VERIFICATION, VALIDATION AND REPLICATION

A WORD ABOUT MODELS..

Essentially, all models are wrong, but some are useful. —George Box

The word model sounds more scientific than the word fable or tale, but I think we are talking about the same thing. —Ariel Rubinstein

THE THREE TESTS

- Model verification is the process of determining whether a given model corresponds to the conceptual model
- Model validation is the process of determining the implemented model corresponds to, and explains, some phenomenon in the real world.
- Model replication is the implementation by one researcher or group of researchers of a conceptual model previously implemented by someone else.

IN OTHER WORDS

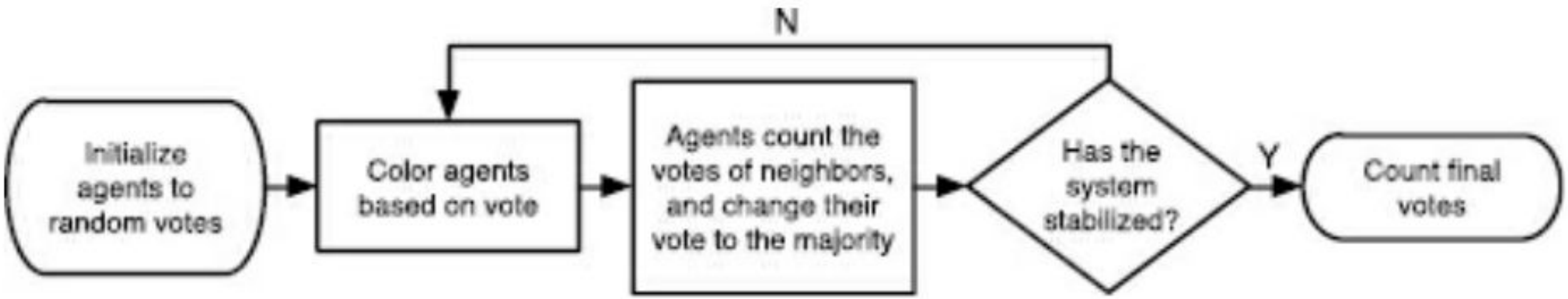
- Verification – Test if your implementation of the concepts is correct
- Validation – Test if your conceptualization itself is correct
- Replication – Peer Review

VERIFICATION

- A general guideline for enabling model verification involves building the model simply to begin with, expanding the complexity of the model only as necessary.
- Two ways of designing
  - The onion
  - The Lego bricks

DESCRIBING CONCEPTUAL MODELS

- We already saw this in ODD
- Flowchart is a good way to describe the conceptual models
- Example: Simple model of voting behavior



TESTING IMPLEMENTATION OF CONCEPTS

- Expand the flowchart into pseudocode
- Compute consequences of extrema in parameters
- Unit Testing – Or Similar
- As we grow we need to spawn unit tests for extrema pairs/tuples
  - Assume you have params A and B
  - You should test for (Amax, Bmin) (Amax, Bmax) (Amin, Bmax) (Amin,Bmin)
- The consequences should be written before testing (probably even test driven development)
- You can have in-model unit testing (next slides) or there are ways to external/offline testing (we *might* look into during advanced topics)

Example: Voting Distribution Model

Voting distribution model is a simple cellular automaton that simulates voting distribution by having each patch take a “vote” of its eight surrounding neighbors.

TESTING IMPLEMENTATION OF CONCEPTS

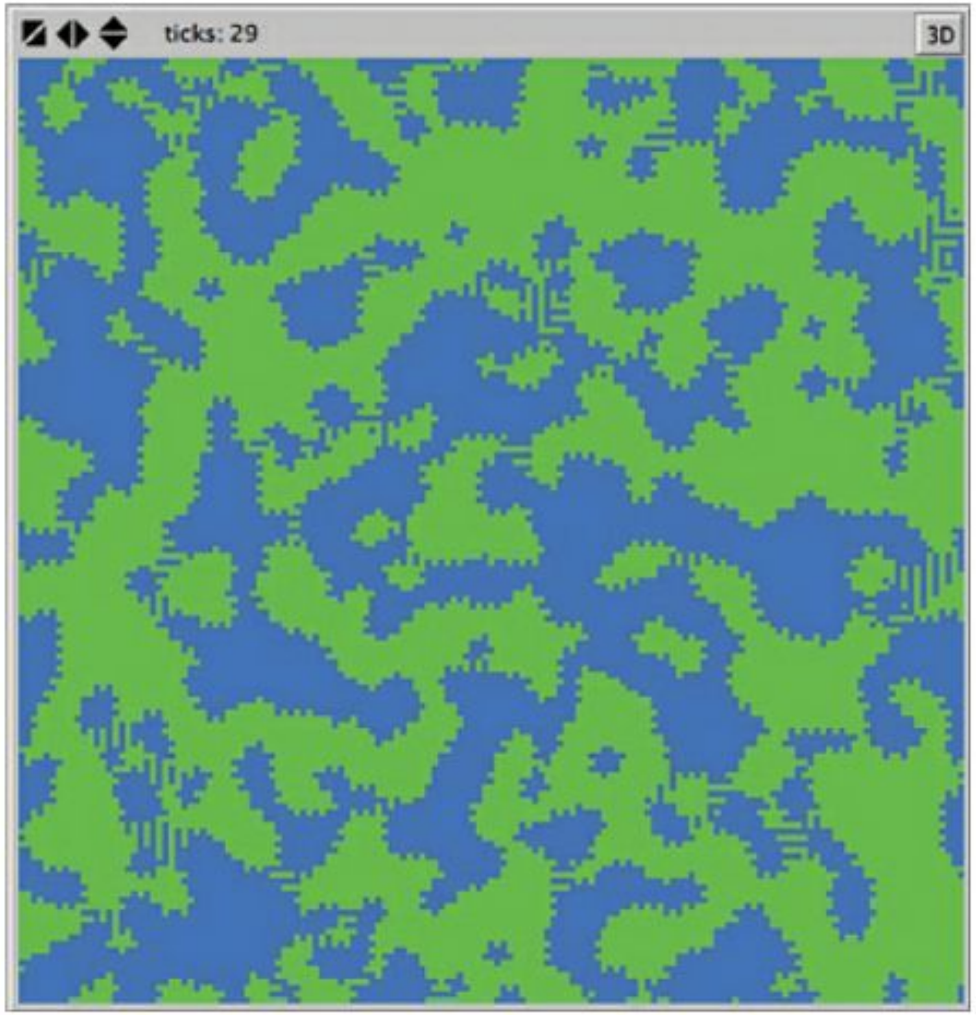
```
patches-own
[
    vote ;; my vote (0 or 1)
    total ;; sum of votes around me
]

to setup
    clear-all
    ask patches [
            if (random 2 = 0) ;; half a chance of this
            [ set vote 1 ]
        ]
    ask patches [
        if (random 2 = 0) ;; half a chance of this
            [ set vote 0 ]
    ]
    ask patches [
        recolor-patch
    ]
end

to recolor-patch ;; patch procedure
    ifelse vote = 0
        [ set pcolor green ]
        [ set pcolor blue ]
end
```

```
to check-setup
  ;; count the difference between the number of green and the number of
  ;; blue patches
  let diff abs (count patches with [ vote = 0 ]      - count patches with
    [ vote = 1 ])
  if diff > .1 * count patches [
    print "Warning: Difference in initial voters is greater than 10%."⁴
  ]
end
```

COMMON MISTAKES (EXAMPLE: JAGGED EDGES)



COMMON MISTAKES

- Ties for limit-points
  - Leverage-points of state/behavior change
- Oscillating patterns
  - Causes of local oscillations or saturation
  - (say) if you help your neighbor good-karma goes up
  - A and B are neighbors, they might go into local cycle
    - A helps/gives B and B helps/gives A
- Implement Tie-breaking logic
- Avoid oscillations using memory/history/memento variables

SENSITIVITY ANALYSIS

- Extremum gives us a quick test
- But, if the model is not computationally too expensive
  - It pays off doing sensitivity analysis
  - Even if the model takes more time, you can adjust the intervals to do a broad sweep
- Sensitivity analysis is an examination of the impact of varying model parameters on model results.
- To determine how sensitive a model is
  - we examine the effect that different initial conditions and agent mechanisms have on the model results.

VALIDATION

- Validation is the process of ensuring that there is a correspondence between the implemented model and reality.
- Validation, by its nature, is complex, multilevel, and relative. Models are **simplifications** of reality
- **Remember** that it is impossible for a model to exhibit all of the same characteristics and patterns that exist in reality

AXES OF VALIDATION (RAND & RUST 2011)

- The level or abstraction of the validation process
  - **Microvalidation** vs **Macrovalidation**
- The level of detail of the validation process
  - (expert) **Face validation** vs **Empirical Validation**

CONSIDERED EXAMPLE: FLOCKING MODEL

- Attempts to recreate patterns of birds flocking behavior
- Its based on Boids model by Reynolds (1987)
- Demonstrated that flocks can arise without necessity of a leader
- Lets see the model
- Just three rules: alignment, cohesion and separation

MICROVALIDATION

- Compare Individual agent (bird) behavior to known behavior observed by experts
- This might explain the sufficiency of modeling the individual
  - Do we need to add additional property (like vision range) ?
  - Or may be one property does not add much value shall we eliminate it
- There are several ways of doing this
  - Let's switch to the other axes and come back….

FACE VALIDITY (EXPERTS)

- On its face does the model (animation) looks correct
- Does the way the flock coalesces somewhat look reasonable?
- Tracing – Trace everything a select subset of individuals see and do
  - Think of it like inspecting a specific turtle using go-once
  - Does the action of the agent is close to real world observations

EMPRICAL VALIDITY (DATA)

- Even with stochasticity the effect of parameters on certain emergent phenomenon should be within an acceptable range
  - Should change in meaningful predictable ways (predictive validation)
  - If unable to see the pattern, you should be able to investigate and explain it
- Internal validity: Run experiments multiple times and subject important numbers to statistical analysis
- Data driven validation: See if the results qualitatively match the real world observed data (Only when Historical data exists)

MACROVALIDATION

- Same idea of doing face validation and data driven validation but on a different scale
- Instead of individual, we analyze on a much higher level
- How does the flock (groups) behave
  - When they come together
  - When they break apart
- Internal validation across runs (we anyway do that as part of testing our hypothesis)
- Think of Micro-Macro validation not as binary but as a spectrum

CASE STUDY - ARTIFICIAL ANASAZI MODEL

- A good model highly appreciated for the level of validation
- Native Indians – Kayenta Anasazi lived in Arizona, US
- The model explains their disappearance from the region based on social and environmental changes
- We will see this tomorrow!

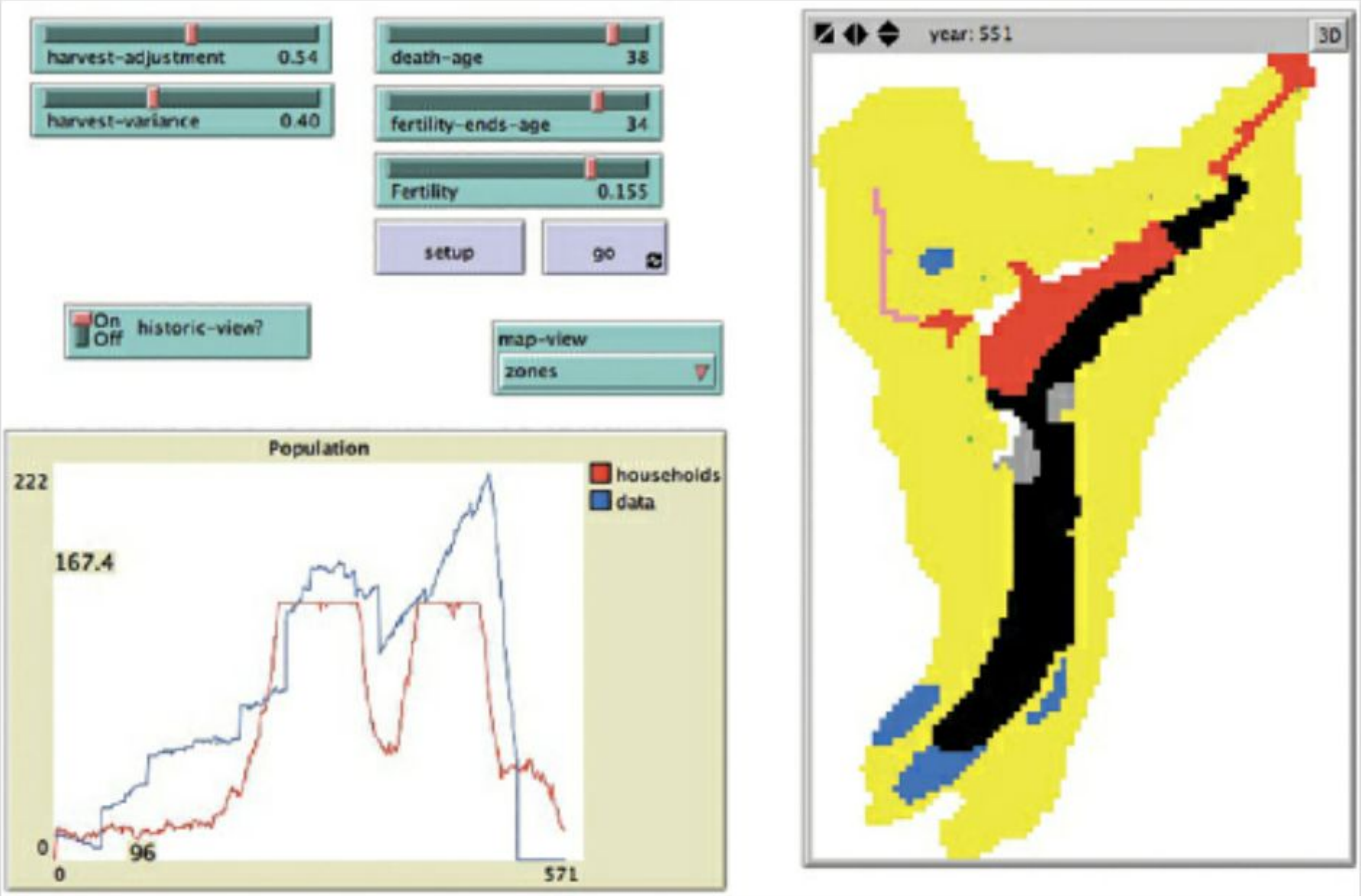


Figure 7.10 NetLogo version of the Artificial Anasazi model. In the population plot, the red line plots the simulated population and the blue line the real population data.

REPLICATION

- Across hardware
- Across languages
- Across platforms