Introduction to agent-based modeling (ABM), the Chitwan ABM, and PyABM

Alex Zvoleff
Postdoctoral Associate
Conservation International

September 27, 2013

San Diego State University

Overview

- Introduction to ABM
- Example: the Chitwan ABM
- Getting started in ABM
  - Recommendations
  - Downloading Chitwan ABM and PyABM
- STACS server data

Introduction to agent-based modeling

What is an agent-based model?

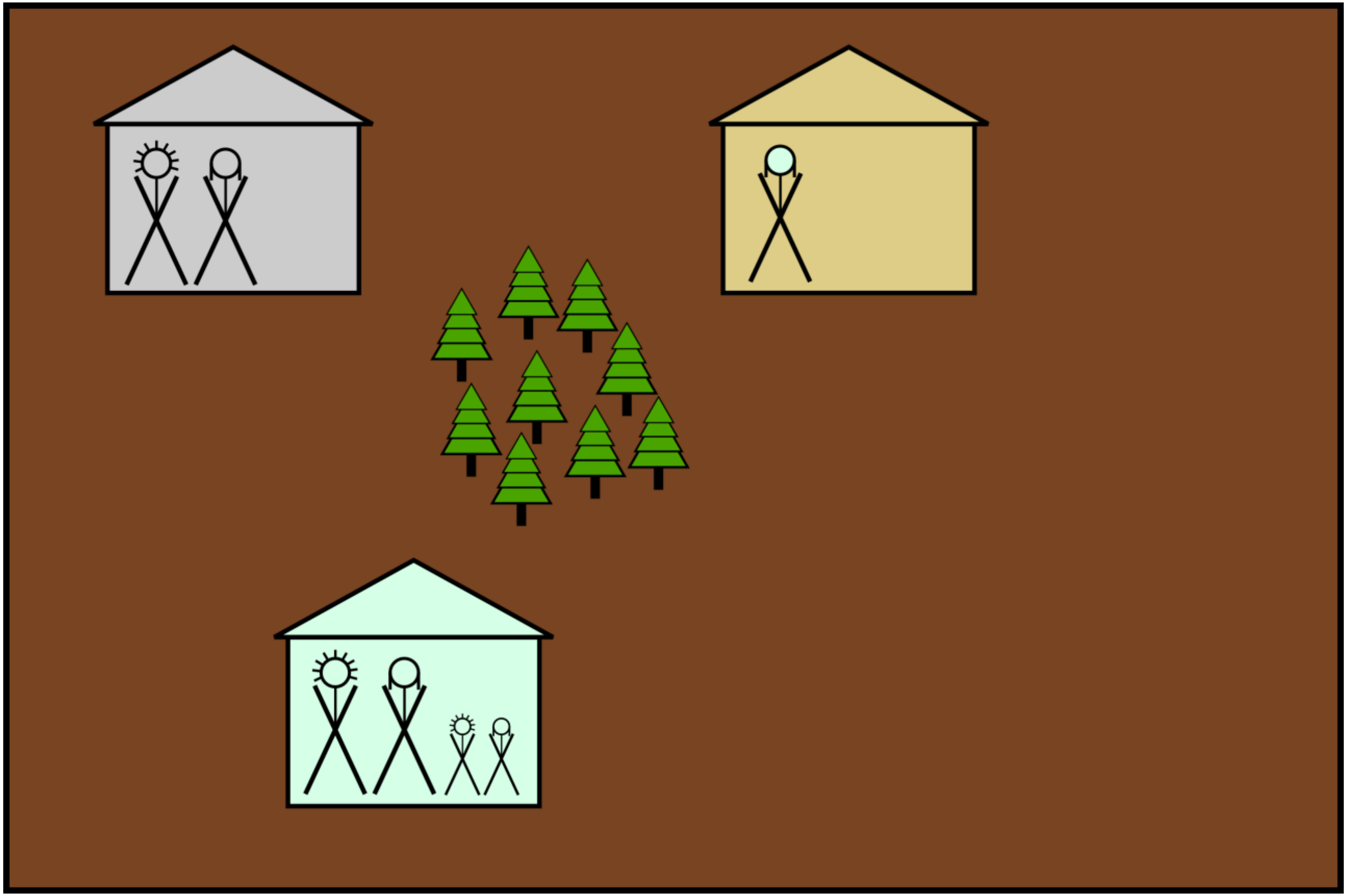
- Agent-based models (ABM) represent individual "agents" and model their interactions
- ABMs model processes from the bottom-up, allowing researchers to study "emergent" phenomena that result from a multitude of lower-level interactions
  – Land use change as a result of individual developers or settlers
  – Habitat change as a result of fuelwood collection
  – Spread of disease
  – Stock market movements

Why use agent-based models in geography?

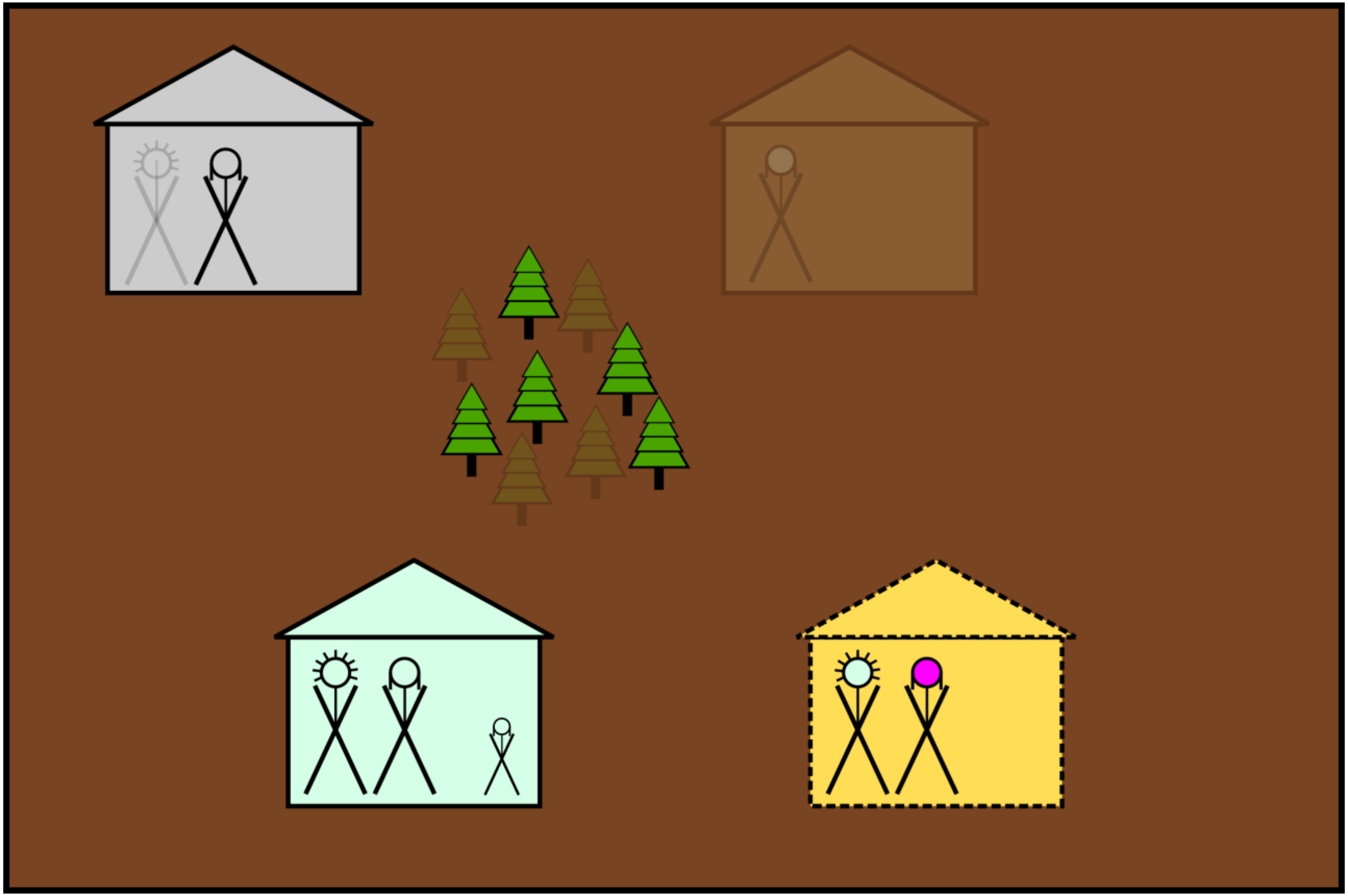
- ABM does a good job
  - Representing alternative models of human-decision making
  - Considering feedbacks
  - Enabling study of system dynamics

- ABM can be used as a “virtual laboratory” where alternative hypotheses about system structure can be tested
  - What if in-migration doubled?
  - If schooling reduces the age of marriage, what happens to land use change?

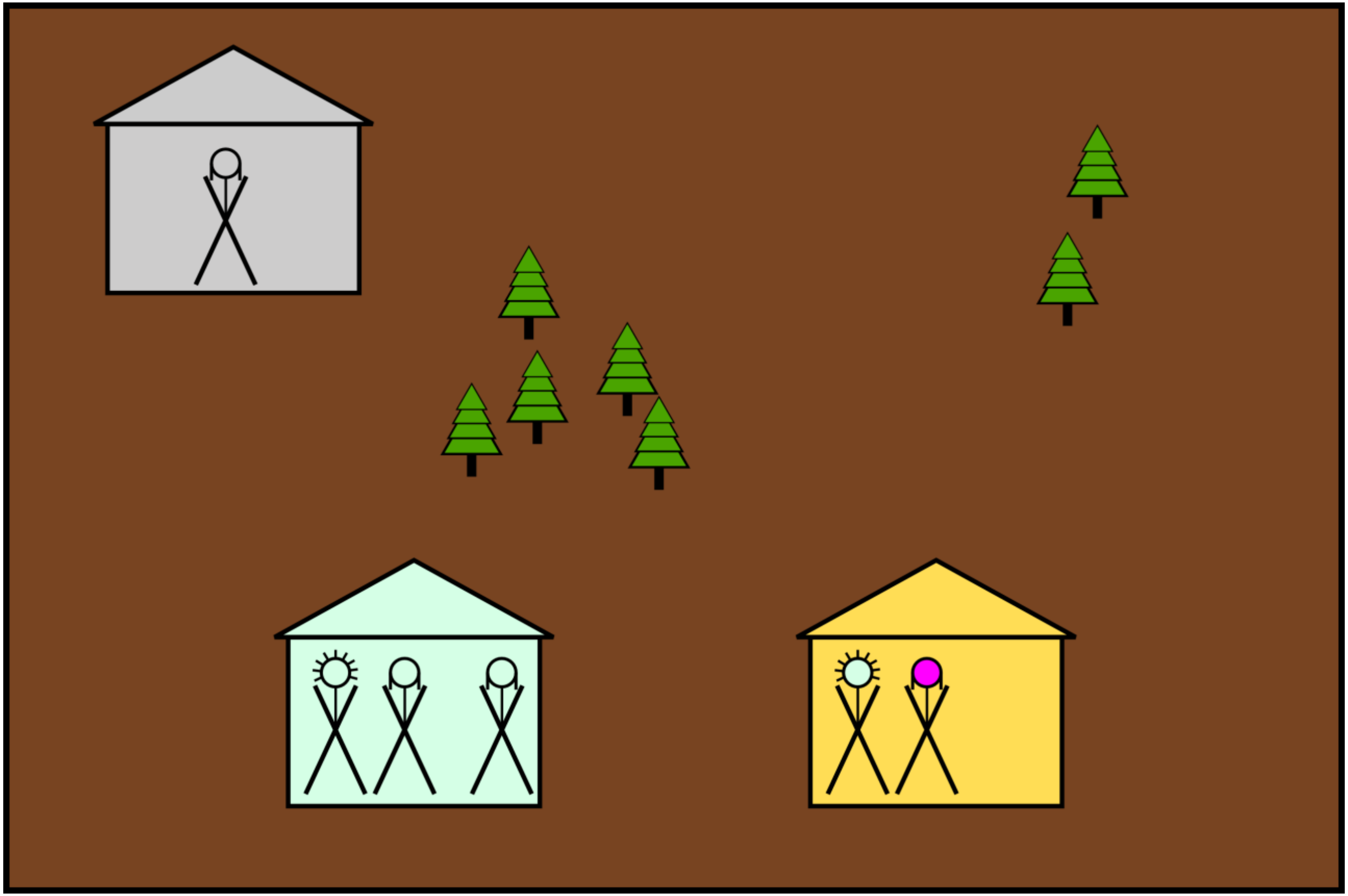
Simplified agent-based model: time 1



Simplified agent-based model: time 1.5

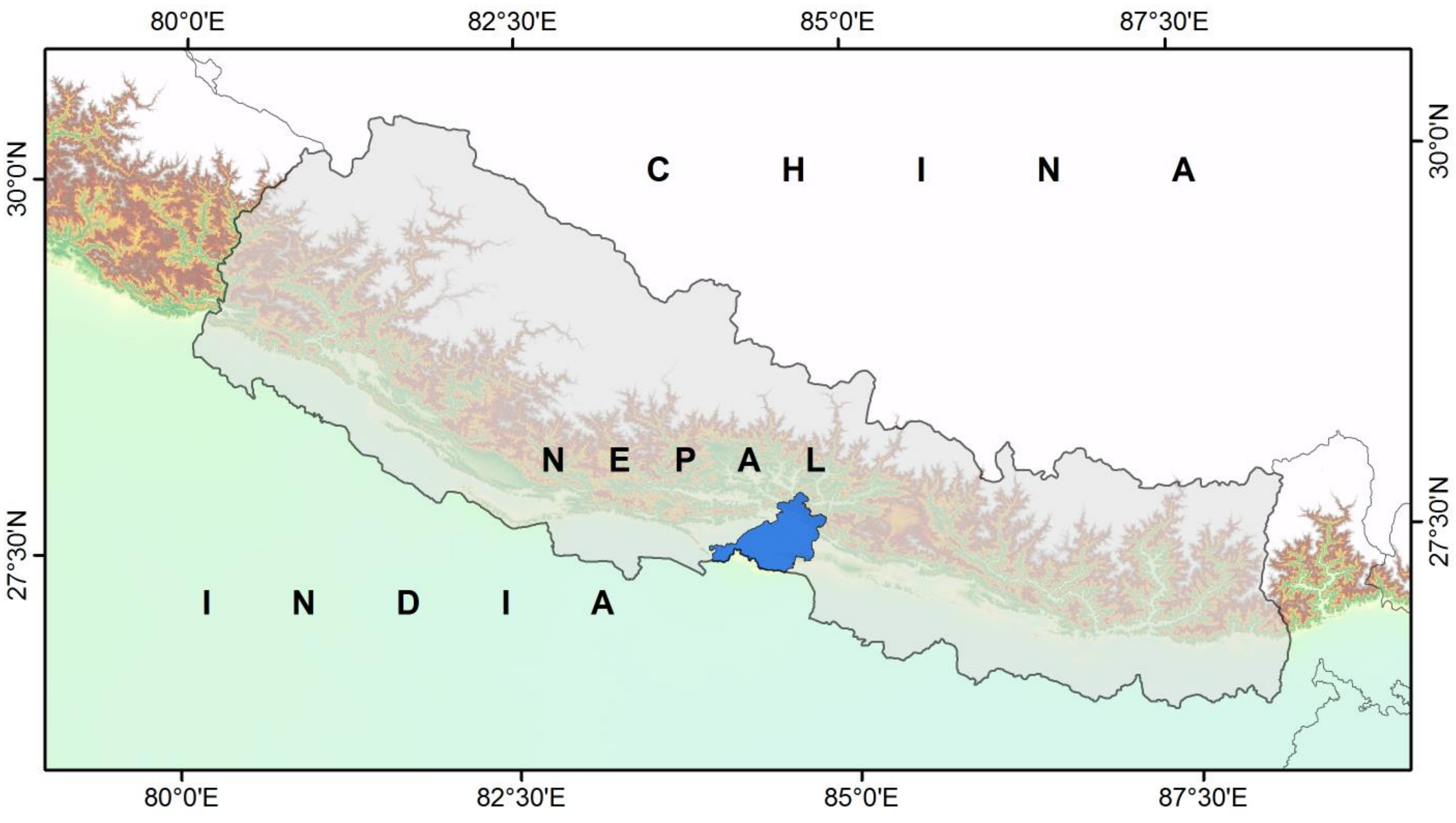


Simplified agent-based model: time 2

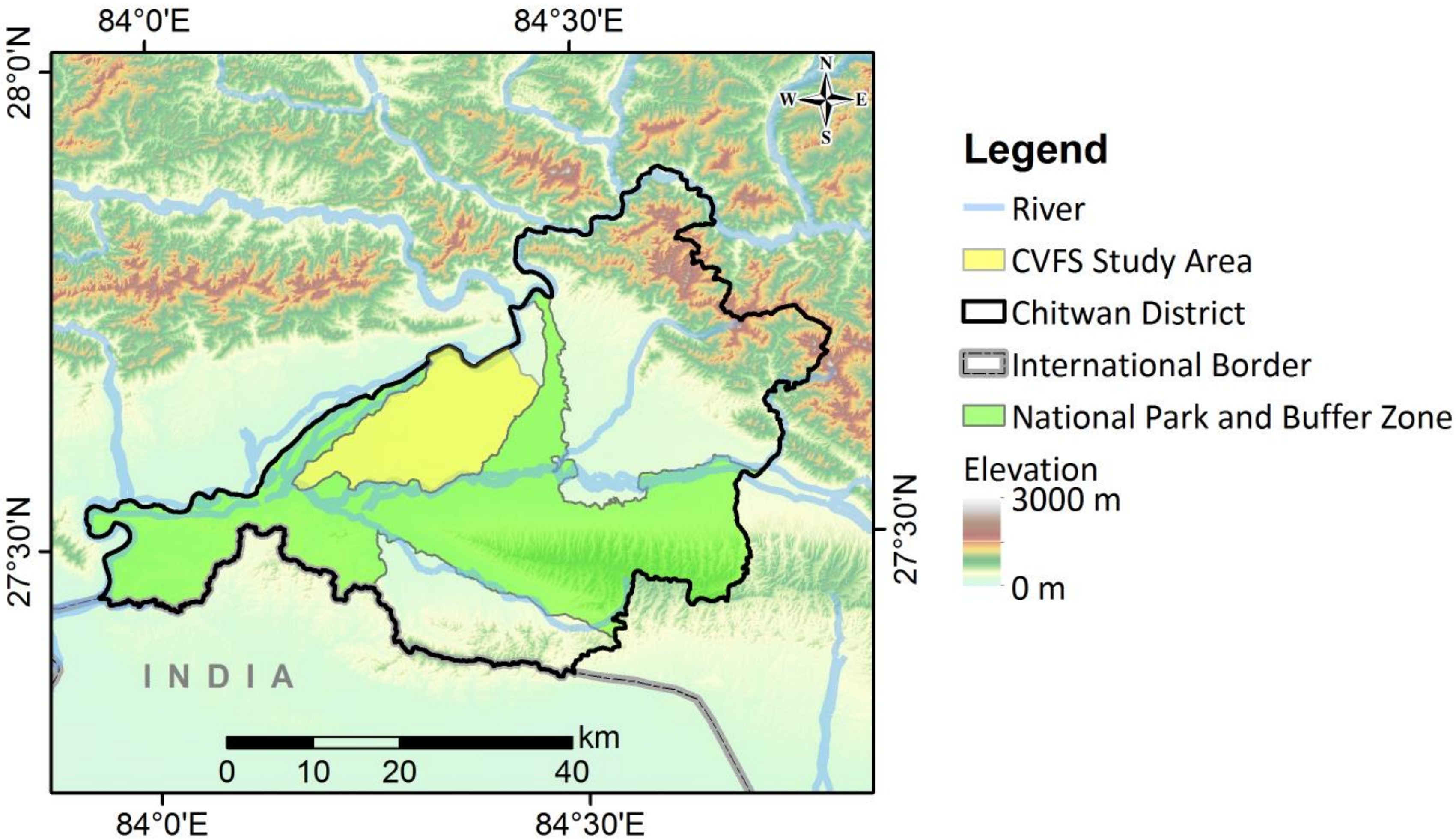


An example: the Chitwan ABM

Study site: Western Chitwan Valley, Nepal



Study site: Western Chitwan Valley, Nepal



Study site overview

- Population
  - 284,939 in 2011
  - 3.1% annual growth rate from 2001-2011
- Agriculture is dominant in local livelihoods
  - 83% of households are in agriculture
  - 80% of land is in agriculture
  - Primarily farm rice and maize

Research questions

1. How are feedbacks between population and land use manifested at the individual-level (focusing on marriage timing and fertility)?

2. What is the dynamic effect of feedbacks between land-use change and micro-level human decision-making?

Social survey datasets

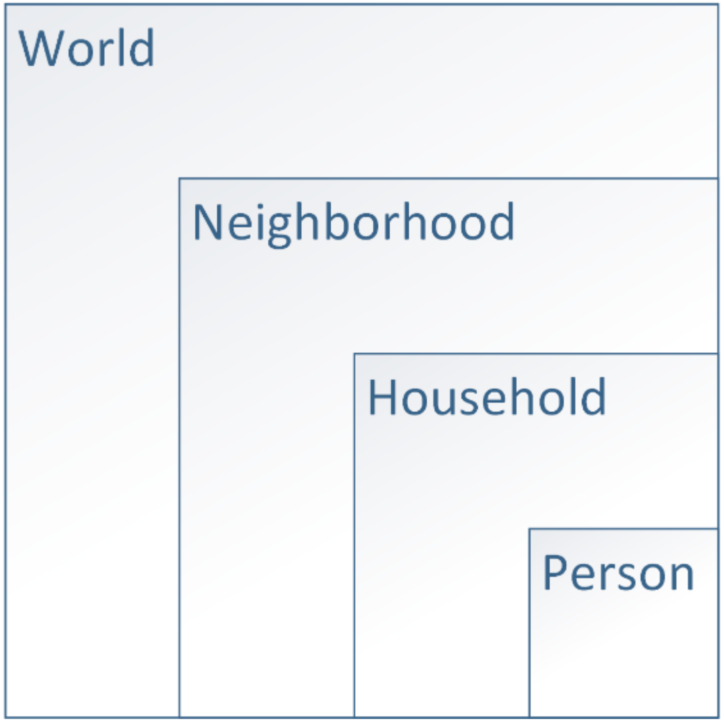
- Chitwan Valley Family Study: 1996-present
  - 1773 households in 1996, 2361 in 2006
  - Monthly demographic event registry
  - 1996, 2001, 2006 household surveys
  - 1996, 2001, 2006 neighborhood history calendars
- Household surveys (author)
  - Fuelwood usage survey: 2009, 2 part survey of 80 households
  - Climate vulnerability and adaptation survey: 2011, 294 households

Overview of the Chitwan ABM

151 neighborhood agents

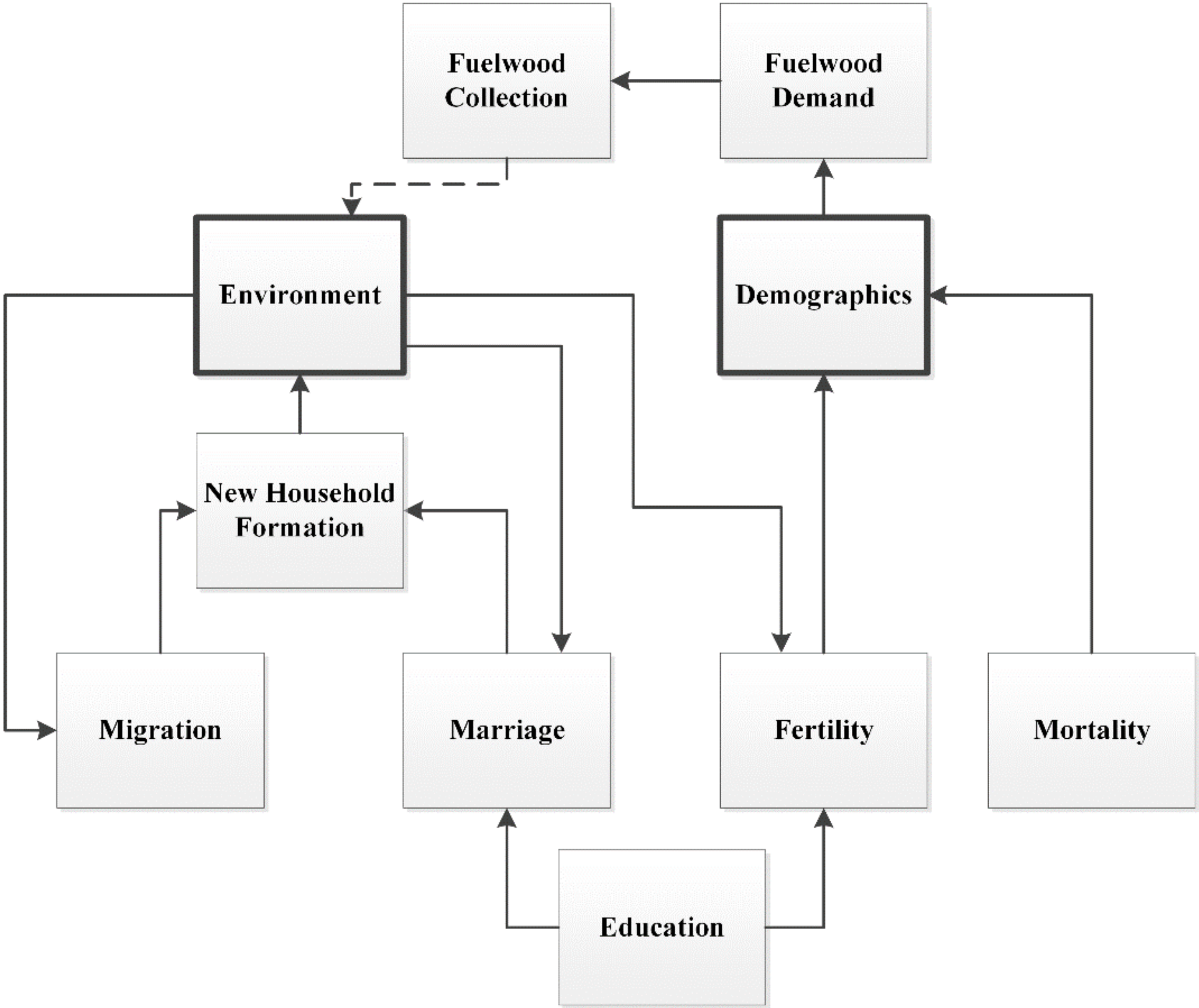
1522 household agents

8245 individual agents

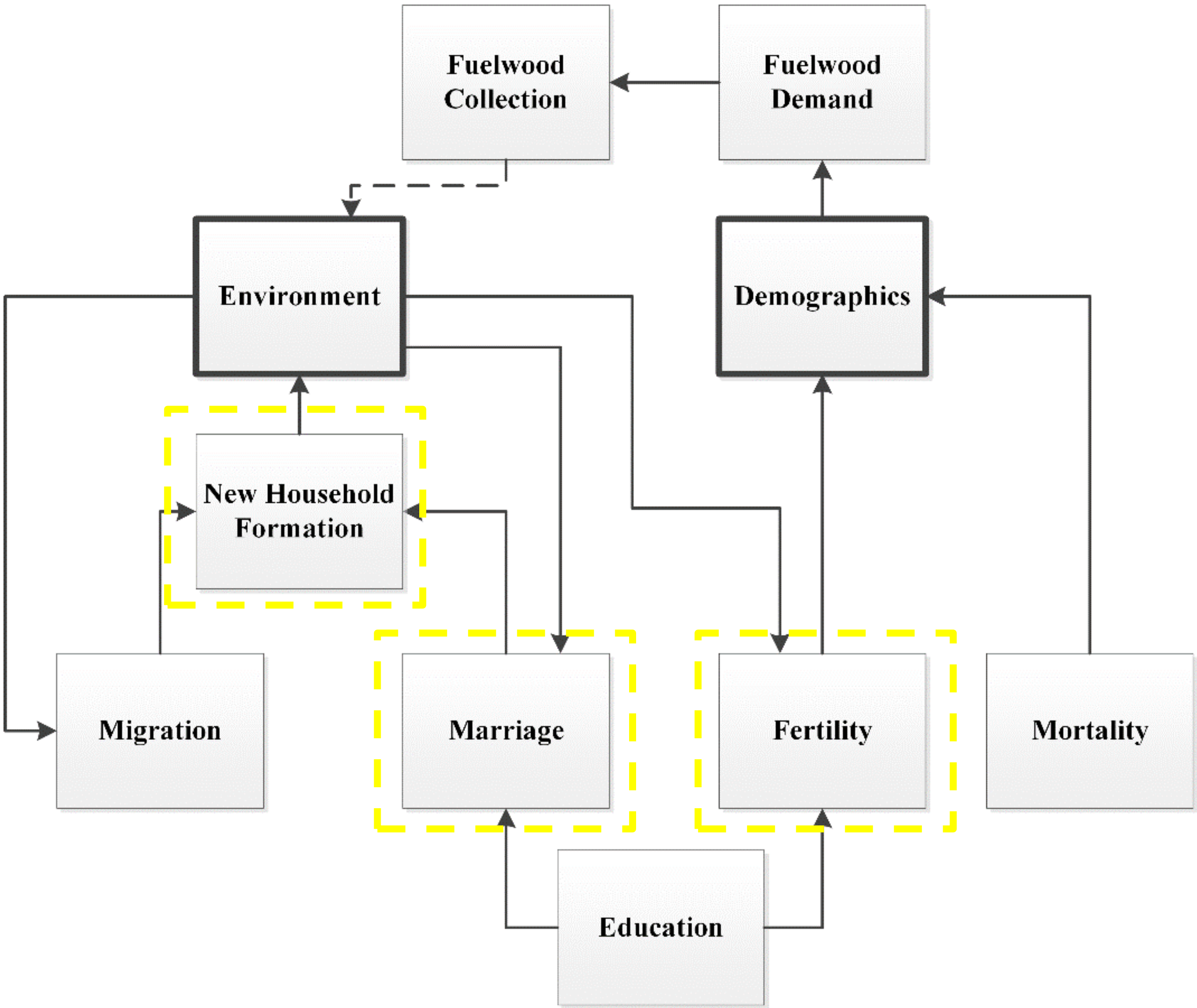


- Model is empirically parameterized
- Model land use at neighborhood level
- One-month timestep
- Runs from 1997-2050

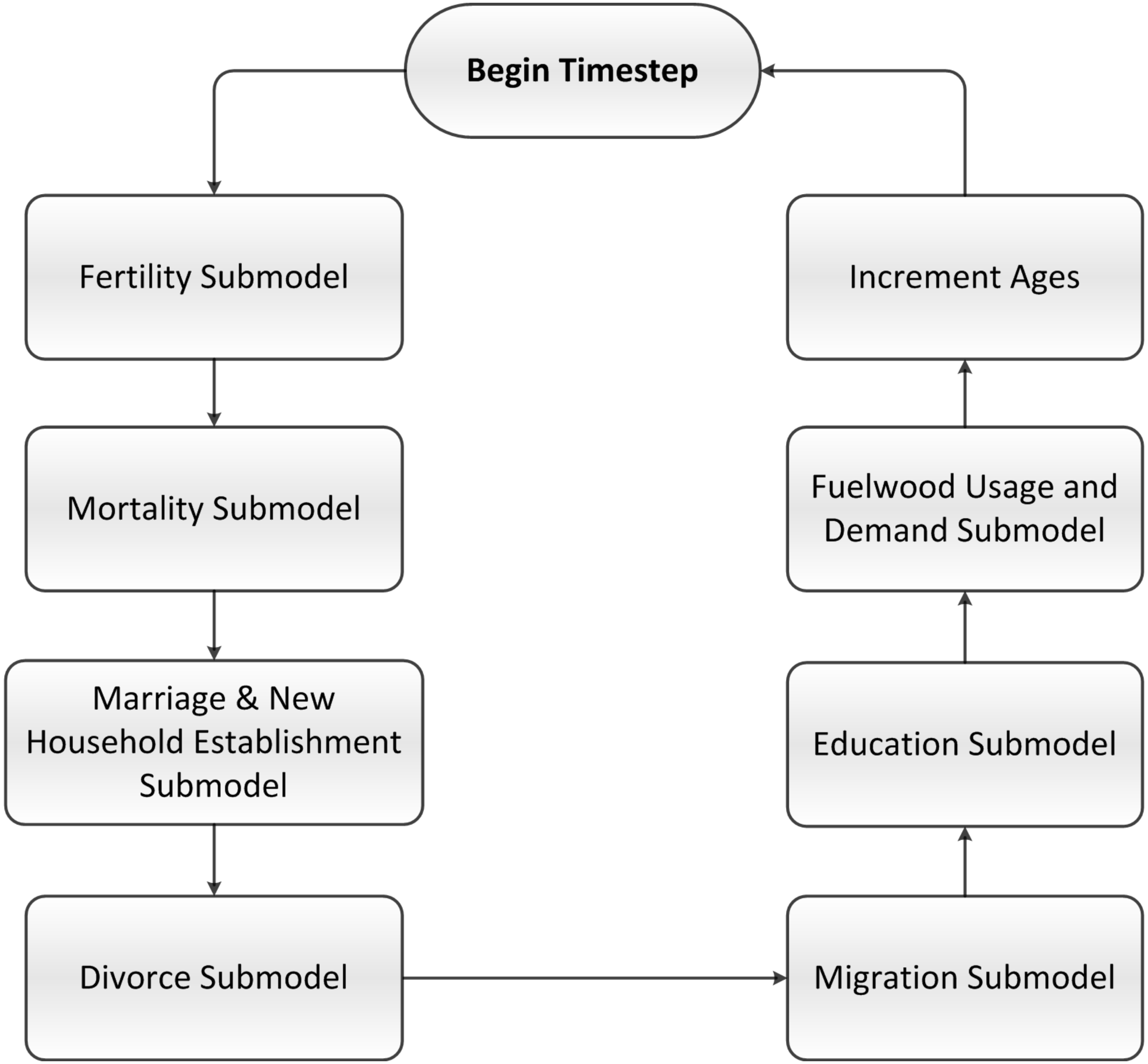
Chitwan ABM structure



Chitwan ABM structure



Chitwan ABM scheduling



Fertility submodel

- First births after marriage are modeled separately from other births
- In each month following marriage, calculate probability of a live birth in that month based on regression results (based on Ghimire and Hoelter 2007 and Axinn and Ghimire 2010)
- Second and subsequent birth intervals follow a probability distribution
- Desired number of children based on empirical probability distribution

Fertility submodel: first-birth timing

Variable	Odds Ratio	2-sided p-value
Percent land area under agriculture	1.002	0.458
Community characteristics		
Neighborhood area	0.420	0.428
Distance to urban center	1.007	0.548
Electricity in 1996	1.298	0.105
Non-family services within a 15-min walk	0.996	0.619
Controls		
Respondent's parents' characteristics		
Mother's number of children	0.977	0.504
Mother's education	0.922	0.678
Mother's work	0.930	0.686
Father's education	0.874	0.324
Father's work	0.721	0.021 *
Parents' contraceptive use	0.968	0.824
Respondent's ethnicity		
Low Caste Hindu	1.026	0.922
Hill Tibeto-burmese	0.539	0.003 **
Newar	0.619	0.210
Terai Tibeto-burmese	0.958	0.816

Continued on next slide. Model based on Axinn and Ghimire, 2010.

Fertility submodel: first-birth timing

Variable	Odds Ratio	2-sided p-value
Respondent's marital experiences		
Age at first marriage	0.968	0.201
Marriage duration		
Marriage duration before 1997	0.994	0.452
Marriage duration during obs. period		
Married for 1–6 months	11.300	<.001 ***
Married for 7–12 months	7.118	<.001 ***
Married for 13–18 months	5.386	<.001 ***
Married for 19–24 months	3.156	<.001 ***
Married for 25–30 months	1.523	0.236
Married for 31–36 months	2.278	0.010 *
Married for 37–42 months	1.300	0.476
Schooling		
4–7 years of schooling	1.621	0.043 *
8–11 years of schooling	2.345	0.063 .
12 or more years of schooling	3.688	0.019 *
Intercept	0.028	<.001 ***

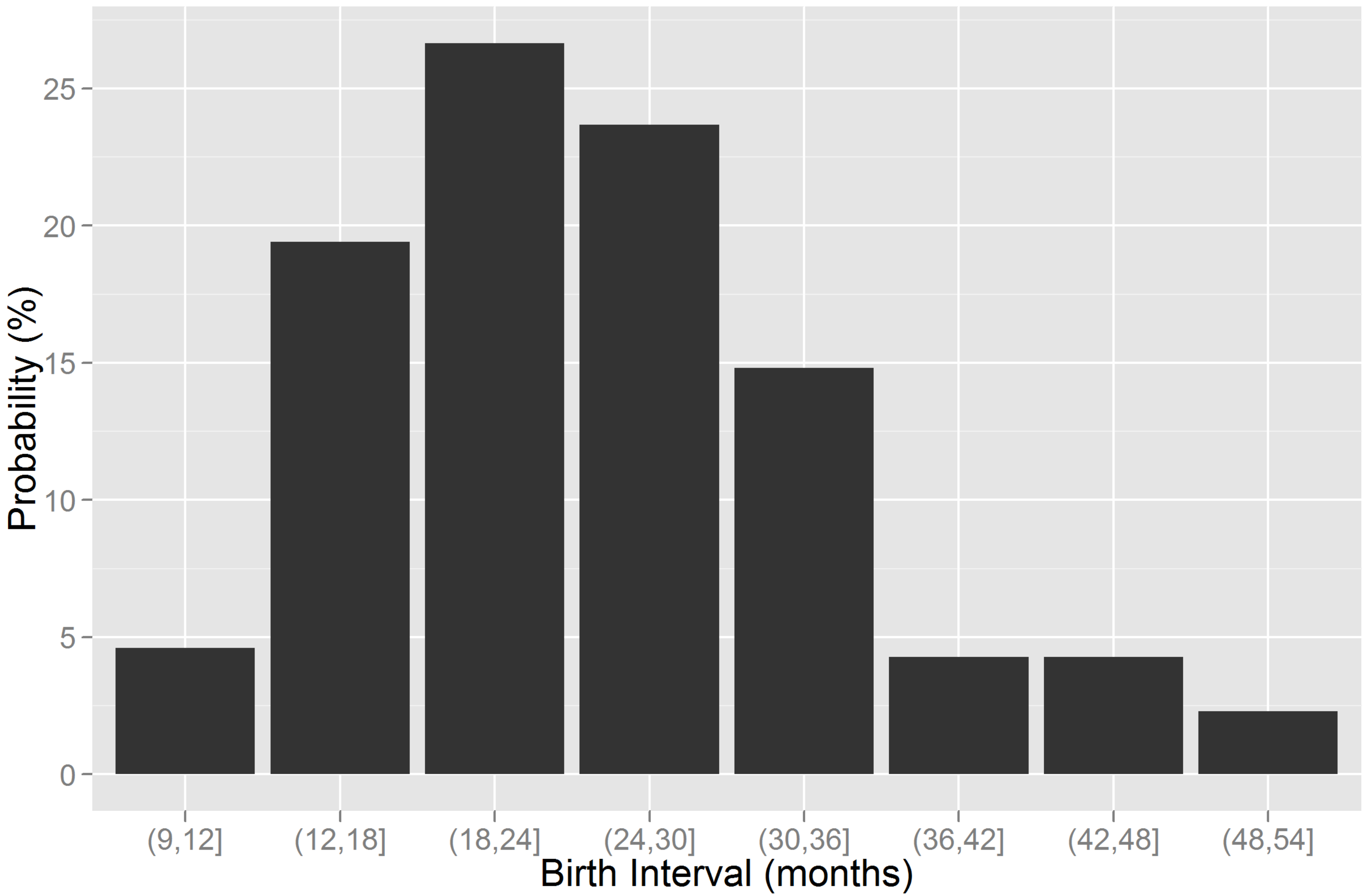
Continued from previous slide. Model based on Axinn and Ghimire, 2010.

Fertility submodel: second and subsequent births

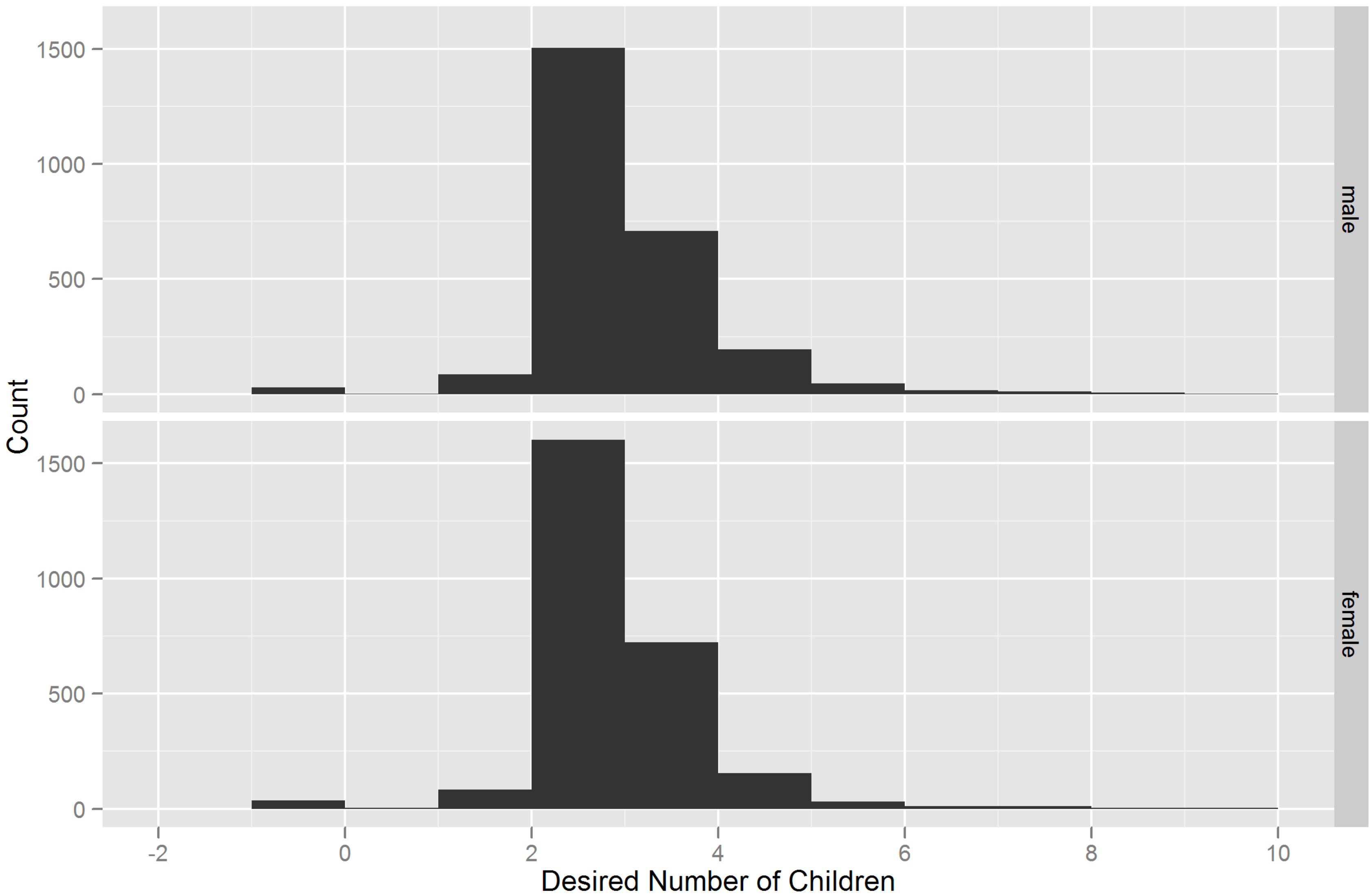
- For all other births, choose the interval until the next child from empirical probability distribution

- Births can occur until:
  - The desired number of children is reached
  - (or) woman reaches *maximum birth age* (45)
  - (or) woman dies/out-migrates

Fertility submodel: birth interval



Fertility submodel: desired number of children



Marriage submodel

1. Minimum marriage age is 15
2. Beginning at minimum marriage age, calculate a probability (**p**) of marriage for each person for that month
3. Draw a random number – if the random number is less than the probability **p**, add person to ‘to be married’ list

Marriage submodel: new household establishment

1. Once a couple is formed, decide if couple will form a new household, by drawing a random number and comparing to the *household fission rate*

2. Find a neighborhood with sufficient free land (agricultural or non-agricultural vegetation, in that order) to build new household

3. Assign new household to chosen neighborhood

Marriage model in Chitwan ABM (based on Yabiku, 2006)

Variable	Odds Ratio	2-sided p-value
Log(Percent Agricultural Land (interpolated))	1.138	0.064 .
Nonfamily organizations (minutes by foot)		
School	1.012	0.138
Health post	0.999	0.711
Bus stop	1.005	0.282
Market	0.999	0.842
Employer	1.003	0.305
Schooling		
Years schooling completed	0.997	0.893
Enrolled in school	0.669	<.001 ***
Female	2.245	<.001 ***
Ethnicity		
Lower Caste Hindu	1.014	0.942
Newar	0.786	0.229
Hill Tibetoburmese	1.187	0.256
Terai Tibetoburmese	0.906	0.508
Age		
Age	2.107	0.004 **
Age-squared	0.986	0.018 *
Intercept	0.000	<.001 ***

Spouse choice

1. Once 'to be married' list is constructed, loop over list in random order
2. Each time through list, choose a person (**psn1**)
3. Calculate probability that **psn1** would marry each other person in the list
   – Assign zero probability to anyone from a different ethnic group or from the same sex
   – Calculate probability for all others based on age difference between spouses (using empirical data)
4. Choose a spouse (**psn2**) by sampling from the potential spouses with each potential spouse weighted by their calculated probability of marriage to **psn1**
5. Unpaired spouses marry in-migrants – with spouse age chosen using spouse age model

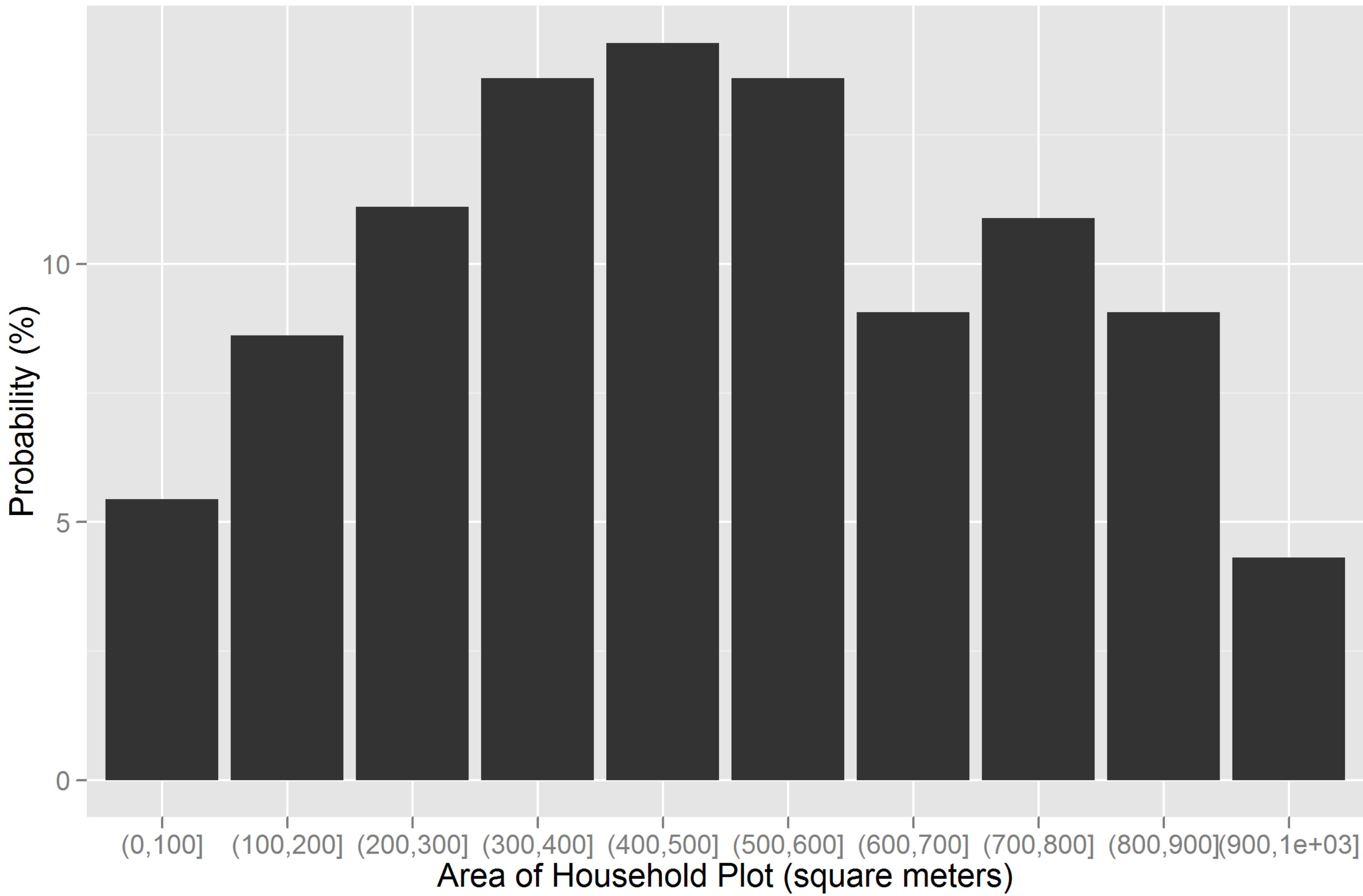
New household establishment (part 1)

1. Once a couple is formed, decide if couple will form a new household, by drawing a random number and comparing to the *household fission rate*

   – If a new household **is not** formed, move the new couple into the husband's household

   – If a new household **is** formed, draw the size (area in sq. m) of the household plot from empirical probability distribution

New household establishment (part 1)



New household establishment (part 2)

2. Find a neighborhood with sufficient free land (agricultural or non-agricultural vegetation, in that order) to build new household
   - First try husband's parent's neighborhood
   - Move outwards by distance from parent's neighborhood until first neighborhood with free land is found
3. Assign new household to chosen neighborhood

Fertility and marriage timing feedbacks

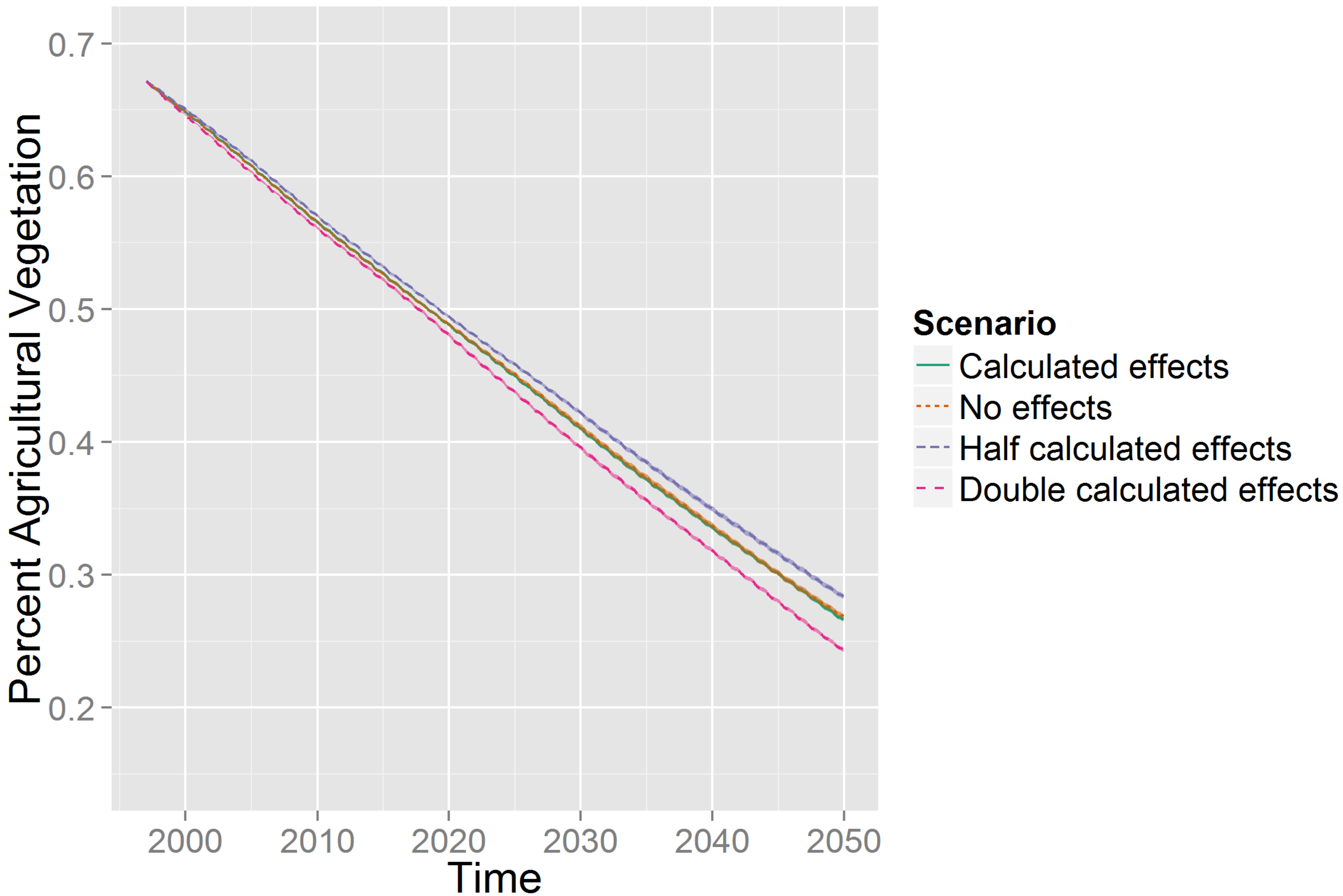
- Decrease in agricultural land leads to:
  - Increase in marriage age[1]
  - Increase in time to first birth after marriage[2]

- Act as negative feedback:
  - New household construction (on agricultural land) occurs after marriage

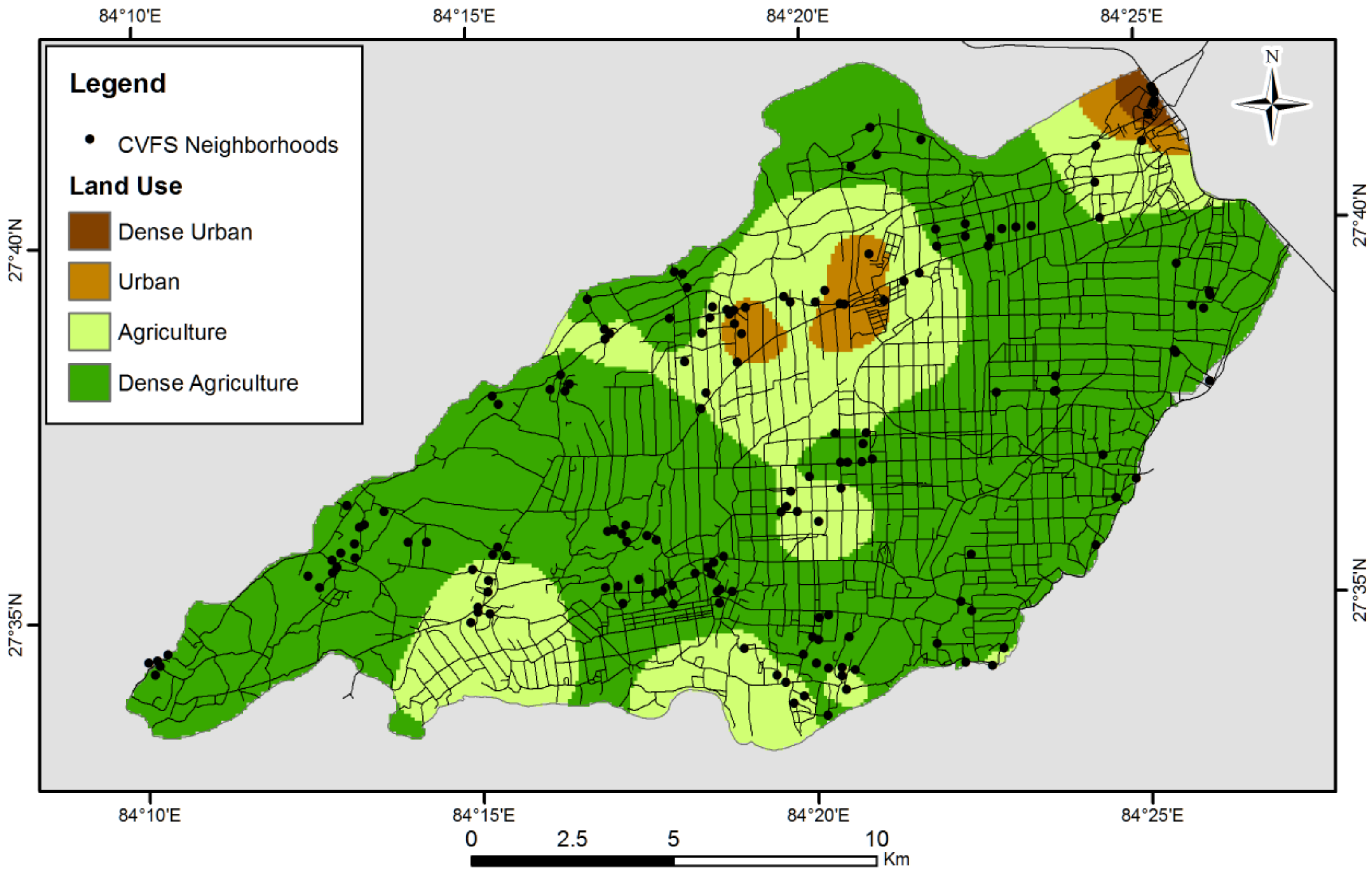
- Consider 4 scenarios (scenario set C)
  - Baseline (no feedbacks)
  - Calculated effect
  - Half effect
  - Double effect

[1] Axinn and Ghimire, 2010
[2] Yabiku, 2006

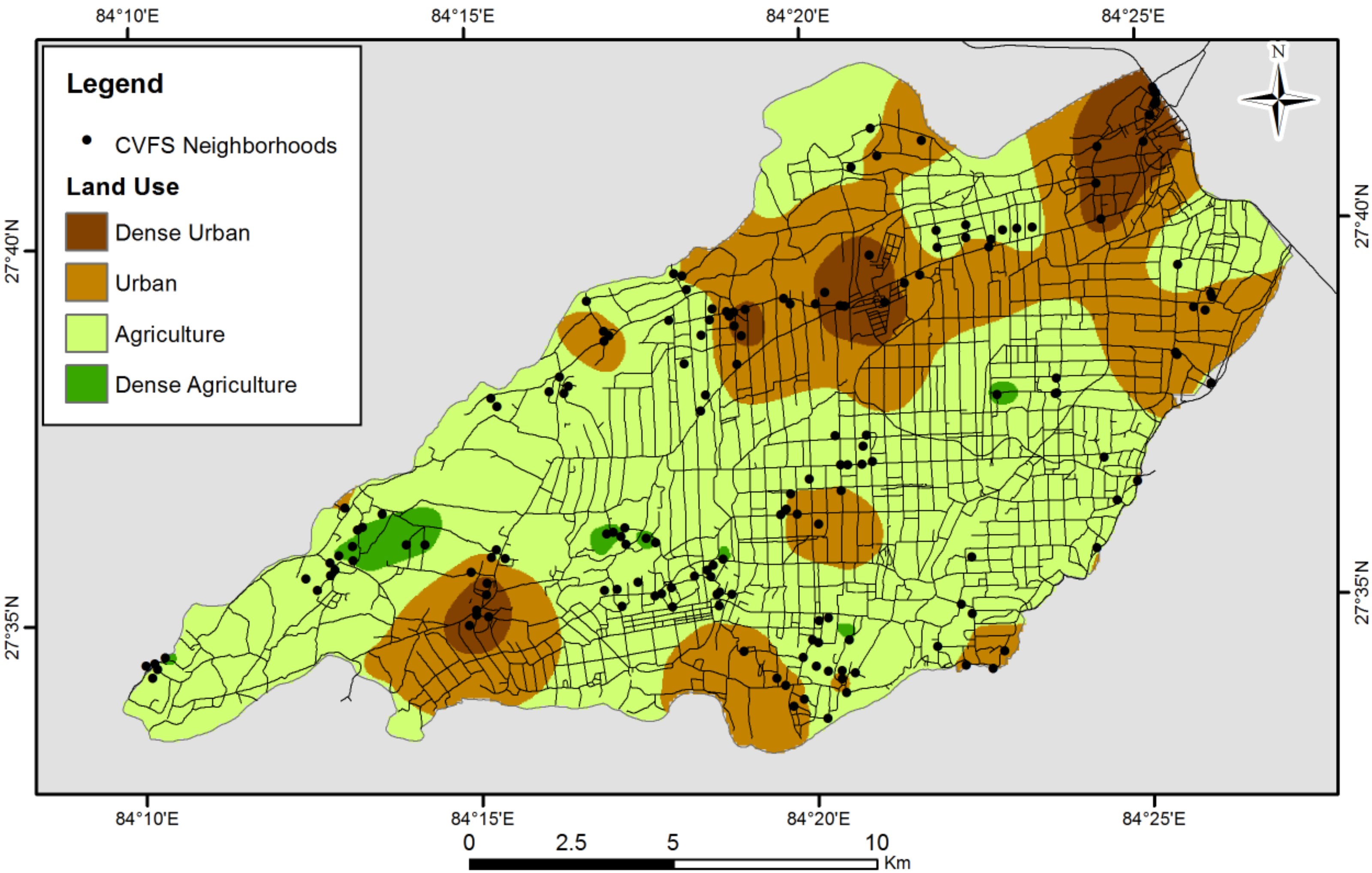
Land use change



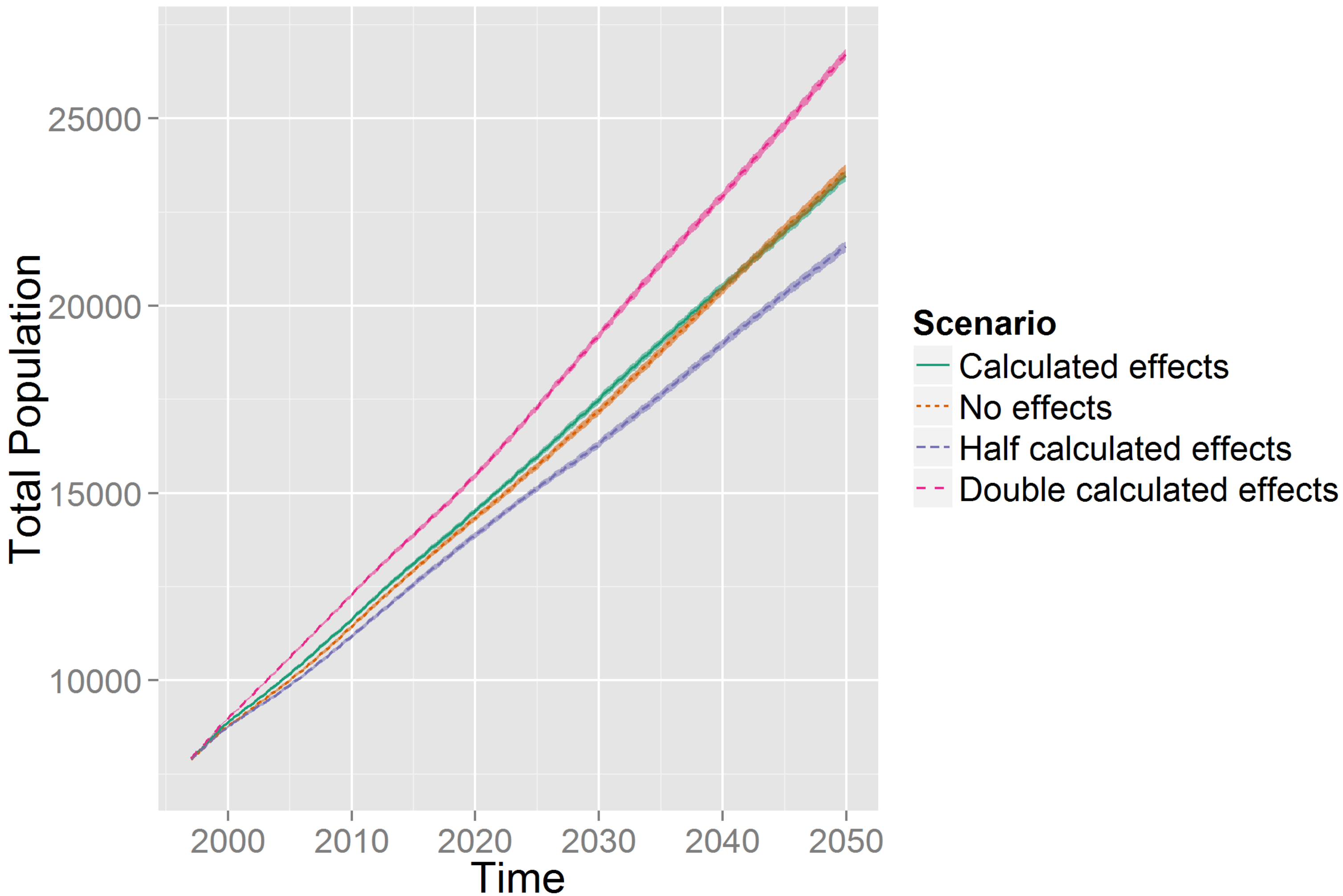
Agricultural land cover: observed data from 1997



Agricultural land cover: model projection for 2020



Population change

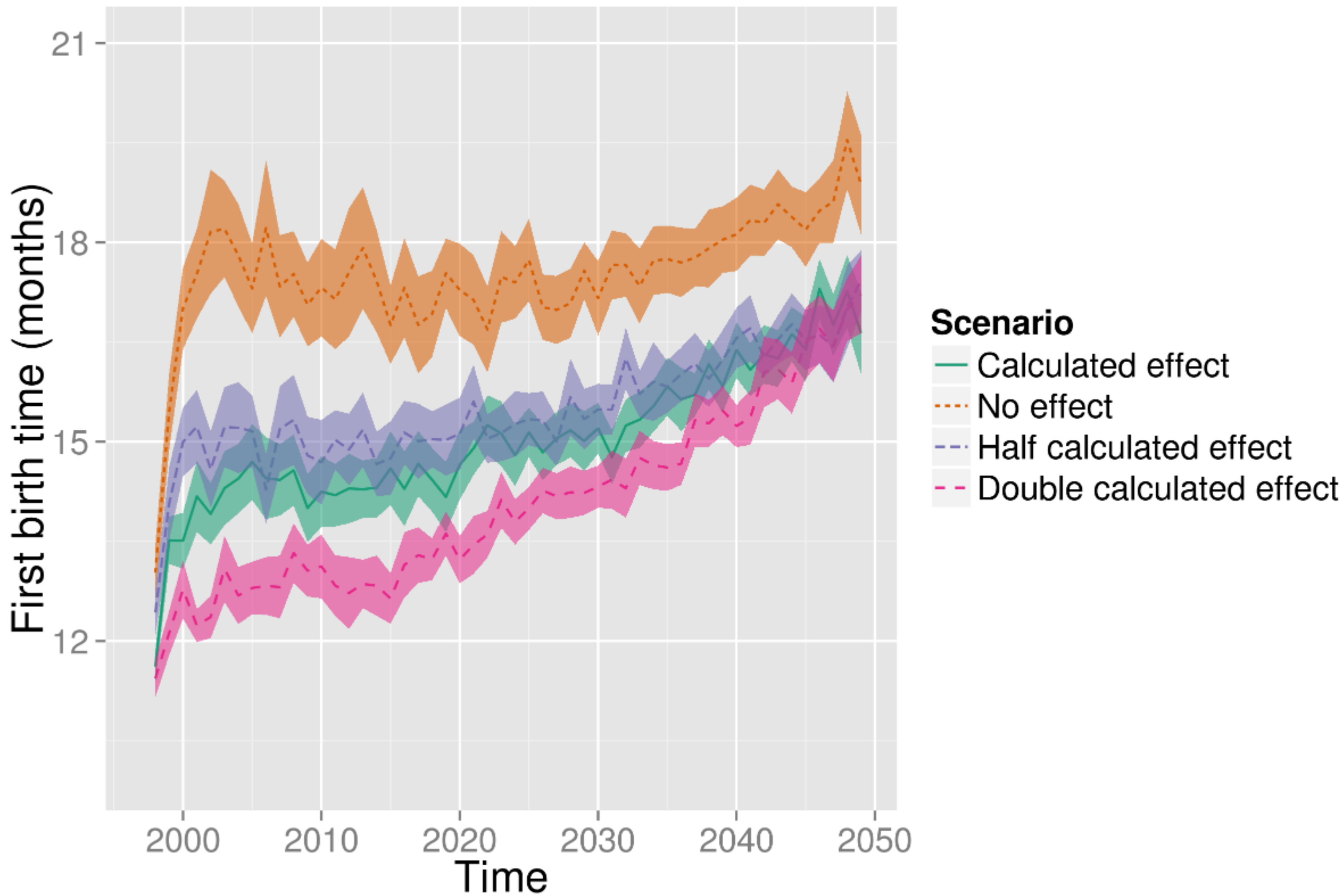


Results (2050)

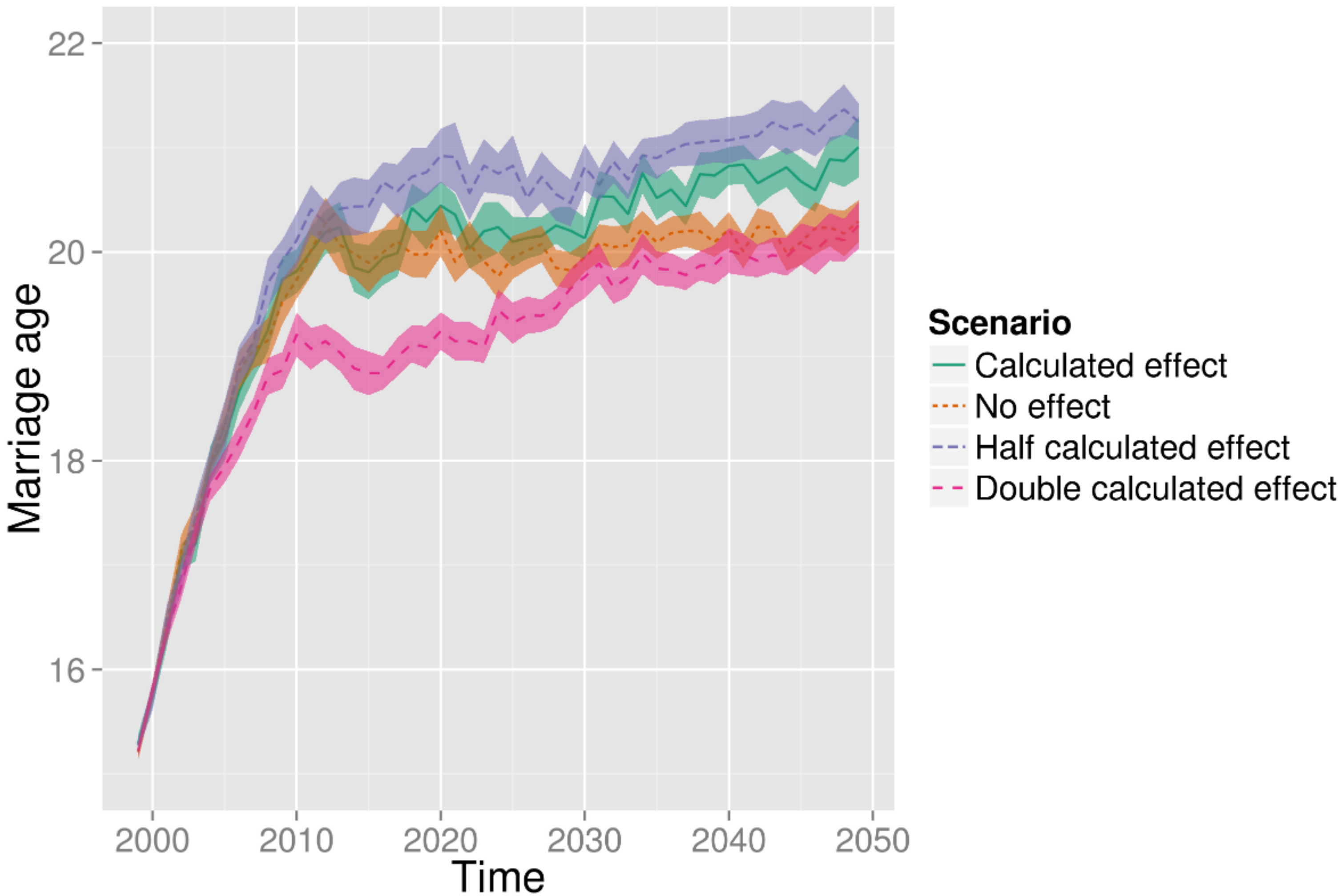
↓ agricultural land → ↑ marriage age
↓ agricultural land → ↑ time to first birth

Scenario	First birth time (months)	Marriage age (years)	Agricultural land use (%)
No effect	18.87	20.29	26.89
Calculated effect	**16.69**	**21.01**	26.66
Half effect	**17.17**	**21.42**	**28.39**
Double effect	**16.82**	20.18	**24.36**

Scenario set A: first birth timing feedback



Scenario set B: marriage age feedback



Conclusions

- Weak feedbacks between decision-making and landscape change can lead to significant differences in model outcomes.

- However, household structure and total fertility rates are stronger controls on land use change than the micro-level connections studied here

- When evaluating results, significance **and** effect size must be considered
  - ABM can be a useful tool to examine the dynamic effect of observed relationships over time and space

Getting started in ABM

Things to consider before building an ABM

- Research question!

- What specific outcomes are you interested in?
  - Land cover? Land use? Habitat change?
  - Population dynamics?
  - Processes? (learning, feedbacks)

- How much detail do you need?
  - Do you need to model individual people?
  - Do you need to model the entire population?

Verification and validation

1. Progressive model building and debugging
2. Uncertainty testing (extreme tests, and extreme combination tests)
3. Empirical validation: comparing model output data to empirical data
4. Sensitivity analysis: examining how model outcomes vary with a small change in key parameters, and
5. Experience or expert opinion

(from An et al. 2005)

Software options for building ABM

- NetLogo
- Swarm (Java based)
- Repast Simphony (ReLogo, Java, C++)
- Python and PyABM
- Any programming language can be used
  - But: using a non-standard language not recommended
  - Limits adoption of your model by others
  - Proliferation of alternative languages makes models difficult to evaluate and replicate

PyABM

- Agent-based modeling toolkit written in Python
- PyABM facilitates
  - Coding basic types of agents
  - Parameter handling (useful in sensitivity testing and model verification)
  - Logging model output
  - Shapefile input / output
- Tested on Windows, Linux, Mac OS

Getting PyABM

- To download code:

  https://github.com/azvoleff/pyabm

- To view documentation, go to:

  http://www.azvoleff.com/

  – Click on "Research"
  – Click on "PyABM"
  – And then scroll down to "Online documentation"

Getting Chitwan ABM

- To download code:

  https://github.com/azvoleff/chitwanabm

- To view documentation, go to:

  http://www.azvoleff.com/

  – Click on "Research"
  – Click on "ChitwanABM"
  – And then scroll down to "Online documentation"

Chitwan ABM dependencies (1 of 2)

- R (download from CRAN), and several R packages. Run the following code from within R to install packages:

```
install.packages(c("plyr", "ggplot2", "lubridate", "reshape2", "reshape", "gstat",
"rgdal", "grid", "RColorBrewer", "scales"))
source("http://bioconductor.org/biocLite.R")
biocLite("rhdf5")
```

- Python (version 2.7), and several Python packages (download packages from Gohlke's webpage):
  - pyabm
  - numpy
  - matplotlib
  - tables

Chitwan ABM dependencies (2 of 2)

- Geospatial Data Abstraction Library (GDAL)
  - Download GDAL binaries - For Python 2.7, download the "MSVC2008 (Win64) -stable" version
  - Add the installation directory 'bin' folder to the system PATH, and make a GDAL_DATA environment variable pointing to the 'data' folder inside the GDAL installation directory

- Chitwan Valley Family Study (CVFS) data
  - Data is restricted access, see Li to be added to agreement we have with ICSPR
  - Metadata online at ICPSR

Notes on coding in Windows…

- If you are editing the code, I suggest using git to track your changes
  - Git is a version control system – logs all changes to source code
  - github.com offers free git hosting for open-source projects

- Other useful open-source software:
  - dngrep – search and replace
  - métamorphose – file renaming

Good ABM references (1 of 2)

- An, L. 2012. Modeling human decisions in coupled human and natural systems: Review of agent-based models. *Ecological Modelling* 229:25–36.
- **An, L., M. A. Linderman, J. Qi, A. Shortridge, and J. Liu. 2005. Exploring complexity in a human environment system: an agent-based spatial model for multidisciplinary and multiscale integration. *Annals of the Association of American Geographers* 95 (1):54–79.**
- An, L., A. Zvoleff, J. Liu, and W. G. Axinn. in press. Agent based modeling in coupled human and natural systems (CHANS): Lessons from a comparative analysis. *Annals of the Association of American Geographers*.
- **Grimm, V., U. Berger, D. L. DeAngelis, J. G. Polhill, J. Giske, and S. F. Railsback. 2010. The ODD protocol: a review and first update. *Ecological Modelling* 221 (23):2760–2768.**
- **Manson, S. M. 2001a. Calibration, verification, and validation. In *LUCC Report Series*, LUCC Report Series., eds. D. C. Parker, T. Berger, and S. M. Manson. Irvine, California, USA.**
- Entwisle, B., G. P. Malanson, R. R. Rindfuss, and S. J. Walsh. 2008. An agent-based model of household dynamics and land use change. *Journal of Land Use Science* 3 (1):73.
- Oreskes, N., K. Shrader-Frechette, and K. Belitz. 1994. Verification, validation, and confirmation of numerical models in the Earth sciences. *Science* 263 (5147):641–646.

Good ABM references (2 of 2)

- **Parker, D. C., B. Entwisle, R. R. Rindfuss, L. K. VanWey, S. M. Manson, E. F. Moran, L. An, P. Deadman, T. P. Evans, and M. A. Linderman. 2008. Case studies, cross-site comparisons, and the challenge of generalization: Comparing agent-based models of land-use change in frontier regions. *Journal of Land Use Science* 3 (1):41–72.**
- Parker, D. C., S. M. Manson, M. A. Janssen, M. J. Hoffmann, and P. Deadman. 2003. Multi-agent systems for the simulation of land-use and land-cover change: A review. *Annals of the Association of American Geographers* 93 (2):314–337.
- Zvoleff, A., and L. An. in review. The effect of reciprocal connections between demographic decision making and land use on decadal dynamics of population and land use change. *Ecology and Society*. (in review)
- Zvoleff, A., and L. An. 2013. Analyzing human-landscape interactions: tools that integrate. *Environmental Management*. http://dx.doi.org/10.1007/s00267-012-0009-1.
- **Edited volume (free through UCSB): Heppenstall, A. J., A. T. Crooks, and L. M. See. 2012. *Agent-based models of geographical systems*. Springer.**
  - Abdou, M., L. Hamill, and N. Gilbert. Designing and Building an Agent-Based Model.
  - Grimm, V., and S. F. Railsback. Designing, formulating, and communicating agent-based models.
  - Manson, S. M., S. Sun, and D. Bonsal. 2012. Agent-based modeling and complexity.

Data

Data

Go over Shared_Data folder

The TEAM Network

The Tropical Ecology Assessment and Monitoring (TEAM) Network

- Global network monitoring change in tropical forests
- Multiple protocols
  - Climate
  - Terrestrial mammals (camera trapping)
  - Vegetation monitoring (1 ha plots)
  - Zone of Interaction (ZOI, anthropogenic effects)
- Network of sites (8 in Latin America, 5 in Africa, 3 in Asia)

CONSERVATION INTERNATIONAL

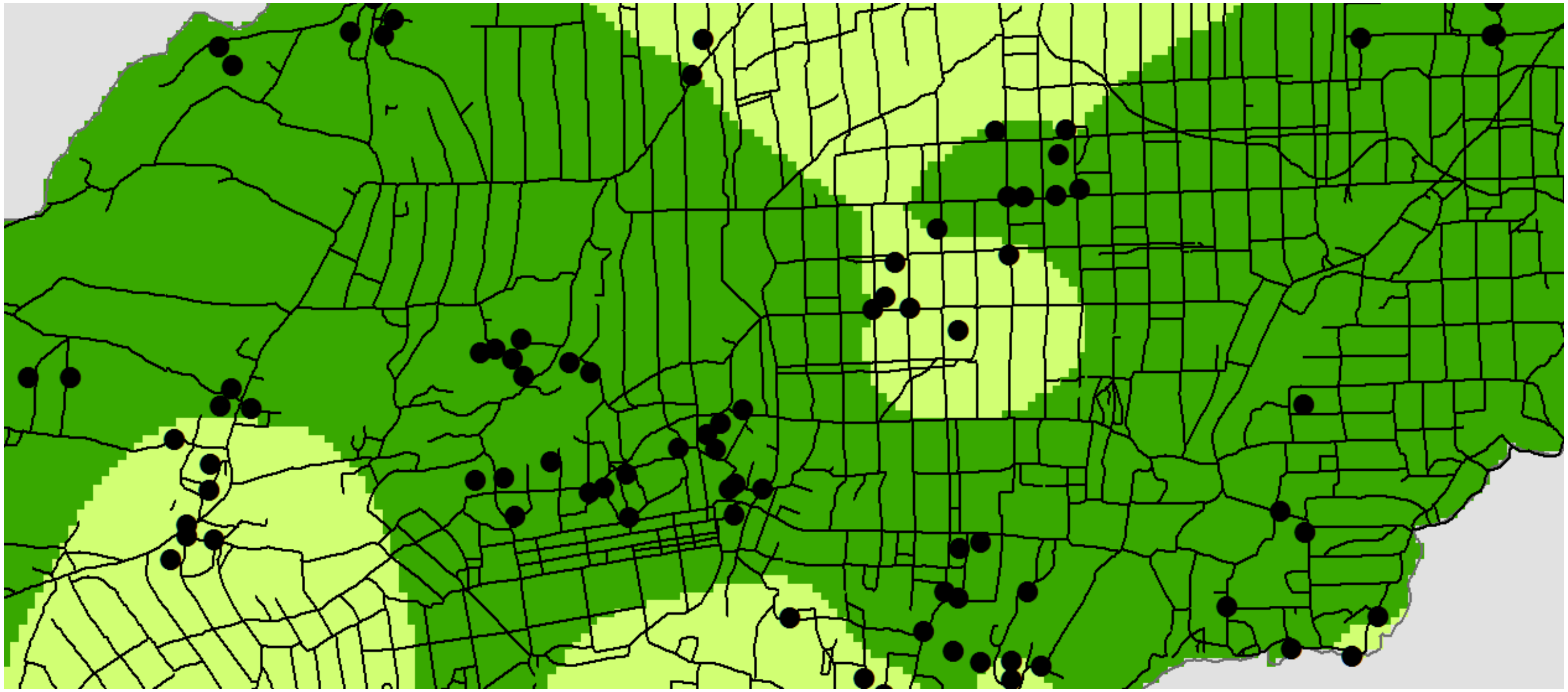
TEAM NETWORK
TROPICAL ECOLOGY ASSESSMENT AND MONITORING

The Tropical Ecology Assessment and Monitoring (TEAM) Network

- All TEAM data is freely available:
  - http://www.teamnetwork.org
- Very interested in graduate student projects
  - Site-specific studies of land cover change
  - Studies of animal behavior
  - Examination of forest responses to climate change
  - Syntheses across sites
- Contact me if interested (or go online)

CONSERVATION INTERNATIONAL

TEAM NETWORK
TROPICAL ECOLOGY ASSESSMENT AND MONITORING



Thank you. Questions?