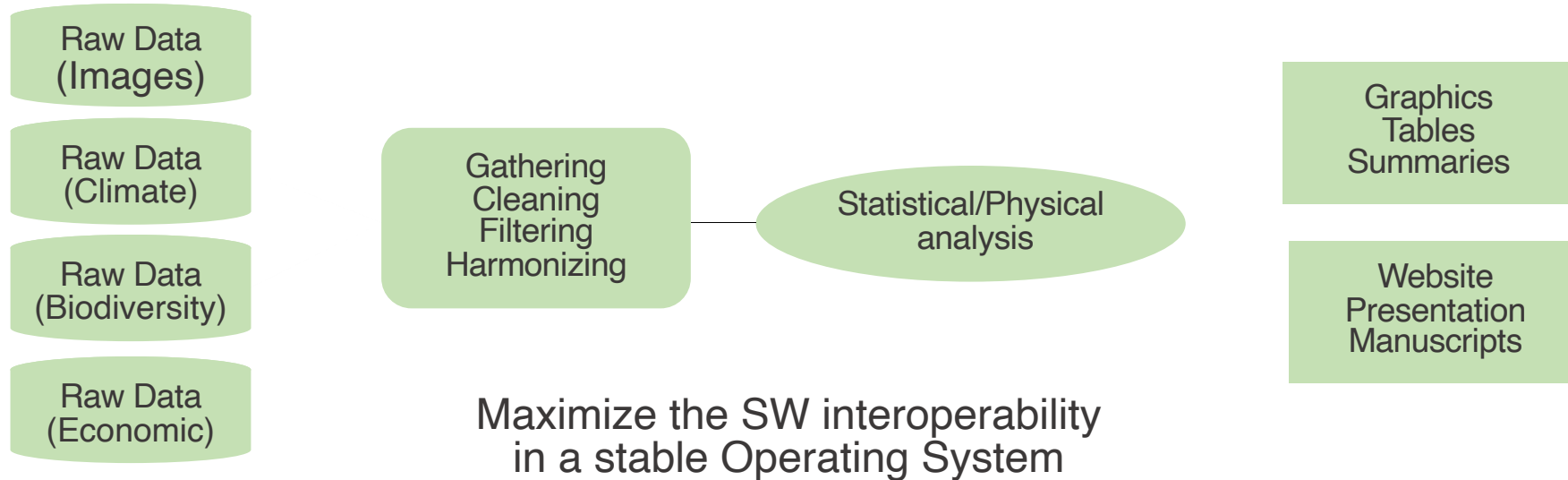
**Freedom:** free to download / test / install / modify / configure / develop / distribute... it's fun!



# Reproducible research & “big data” processing

**Codes that are easily published > no license constraints**

**Complex work-flows > integrate different data analysis methods**



# Freedom? and why it's fun

## Code

- Understating the code beyond a process
- Be able to modify the code
- Build up your own algorithm.
- Use all the SW that I want without license constraints

## Help

- Get help from mailing list
- Keep in touch with developers for code adjustment and improvement

## Process

- Job priority processing
- Job scheduling
- RAM management

## Remote server

- Automatic connection to remote servers
- Overpassing quota issues in remote servers, by creating a folder linked to your PC

## Hardware resources

- Storing temporal file in ram rather in the hard-disk, by creating a folder in the ram
- Get the best of different programing languages and create a unique work flow.

## Last but not least

- Enjoy life while the PC is working for you!



# Ubuntu Linux operating system

## Geocomputation with Open Source: Optimal latitude & interactivity

