



Swiss Tropical Institute  
Institut Tropical Suisse  
Schweizerisches Tropeninstitut



## Case Study: [malariacontrol.net](http://malariacontrol.net)

Nicolas Maire  
Swiss Tropical Institute  
[asia@home 09](mailto:asia@home.09)





Swiss Tropical Institute  
Institut Tropical Suisse  
Schweizerisches Tropeninstitut



# Swiss Tropical Institute

- Research, services, and teaching and training, contribute to the improvement of the health of populations internationally and nationally
- Associated with University of Basel
- 2 research departments
  - Medical Parasitology and Biology of Infection
  - Public Health and Epidemiology



Swiss Tropical Institute  
Institut Tropical Suisse  
Schweizerisches Tropeninstitut



## Biostatistics and Epidemiology group

- Modeling malaria epidemiology and control
- Spatio-temporal statistical modeling of infectious diseases
- Demand for high performance computing
  - Simulation models of malaria
  - Mapping malaria transmission
  - Estimation of epidemiological parameters



## Simulation models of malaria transmission and control

- July 2003: Mathematical modeling of the impact of malaria vaccines on the clinical epidemiology and natural history of *Plasmodium falciparum* malaria
  - Funded by MVI/GSK
  - Effectively 2 year project
- June 2006: Simulation modeling of the epidemiological impact and cost-effectiveness of malaria interventions
  - Funded by BMGF
  - 3 year project
- March 2009: A stochastic simulation platform for predicting the effects of different malaria intervention strategies
  - Funded by BMGF
  - 2 year project

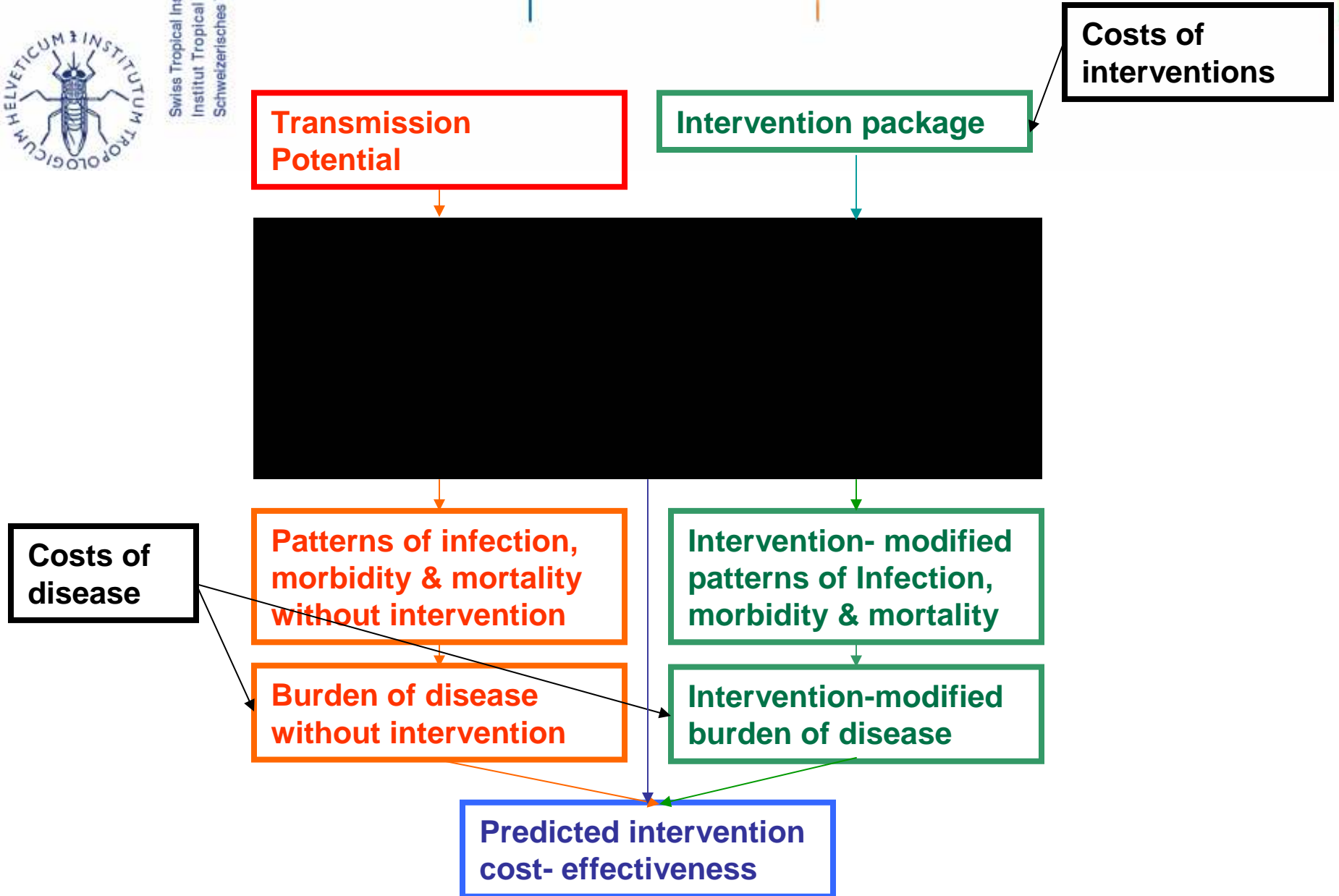


## Malaria and malaria models

- Malaria: a mosquito-borne disease caused by a parasite
- One to three million deaths per year
- Hundreds of millions of illness episodes per year
- Up to 40% of health expenditure
- Mathematical models have been a valuable decision making tool in public health, especially planning vaccination strategies



Swiss Tropical Institute  
Institut Tropical Suisse  
Schweizerisches Tropeninstitut





## Modeling approach

- Discrete time stochastic individual-based simulations
  - Human hosts are characterized by a set of state variables (age, parasite densities, immune status variables, infectiousness)
- Models for the effect of acquired immunity on parasite densities, for transmission to the vector, for morbidity, and for mortality, as functions of parasite density
- Fit model to data from field studies
- Predict impact of control strategies by comparing simulated interventions with baseline scenarios



Swiss Tropical Institute  
Institut Tropical Suisse  
Schweizerisches Tropeninstitut

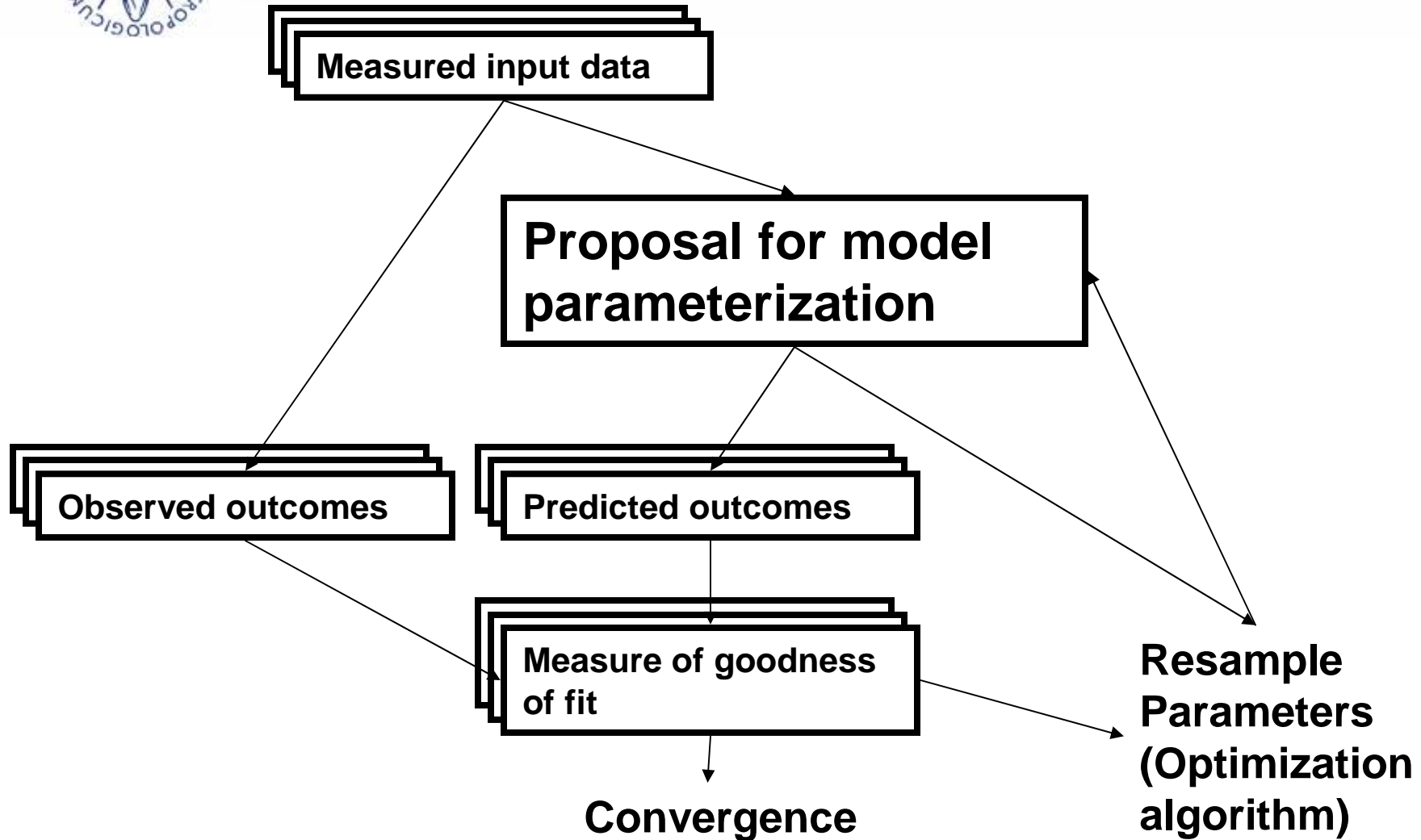


## Simulation objectives

- Fitting of model parameters to field data
- Prediction for a range of scenarios
- Sensitivity analysis
- Model comparison



# Estimating model parameters from field data

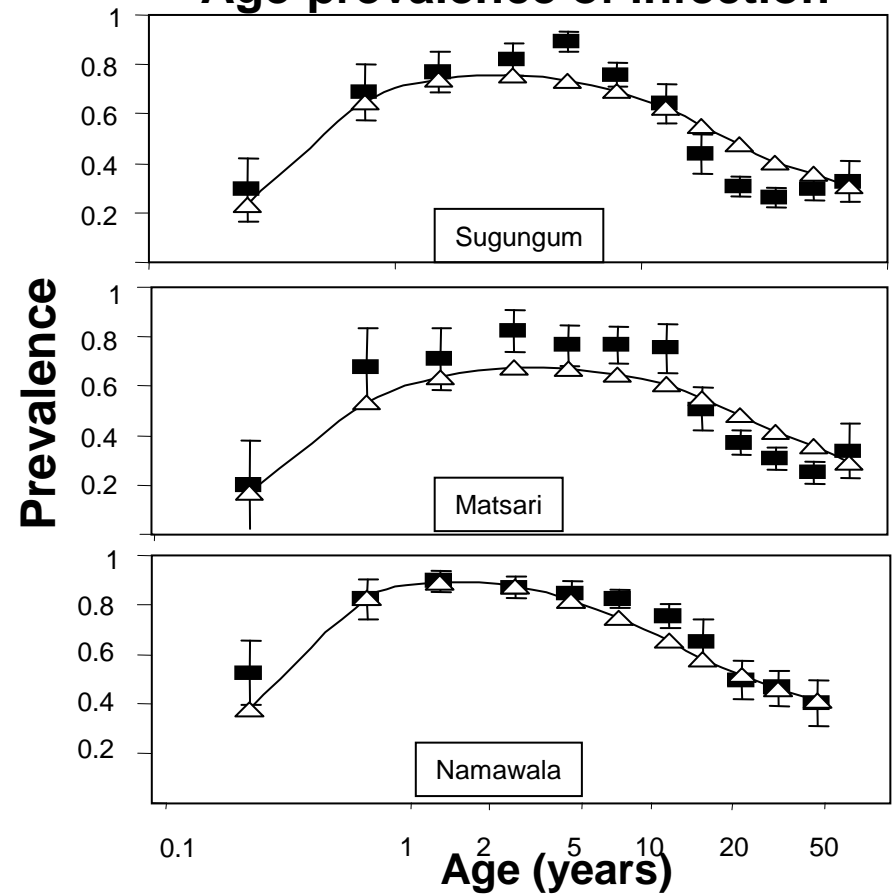




# Calibration

- 61 datasets from field studies, different objectives\*
  - Incidence of infection
  - Age-prevalence of parasitemia
  - Seasonality of parasitemia
  - Age-density of parasites
  - Age-incidence of clinical disease, hospitalisation and mortality

## Age-prevalence of infection

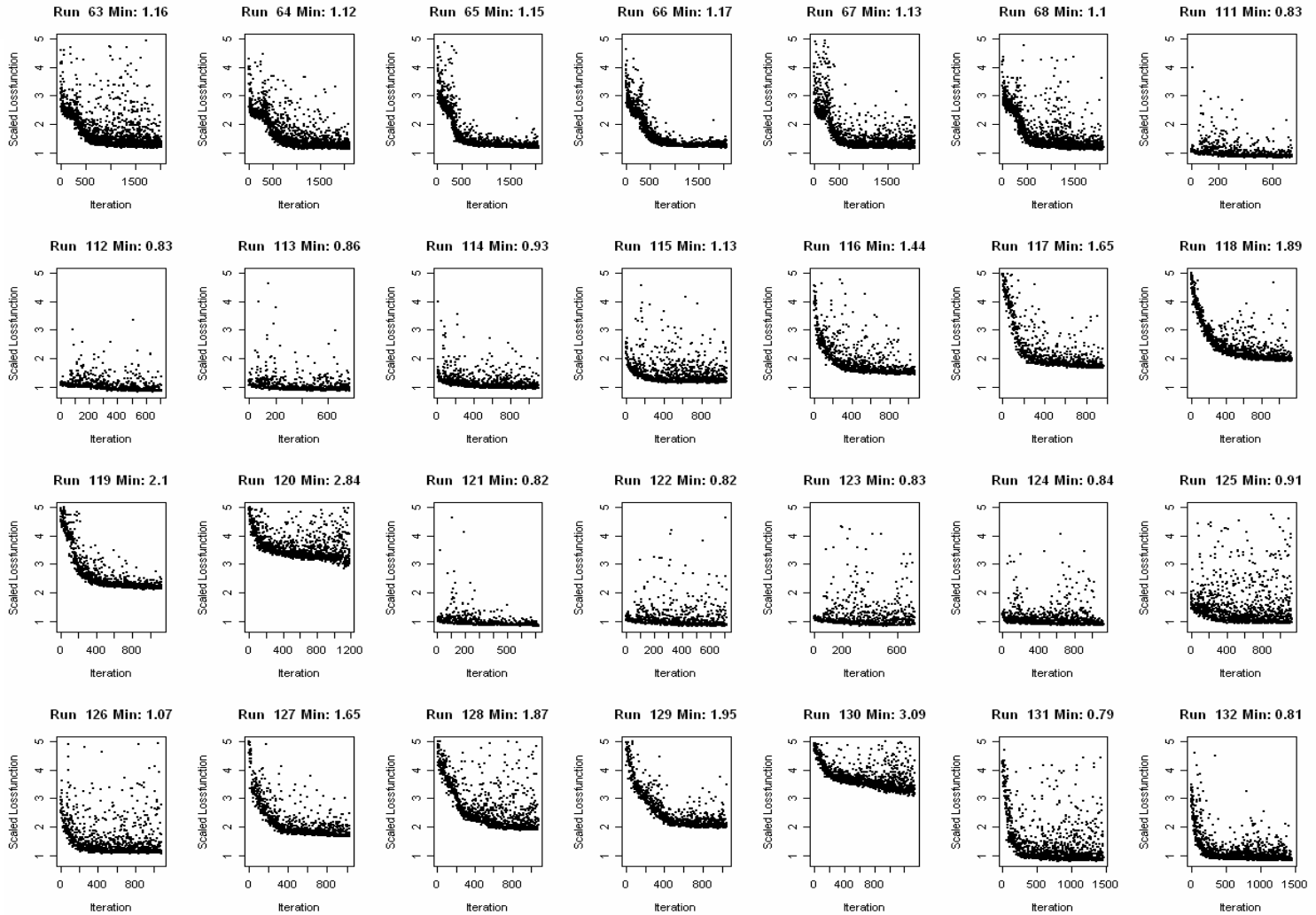


■ Field data (CI)    △ Model

\*all related to seasonal patterns of transmission



# Fit by iteration for alternative models





# Computation needs

- Several hours to run a single simulation
- Large number of simulation runs
  - 1'000s of iterations necessary to fit parameters to field data
  - 10'000s to 100'000s of simulations for predictive runs and sensitivity analysis



# **malariacontrol.net: Suitability for Volunteer computing**

- Independent parallelism
  - Divisible into parallel parts with few or no data dependencies
- Low data/compute ratio
  - Less than a gigabyte of data per day of CPU time
- Not dependent on short turnaround time
  - Several days per results, possibly resend a few times
- Public appeal



## From distrifit to malariacontrol.net

- April 2003: Fortran model implementation
- July 2003: Distrifit, a Java RMI based desktop grid, Intranet Swiss Tropical Institute
- Spring 2005: malariacontrol.net as the first application within the africa@home project (CERN, Uni Geneva, ICV, ISF, STI)
- August 2005: Student team in Geneva and Basel helps port the application to BOINC
- November 2005: Start of testing



## Port of existing application to BOINC

- Science application/Client
  - Reimplementation of some components
  - Communication with core client (BOINC-API)
  - Implementation of checkpointing
  - Screensaver graphics
- Project Server setup
  - Hardware provided by CERN/CUI Geneva
  - Hosting provided by CUI
  - Development of project specific server components



Swiss Tropical Institute  
Institut Tropical Suisse  
Schweizerisches Tropeninstitut



## Science application

- 2003: Fortran core model, Java Wrapper (XML-Databinding, GUI, RMI), C++/JNI-Bridge
- 2005: Fortran core, C/C++ Wrapper (XML-Databinding, BOINC-API)
- 2008: C++
- <http://openmalaria.googlecode.com>



Swiss Tropical Institute  
Institut Tropical Suisse  
Schweizerisches Tropeninstitut



## Project specific server components: Assimilator

- Processes validated results
- Reads the simulation output file
- Compares the predictions with the corresponding field data
- Computes a measure of fit



## Project specific server components: Work generator

- Creates new work if necessary
- Sample a new model parameterization and store it to the backend database
- Based on the completed parameterizations in the backend database
- Create workunit files, copy them to the download directory and all create\_work to add to the BOINC database



Swiss Tropical Institute  
Institut Tropical Suisse  
Schweizerisches Tropeninstitut



## malariacontrol.net statistics

**Volunteers:** 25'000 total, 10'000 active

- Sign up rate: up to 400 new users per day
- Currently no new users

**Host PCs:** 15,000 active,

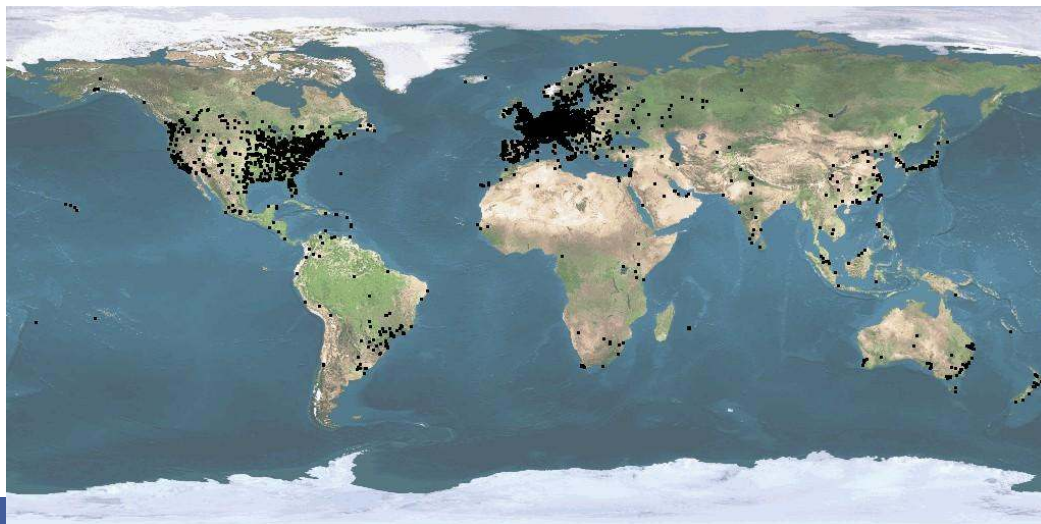
- 85% Windows, 15% Linux, Mac

**CPU power:10(5)**  
TeraFLOPS

- delivered to date  
10,000 CPU years (Oct 08)

**Simulations per day:**  
75,000

**...+ huge public/press interest!**





Swiss Tropical Institute  
Institut Tropical Suisse  
Schweizerisches Tropeninstitut



## Outlook I: Optimize parameter estimation process

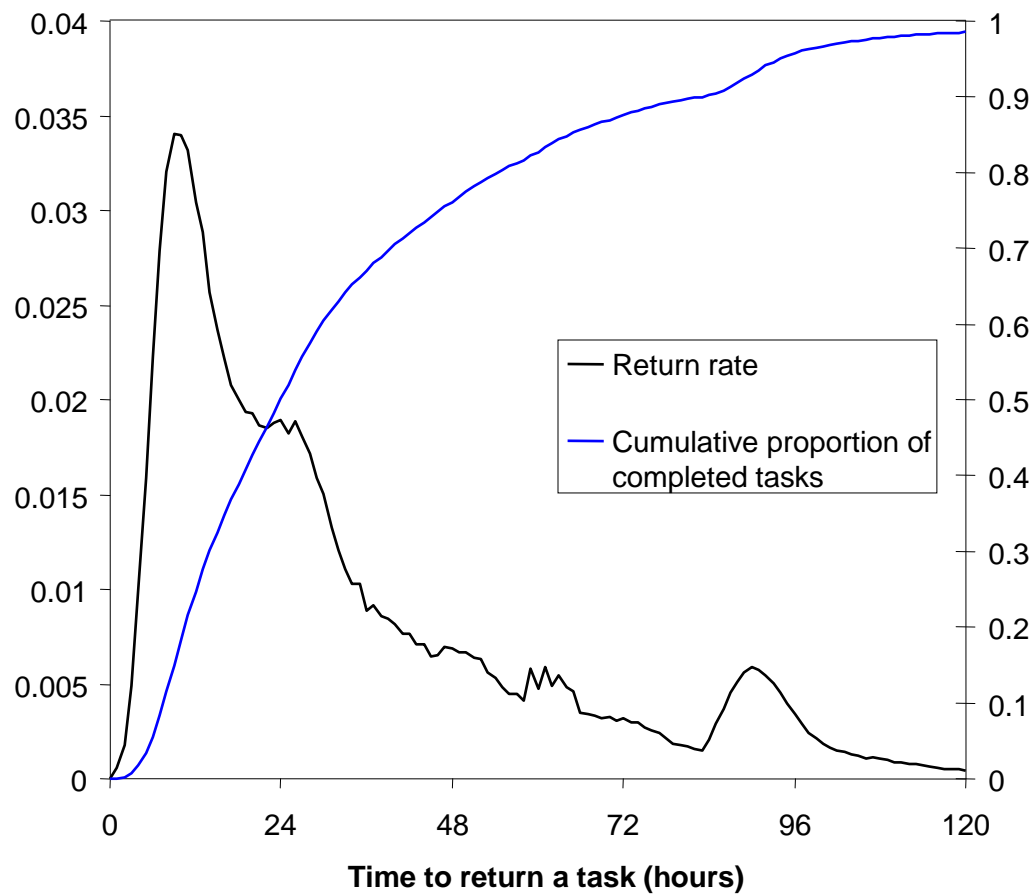
- Computationally intense
- Highly parallel
- Timing of completion of results long and unpredictable



Swiss Tropical Institute  
Institut Tropical Suisse  
Schweizerisches Tropeninstitut



# Waiting time for results





## BOINC approaches to minimize latency

- Reliable host concept
  - Track history of hosts
  - Define reliable hosts: Low average return time, low error rate
  - Assign high priority tasks to those hosts
  - Optionally reduce task deadline
- Redundancy elimination/adaptive replication
  - Replicate tasks for unknown or unreliable hosts
  - Replicate a low proportion of tasks for reliable hosts



## Outlook II: Design and analysis of predictive simulation studies

- Study design and data analysis is currently a bottleneck, and error prone
- Increasing demand for model predictions for decision-making (e.g. BMGF MalERA)
- Need for better accessibility for users from various domains

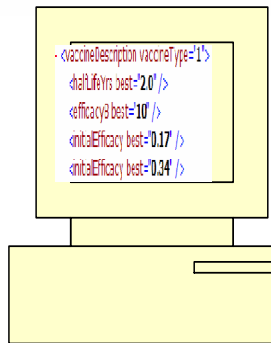


Swiss Tropical Institute  
 Institut Tropical Suisse  
 Schweizerisches Tropeninstitut



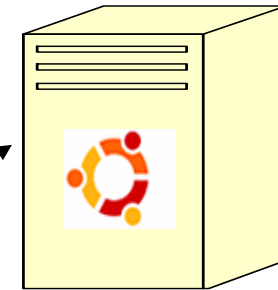
# Platform and workflows

Design of simulation study



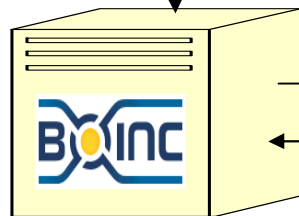
Malaria model application and run (set of workunits) are sent to BOINC Server

Data analysis

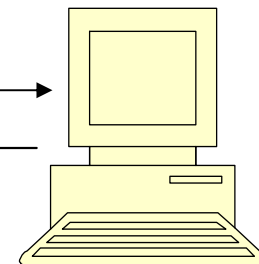


Data server receives results from BOINC server. (Storage/Analysis)

Workunit management, dynamic wu generation



BOINC sends workunits/app to volunteers and collects results





## Outlook

- Collaboration with NDSSL, Virginia Tech
  - NDSSL previous projects:
    - DIDACTIC: HPC Services Based Epidemiological Planning Environment
    - User-friendly interface capable of generating complex study designs as a front-end to HPC resources
- Implementation of a web-based job-submission system for the malaria simulation application



Swiss Tropical Institute  
Institut Tropical Suisse  
Schweizerisches Tropeninstitut



## Current Research Team

### **Applied Mathematics**

Melissa Penny  
Nakul Chitnis (MACEPA)

### **Epidem./Public Health**

Allan Schapira  
Blaise Genton  
Christian Lengeler  
Don de Savigny  
Marcel Tanner  
Rick Steketee (MACEPA)

### **Quantitative biology**

Ian Hastings (Liverpool)  
Michael Bretscher

### **Statistics**

Amanda Ross  
Penelope Vounatsou  
Tom Smith

### **Computer Science**

Diggory Hardy  
Nicolas Maire  
Tiago Antão (Liverpool)

### **Health Economics**

Fabrizio Tediosi  
Josh Yukich  
Lesong Conteh  
Valerie Crowell

### **Financial support**

Bill & Melinda Gates Foundation, PATH-MACEPA, Swiss National Science Foundation