MODELING IN PRACTICE: THE LIFE CYCLE OF A MODELING PROJECT, FROM CONCEPTION TO PUBLICATION - The example of Buruli ulcer in Cameroon -

Andrés Garchitorena
Researcher, Institut de Recherche pour le Développement
Associate Scientific Director, PIVOT Madagascar

E²M² Workshop, Ranomafana, March 2024
Steps in a modeling project

1. Development of the study concept and question
2. Literature review
3. Data collection
4. Construction of model framework
5. Model analyses and selection
6. Model validation
7. Manuscript writing and submission
Development of the study concept

• What is your question?
• Why is it interesting?
• Who is interested?
• Can it be narrowed down to a question about specific quantitative relationships?
Literature review

- Who has tried to answer this before and how did they do it?
  - Empirical studies
  - Modeling studies

- What are these studies short-comings?

- Are there already parameter estimates or data sets to help you answer your question?
Data collection

- What do you need to characterize?
  - Spatial and/or temporal dynamics
  - Relationships between parameters or systems
Construction of model framework

• What drawbacks of previous studies can I mitigate?
• What type of modeling is necessary to answer my question?
  o Statistical: GLM, spatial, time-series, etc.
  o Mathematical: population based, individual based
• What modeling elements are necessary for my question?
  o Stochasticity
  o Compartments and complexity
Model analysis, selection and validation

- What model(s) best fit my data and explain my question?
  - Comparison of alternative models and application of selection procedures
- Does the selected model suffer from any substantial drawbacks?
  - Statistical models: verification of model assumptions
  - Mathematical models: sensitivity analyses and out-of-sample predictions
Types of modeling studies

Without data collection

1. Purely theoretical studies
2. Parametrization based on published studies
   - Systematic reviews and meta-analyses
   - Experimental and field studies

1. Development of the study concept
2. Literature Review
3. Data collection
4. Construction of model framework
   - Dynamic equations and code
   - Relationships between parameters
5. Model analyses and selection
   - Parametrization
   - Simulations and debugging
6. Model validation
   - Model validation
   - Sensitivity analyses
7. Manuscript writing and submission
Types of modeling studies

1. Development of the study concept
2. Literature Review
3. **Data collection**
4. Construction of model framework
   - Statistical vs. Mathematical model
   - Model better adapted to our data
5. Model analyses and selection
   - Descriptive, univariate and multivariate
   - Parametrization and simulations
6. Model validation
   - Model validation, comparison
   - Sensitivity analyses
7. Manuscript writing and submission

**With data collection**

1. Data already collected for other purposes
   - Focus only on analyses
   - Need to understand data limitations and quality
   - Need to adapt modeling to the available data
2. Data collected for the modeling project
   - Very time consuming
   - Modeling is generally more straightforward
THE EXAMPLE OF BURULI ULCER IN CAMEROON
Buruli ulcer

Introduction

**Most affected**: Children <15 years

**25% cases with functional limitations**

Source of images: [www.who.int](http://www.who.int) (2014)
Buruli ulcer: an emergent and neglected disease

Introduction

Around 5000 new cases each year

Cases in more than 30 countries

Focal distribution

WHO meeting on BU control and research (2013)

Landier et al. (2014, PLoS NTDs)
1. LITERATURE REVIEW & IDENTIFICATION OF THE PROBLEM

What is my question?

Why is it interesting?

Who has tried to answer this before and how?

What are these studies short-comings?
Buruli ulcer: a disease linked to aquatic ecosystems

**BU Risk factors**

Proximity to stagnant or slow flowing waters

Activities near water

*M. ulcerans*
Buruli ulcer: a mysterious disease

Hypothesis 1
Direct transmission from the environment

Hypothesis 2
Aquatic biting bugs as vectors

Source: L. Marsollier
Objectives of the project

General objective
To gain insights on the ecological determinants of Buruli ulcer disease.

Specific objectives
1. To understand the effects of environmental factors on *M. ulcerans* ecology
2. To study the transmission of *M. ulcerans* from the aquatic environment to humans
What do I need to characterize?

Spatial and/or temporal dynamics?

Relationships between parameters or systems?

DATA COLLECTION & DESCRIPTIVE ANALYSES
1 Introduction

Characterization of MU in the environment

Favourable environments

Unfavourable environments
Regions of study

Akonolinga

- Landscape: Tropical rainforest
- Historically endemic area (>40 years)

Bankim

- Landscape: Savannah-Forest
- New endemic area (10 years)

Characterization of MU in the environment

Marion et al. (2011, EID)

Landier et al. (2014, PLoS NTDs)
1. **Fieldwork**: Environmental sampling

2. **Laboratory (CPC)**: Taxonomic identification & Pool composition

3. **Laboratory (Angers)**: DNA extraction & Amplification

**Characterization of MU in the environment**
**M. ulcerans** geographical distribution

<table>
<thead>
<tr>
<th>Type of ecosystem sampled</th>
<th>Percentage of Positive Pools</th>
<th>Elevation (m)</th>
<th>Basin</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stream</td>
<td>0</td>
<td>&lt; 300</td>
<td></td>
</tr>
<tr>
<td>Flooded Area</td>
<td>1-10</td>
<td>1-300</td>
<td></td>
</tr>
<tr>
<td>Swamp</td>
<td>10-20</td>
<td>300-700</td>
<td></td>
</tr>
<tr>
<td>River</td>
<td>&gt;20</td>
<td>&gt; 700</td>
<td></td>
</tr>
</tbody>
</table>

Garchitorena et al. (2014, PloS NTDs)
Seasonal fluctuations of *M. ulcerans* in freshwater ecosystems

Characterization of MU in the environment

Garchitorena *et al.* (2014, *PloS NTDs*)
What type of modeling is necessary to answer my question?

- **M. ulcerans** dynamics
- Environmental factors
- Buruli ulcer in humans

STATISTICAL ANALYSES TO UNDERSTAND M. ULCERANS ECOLOGY
Environmental drivers of *M. ulcerans*

- **Seasonal Climatic Factors**
  - Temperature
  - Rainfall
  - Others

- **Water Characteristics**
  - Temperature
  - Oxygen
  - pH
  - Conductivity
  - Others

- **Biotic Factors**
  - Abundance of species
  - Community assemblages
  - Trophic interactions
  - Others

- **M. ulcerans positivity**
Methodology: Multi-model approach

Model Definition
- Generalized linear mixed model (binomial)
- Random effect: Sample site

Multi-model Selection
- Screening of possible combinations in a given set
- Selection of the best models (AIC)

Multi-model Inference
- Model-averaged estimates
- Importance of terms
- Unconditional SE, lower and upper CI

Environmental drivers of MU
Environmental drivers of *M. ulcerans*: Akonolinga

### Variable Estimates

<table>
<thead>
<tr>
<th>Variable</th>
<th>Avg. beta</th>
<th>Lower.CL</th>
<th>Upper.CL</th>
<th>Relative Importance</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Intercept)</td>
<td>-13.66</td>
<td>-22.50</td>
<td>-4.82</td>
<td>1</td>
</tr>
<tr>
<td>SEASONALITY</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sin(2pi*Mois/12)</td>
<td>0.35</td>
<td>0.02</td>
<td>0.69</td>
<td>1</td>
</tr>
<tr>
<td>Sin(2pi*Mois/4)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cos(2pi*Mois/12)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cos(2pi*Mois/4)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Environmental Drivers

**Biotic Factors**

- **M. Ulcerans** positivity
- Seasonal Factors
- Water Characteristics

**Environmental Factors**

- Environmental drivers of *M. ulcerans*
- Akonolinga

**Seasonal Factors**

- *M. Ulcerans* positivity
- SEASONALITY

**Water Characteristics**

- AQUATIC COMMUNITY
- Abundance
- Relative Importance

<table>
<thead>
<tr>
<th>Variable</th>
<th>Avg. beta</th>
<th>Lower.CL</th>
<th>Upper.CL</th>
<th>Relative Importance</th>
</tr>
</thead>
<tbody>
<tr>
<td>AQUATIC COMMUNITY Abundance</td>
<td>-0.70</td>
<td>-1.06</td>
<td>-0.34</td>
<td>1</td>
</tr>
<tr>
<td>Biodiversity</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gastropoda</td>
<td>-22.27</td>
<td>-36.86</td>
<td>-7.68</td>
<td>1</td>
</tr>
<tr>
<td>Hemiptera</td>
<td>-7.33</td>
<td>-13.42</td>
<td>-1.23</td>
<td>0.92</td>
</tr>
<tr>
<td>Hirudinea</td>
<td>0.57</td>
<td>0.06</td>
<td>1.08</td>
<td>0.87</td>
</tr>
<tr>
<td>Diptera</td>
<td>1.84</td>
<td>-1.27</td>
<td>4.94</td>
<td>0.81</td>
</tr>
<tr>
<td>Hydracarina</td>
<td>18.07</td>
<td>-2.07</td>
<td>38.22</td>
<td>0.77</td>
</tr>
<tr>
<td>Oligochaeta</td>
<td>0.27</td>
<td>-0.28</td>
<td>0.82</td>
<td>0.76</td>
</tr>
<tr>
<td>Tricoptera</td>
<td>1.29</td>
<td>-5.20</td>
<td>7.78</td>
<td>0.73</td>
</tr>
<tr>
<td>Coleoptera</td>
<td>3.13</td>
<td>-0.31</td>
<td>6.57</td>
<td>0.49</td>
</tr>
<tr>
<td>Odonata</td>
<td>1.35</td>
<td>-1.92</td>
<td>4.62</td>
<td>0.44</td>
</tr>
<tr>
<td>Anura</td>
<td>-3.83</td>
<td>-9.11</td>
<td>1.45</td>
<td>0.26</td>
</tr>
<tr>
<td>Ephemeroptera</td>
<td>3.08</td>
<td>0.22</td>
<td>5.93</td>
<td>0.26</td>
</tr>
<tr>
<td>Decapoda</td>
<td>-1.01</td>
<td>-1.81</td>
<td>-0.21</td>
<td>0.21</td>
</tr>
</tbody>
</table>

**PHYSICOCHEMICAL PARAMETERS**

- pH
- Flow
- Temperature
- Dissolved Oxygen
- Conductivity

- Rainfall
- 3-Month Cumulative Rainfall
- Seasonal Effect
Environmental drivers of *M. ulcerans*: Bankim

**Biotic Factors**

**M. Ulcerans positivity**

**Seasonal Factors**

**Water Characteristics**

### Variable | Avg.beta | Lower.CL | Upper.CL | Relative.Importance
--- | --- | --- | --- | ---
(Intercept) | -10,13 | -18,94 | -1,32 | 1

<table>
<thead>
<tr>
<th>PHYSICO-CHEMICAL PARAMETERS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Water Flow (lentic)</td>
</tr>
<tr>
<td>Water Flow (lotic)</td>
</tr>
<tr>
<td>pH</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Biotic Factors</th>
</tr>
</thead>
<tbody>
<tr>
<td>Comp3</td>
</tr>
<tr>
<td>Comp1</td>
</tr>
<tr>
<td>Comp2</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Community</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abondance</td>
</tr>
<tr>
<td>Shannon</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Aquatic Taxa (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Coleoptera</td>
</tr>
<tr>
<td>Odonata</td>
</tr>
<tr>
<td>Diptera</td>
</tr>
<tr>
<td>Anura</td>
</tr>
<tr>
<td>Fish</td>
</tr>
<tr>
<td>Hydracarina</td>
</tr>
<tr>
<td>Hirudinea (Presence)</td>
</tr>
<tr>
<td>Tricoptera</td>
</tr>
<tr>
<td>Ephemeroptera</td>
</tr>
<tr>
<td>Gastropoda</td>
</tr>
<tr>
<td>Decapoda (Presence)</td>
</tr>
<tr>
<td>Oligochaeta (Presence)</td>
</tr>
<tr>
<td>Hemiptera</td>
</tr>
</tbody>
</table>

### MU Positivity in Ecosystems

![Box plots showing MU positivity in different ecosystem categories](image)

**Categories (Comp 1):** Lotic, Lentic, Stagnant

**Prop Positive Pools:**
- Lotic: 0.04
- Lentic: 0.12
- Stagnant: 0.12
Why the two regions are so different?

Optimal pH for MU [5.8-6.5]

Environmental drivers of MU
Our theory

Scenario 1: Favourable physico-chemical conditions

Free living stages & Environmental transmission to aquatic organisms

↓ Water flow
↓ O2
↓ pH (optimal)

Environmental drivers of MU
Scenario 2: Adverse physico-chemical conditions

Mostly intra-host & Trophic transmission
What type of modeling is necessary to answer my question?

MATHEMATICAL MODELING TO UNDERSTAND BU TRANSMISSION
Introduction

2 possible routes of transmission to humans
Introduction

Dynamic model

Transmission of MU to humans
Temporal data

Transmission of MU to humans

10 year database

BU cases diagnosed at Akonolinga Hospital

Graph showing Buruli Ulcer cases over months.
Mathematical model framework

- **MU in the environment**
- **MU in waterbugs**

Transmission of MU to humans

\( \beta_{CH} \)  
\( \beta_{VH} \)  
\( \delta \)  
\( \varepsilon \)  
\( \rho \)
Mathematical model framework

Mathematical Model

\[
\begin{align*}
\frac{dS}{dt} &= \mu N - \lambda_{CH}(Month_i) S - \lambda_{VH}(Month_i) S - \mu S \\
\frac{dE}{dt} &= \lambda_{CH}(Month_i) S + \lambda_{CH}(Month_i) S - \sigma E - \mu E \\
\frac{dI}{dt} &= \sigma E - \varepsilon I - \mu I \\
\frac{dT}{dt} &= \varepsilon I - \gamma T - \mu T
\end{align*}
\]

Transmission of MU to humans
Mathematical model framework

Model simulations to account for:
- A range of initial parameters
- Uncertainties in rates of incubation ($\delta$) and seeking treatment ($\varepsilon$)
- Different proxies of waterbug transmission and environmental transmission
- Linear risks or thresholds in the relationship MU-BU

Comparison of model fit using AIC and selection of best performing (2 AIC)
Results for Buruli ulcer temporal dynamics

Best temporal fit

Environmental transmission >>>>> water bug transmission

MU environmental concentration as linear predictor of BU cases
AT THIS STAGE WE ARE ALMOST DONE...
Manuscript writing and submission

• What are the main results that provide the answer to my question?
  • 1 to 3 graphs
  • 1 to 3 tables

• What is the journal that best fits my study?
  • Scope, audience, impact factor, math focus

• How do I present my manuscript?
  • Introduction: set the stage to your question
  • Methodology: describe explicitly all steps for replicability
  • Results: clear and concise
  • Discussion: explain how your study improves previous knowledge
MODELING IN PRACTICE: THE LIFE CYCLE OF A MODELING PROJECT, FROM CONCEPTION TO PUBLICATION

- The example of Buruli ulcer in Cameroon -

Andrés Garchitorena

Researcher, Institut de Recherche pour le Développement

Associate Scientific Director, PIVOT Madagascar

$E^2M^2$ Workshop
Ranomafana, January 2020