

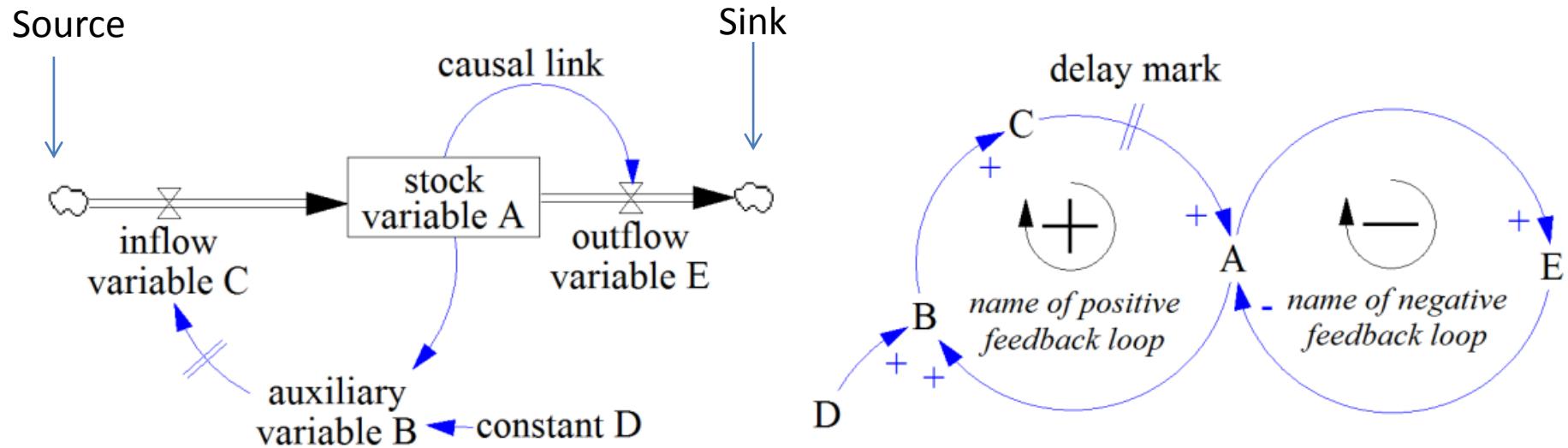
Simulation and Analysis on Steroids with Python

Erik Pruyt

Overview

1. Traditional System Dynamics Model. & Simulation
2. EMA Workbench & AS: Simulation & analysis with Python scripts and ipython / jupyter notebooks
3. PySD: Translating Vensim models to Python + Calibration, Simulation & Analysis in Python
4. From Aggregated to Disaggregated and Hybrid Modelling & Simulation
 - (3) Database + 1 model for many regions (wo interaction)
 - (4) MSc thesis of Philipp Schwarz
 - (5) Hierarchical geo-spat. SD-ABM-DES M&S in AnySim

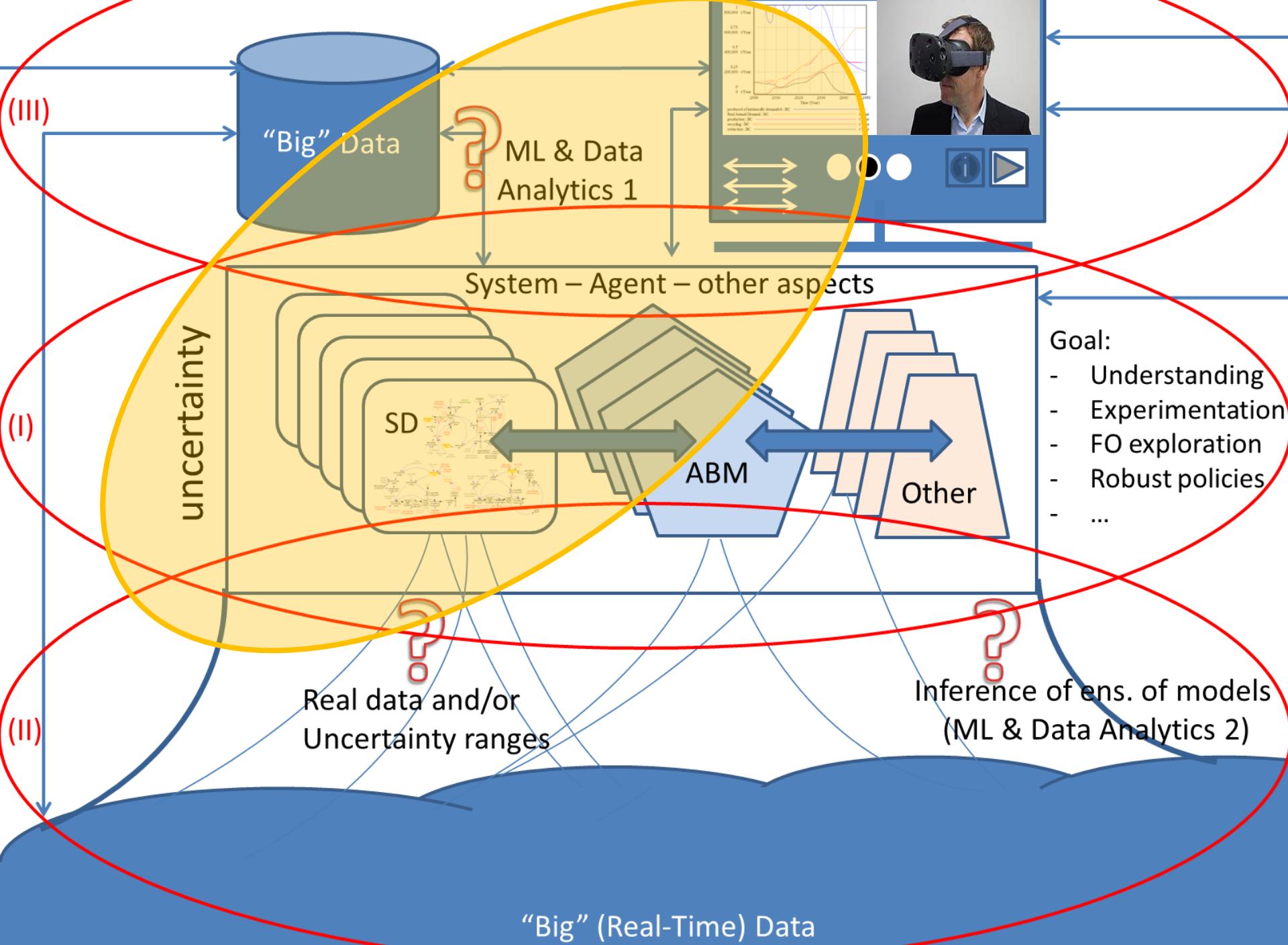
Traditional System Dynamics Modelling & Simulation



- Causal Loop Diagrams -> causal effects and feedback loops
- Stock Flow Diagrams -> ``Simulation Models``:
Stocks, flows, delays, auxiliaries, parameters, (sources and sinks)
- Hybrid Diagrams -> Causal Loop Diagrams + Stocks (&Flows)

Overview II: Enhancing SD with Python

- Pre 2009: “traditional” System Dynamics modelling
- “Trying” to do uncertainty & robustness analysis
- 2009: Dynamic complexity x “Deep Uncertainty”,
first w Caner Hamarat , later w Jan Kwakkel et al.
- Mexican Flu: model in 30’ + 3 weeks of exploration
- Post ” ”: case-based development -> **EMA wrkbnnch**
- 2015: Workshop setting? => “**Adaptive Sampling**”
- 2015: **pySD** (James Houghton @ MIT)
- 2016: **python+** (Philipp Schwarz, Bas Keijser, WLA)
- 2016: **AnySim** (Reza Hesan)



1) EMA Workbench, AS2, etc: Simulation & Analysis with Python & jupyter notebooks

- Exploratory Modelling and Analysis
 - Why? What? How?
 - Examples
 - Example 1 - Flu: ExampleFluWithPoliciesV.ipynb & PrimFlu.ipynb
 - Example 2 - Ebola:
EbolaV26Models123456wPOLSprimmedV03_overtimeD291UncRed
- Adaptive Sampling (AS2)
 - Why? What? How?
 - Examples
 - Energy/Climate to State Instability
 - BSSv02b_Burnout01short_v01.ipynb

Why?

Deep uncertainty: Uncertainty & Ambiguity

- Decision makers and stakeholders do not know or cannot agree on the outcomes of interest, the system under study, or future developments
- How to let planning and decision making proceed despite the presence of different models, different beliefs about the future, and disagreement on which outcomes of interest should be included?
 - ⇒ Not just sampling across input parameter ranges: inputs (lack/conflict of data/info), perspectives (different models), evaluation of desirable, methods,...
 - ⇒ Sets of plausible models of the system, of outcomes of interest without a priori weighting, and of scenarios



Approaches for Dealing with Uncertainty, Ambiguity and Surprises in modelling

- Explorative modelling: embrace uncertainty, model un/knowns and multiple perspectives, what-if explorations / test policy robustness
=> Inclusion of different models, relations, inputs,...
- Consolidative / consensual modelling: reduce uncertainty as much as possible, model “known”/agreed upon, and test sensitivity
- Steps still +/- the same, but everything changes

What? Exploratory Modelling & Analysis

Model-based methodology for systematically exploring deep uncertainty & testing policy robustness

1. Identification of ‘certainties’ and ‘uncertainties’ & construction of ‘plausible’ SD models
 - 2a. Open exploration
 1. Model-based generation of many plausible scenarios
 2. Analysis of ensemble of plausible scenarios using all sorts of algorithms (clustering, PRIM, ...)
 3. Identification of interesting types of scenarios + origins
 - 2b. Direct searches
 3. Policy design and testing across all plausible scenarios
- Need for fast and manageable models... => SD!!!

Open exploration



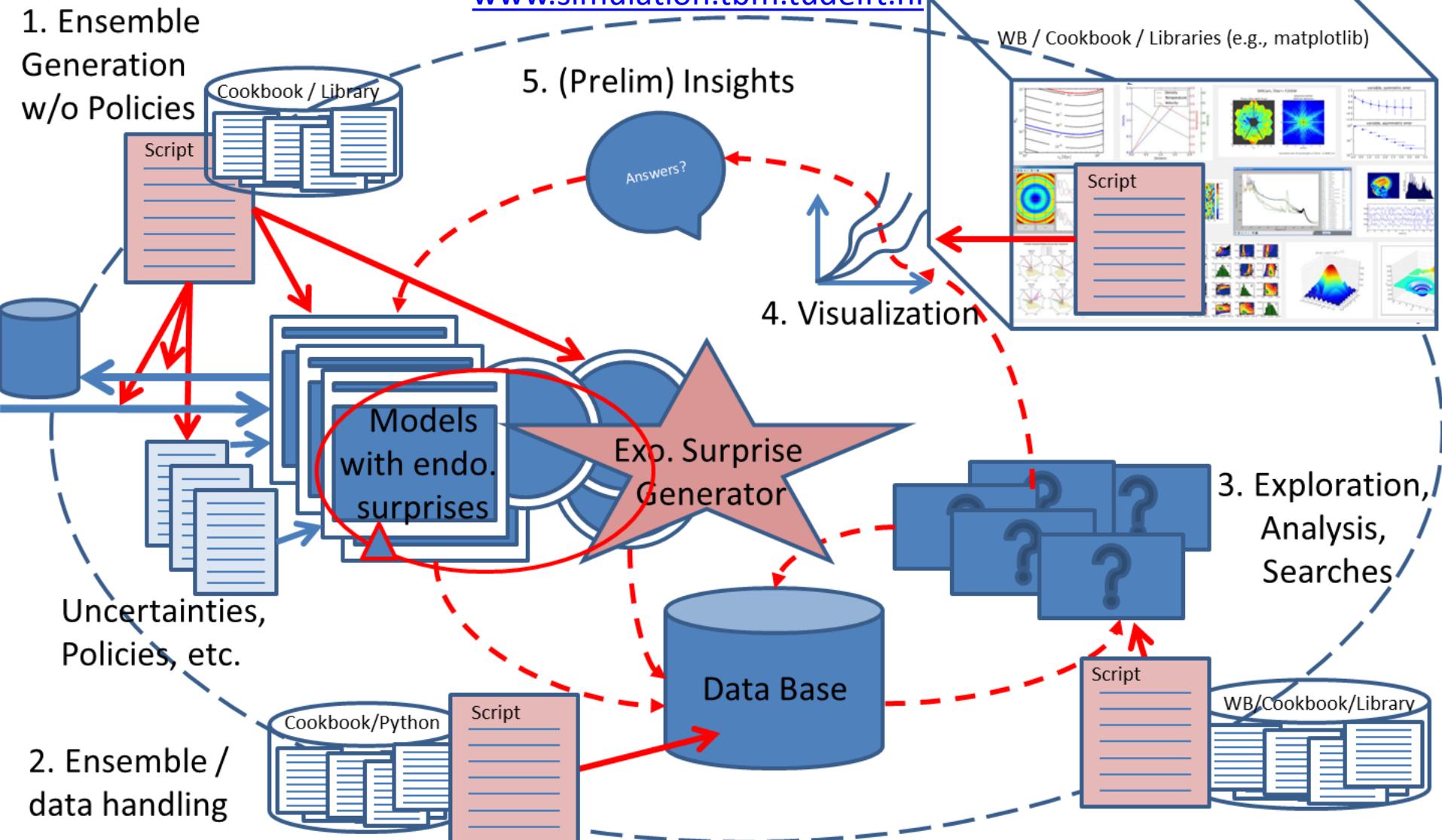
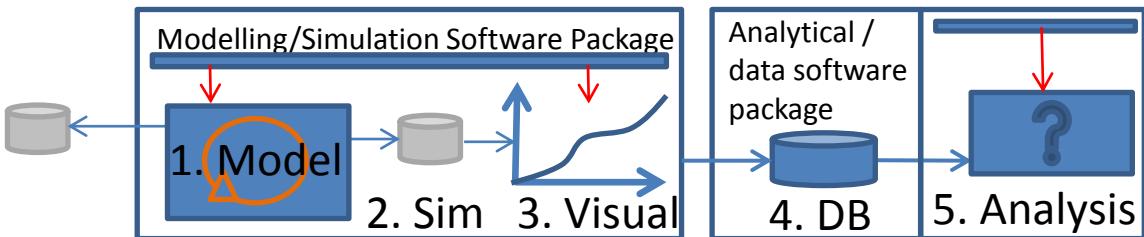
- Design of experiments
 - Factorial methods
 - Monte Carlo sampling
 - Latin Hypercube sampling
 - Behaviour Space sampling
- Used for
 - Identification of bandwidth outcomes
 - Identification of types of behavior
- Subsequent analyses
 - Scenario discovery
 - (Policy analysis + testing)

Directed search



- Optimization for search
 - Multi-objective optimization
 - Robust optimization
- Used for
 - Worst case discovery
 - Identification of boundaries where behavior switches
 - Policy design
- Subsequent analyses
 - Trade-offs
 - Tipping points

How? With our EMA WorkBench



EMA WB, Python, Jupyter Notebooks

Example.py - Eclipse SDK

E Search Project Pydev Run Window Help

flu_multiplot flu_vensimExample envelope_lines

```
class FluModel(VensimModelStructureInterface):  
    #base case model  
    modelFile = r'\\FLUvensimVibasecase.vpm'  
  
    outcomes = [Outcome('deceased population region 1', time=False),  
                Outcome('infected fraction R1', time=True)]  
  
    uncertainties = [  
        ParameterUncertainty((0, 0.5), "additional se"),  
        ParameterUncertainty((0, 0.5), "additional se"),  
        ParameterUncertainty((0.0001, 0.1), "fatality ratio"),  
        ParameterUncertainty((0.0001, 0.1), "fatality rate"),  
        ParameterUncertainty((0, 0.5), "initial immun"),  
        ParameterUncertainty((0, 0.5), "initial immun"),  
        ParameterUncertainty((0, 0.9), "normal interc"),  
        ParameterUncertainty((0, 0.5), "permanent imm"),  
        ParameterUncertainty((0, 0.5), "permanent imm"),  
        ParameterUncertainty((0.1, 0.75), "recovery time"),  
        ParameterUncertainty((0.1, 0.75), "recovery time"),  
        ParameterUncertainty((0.5, 2), "susceptible t"),  
        ParameterUncertainty((0.5, 2), "susceptible t"),  
        ParameterUncertainty((0.01, 5), "root contact"),  
        ParameterUncertainty((0.01, 5), "root contact"),  
        ParameterUncertainty((0, 0.15), "infection rat"),  
        ParameterUncertainty((0, 0.15), "infection rat"),  
        ParameterUncertainty((10, 100), "normal contac"),  
        ParameterUncertainty((10, 200), "normal contac")]  
  
    if __name__ == "__main__":  
        logging.basicConfig(level=logging.INFO)  
  
        model = FluModel(r'..\..\models\JAN\flu', "fluCase")  
        ensemble = ModelEnsemble()  
        ensemble.set_model_structure(model)  
  
        policies = [{"name": "no policy",  
                    "file": r'\\FLUvensimVibasecase.vpm'},  
                    {"name": "static policy",  
                    "file": r'\\FLUvensimVistatic.vpm'},  
                    {"name": "adaptive policy",  
                    "file": r'\\FLUvensimVodynamic.vpm'}]  
        ensemble.add_policies(policies)  
  
        ensemble.parallel = True  
  
        import time  
        start_time = time.time()  
        results = ensemble.perform_experiments(10000)  
        print time.time() - start_time  
        save_results(results, r"10000 flu cases policies.cPickle")
```

Problems Console Hierarchy View PyUnit

<terminated> D:\eprj\workspace\EMAProjects\JAN\flu\FLUvensimExample.py
[INFO/MainProcess] generating cases

IPy notebooks/ IPy EbolaV26SRBScorrect

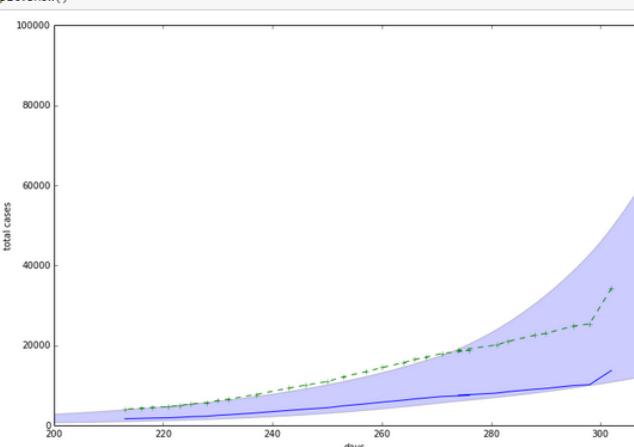
localhost:8888/notebooks/notebooks/EbolaV26SRBScorrect.ipynb

IP[y]: Notebook EbolaV26SRBScorrect Last Checkpoint: Dec 01 16:11 (autosaved)

File Edit View Insert Cell Kernel Help

Cell Toolbar: None

plt.show()



total cases

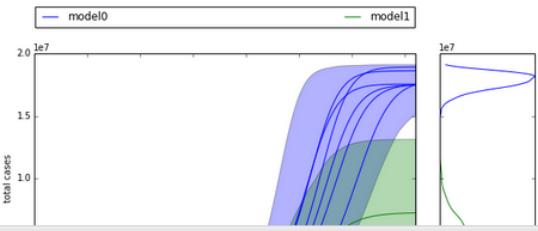
days

So yes, until today a good fit (both). The question is: are they different from here on?

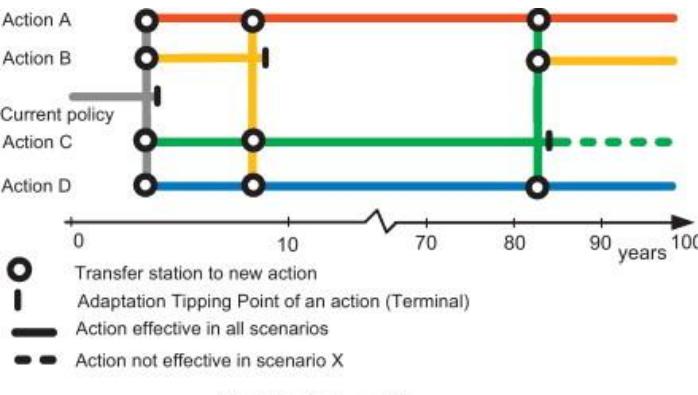
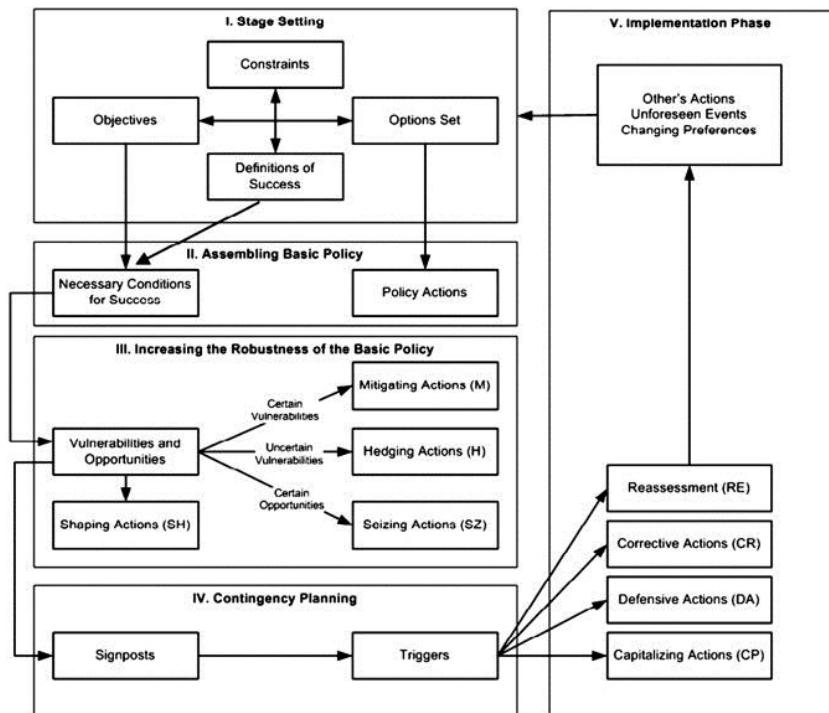
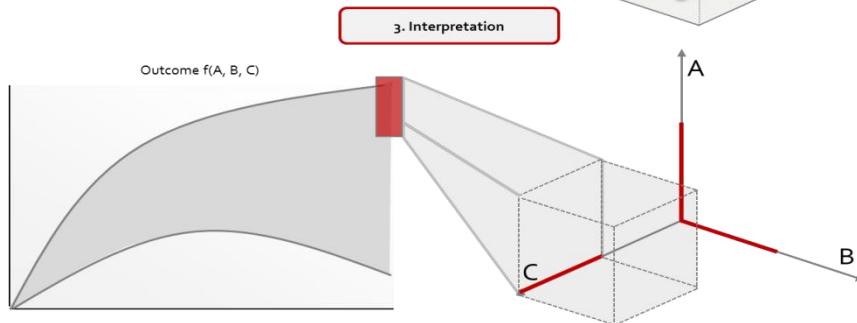
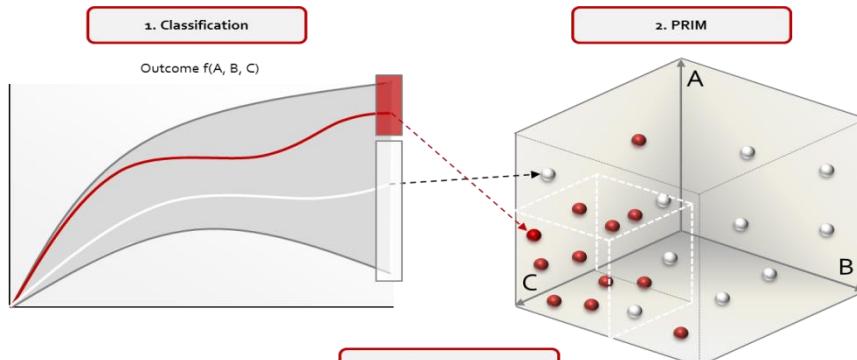
In [24]:

```
experiments, outcomes = results  
  
desired_nr_lines = 10  
nr_cases = results[0].shape[0]  
indices = np.arange(0, nr_cases, nr_cases / desired_nr_lines)  
  
plotting.TIME_LABEL = "Time (days)"  
outcomes = results[1].keys()  
outcomes.pop(outcomes.index("TIME"))  
for outcome in outcomes:  
    fig, axes_dict = plotting.lines(results, outcomes_to_show=outcome, group_by='model', show_envelope=True,  
                                    experiments_to_show=indices, titles="", density=plotting.KDE)  
    axes_dict[outcome].set_xlim(200, 300)  
    axes_dict[outcome].set_ylim(ymin=0)  
    fig.set_dpi(150)  
    fig.set figsize(10)  
    fig.set_fighight(5)  
  
plt.show()
```

model0 model1

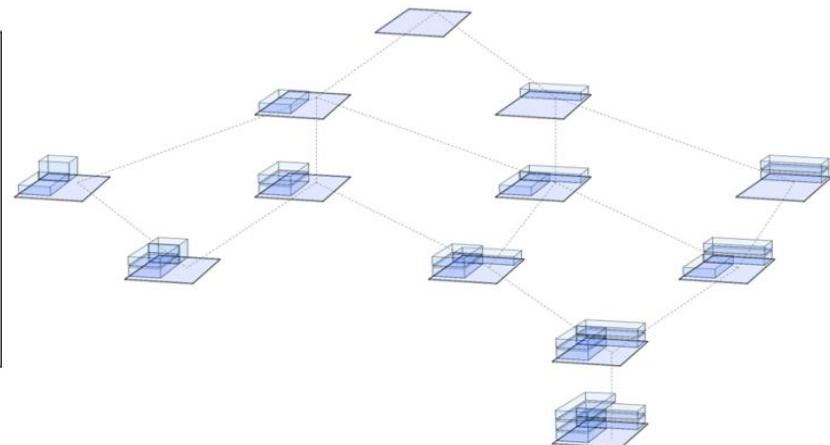


New Methods, Techniques, Algorithms



Path actions	Relative Costs	Target Side effects	Side effects
1 ●	+++	+	0
2 ○●	++++	0	0
3 ○○●	+++	0	0
4 ○○○●	+++	0	0
5 ○○○○●	0	0	-
6 ○○○○○●	++++	0	-
7 ○○○○○○●	+++	0	-
8 ○○○○○○○●	+	+	---
9 ○○○○○○○○●	++	+	---

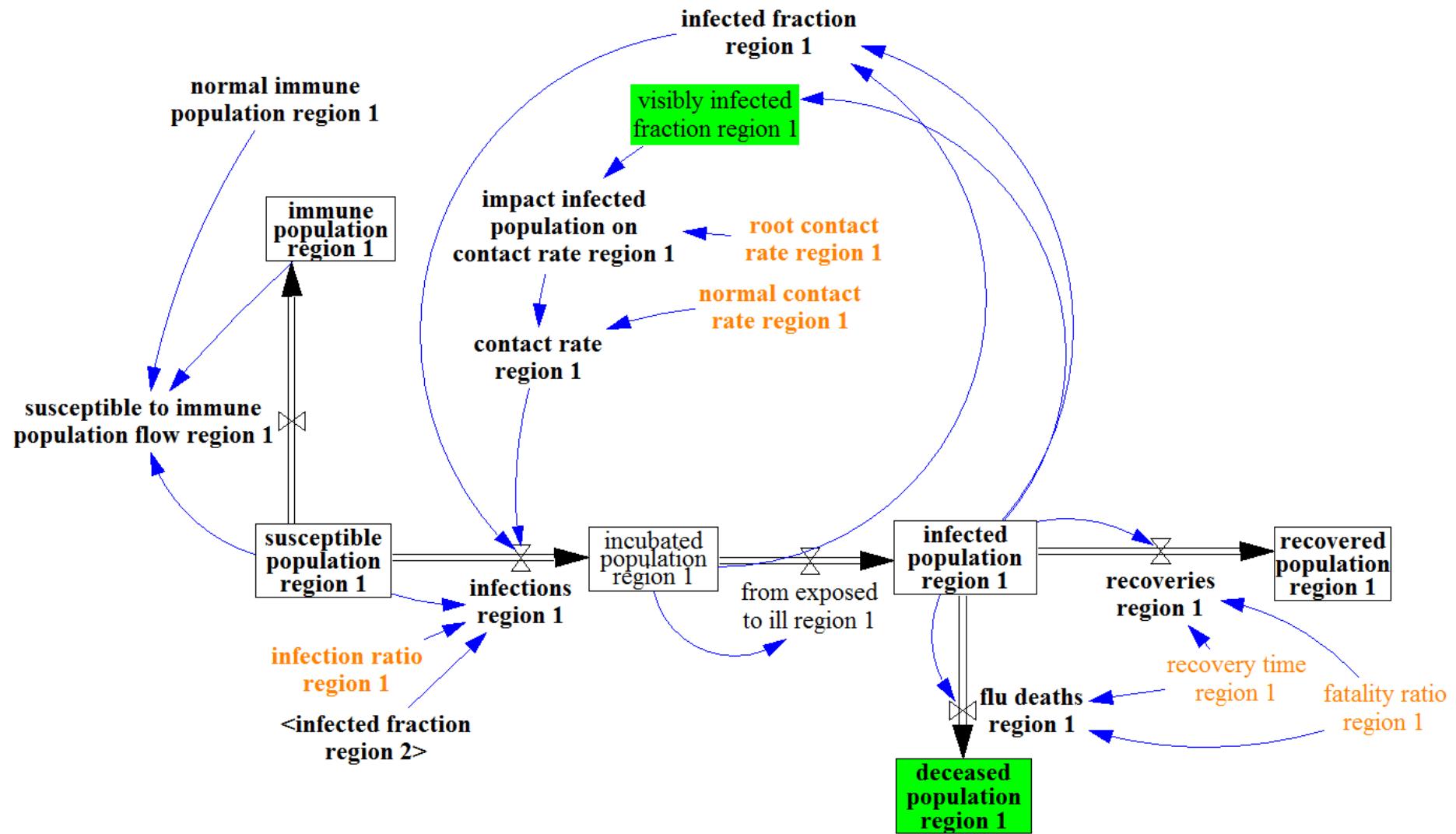
Scorecard pathways



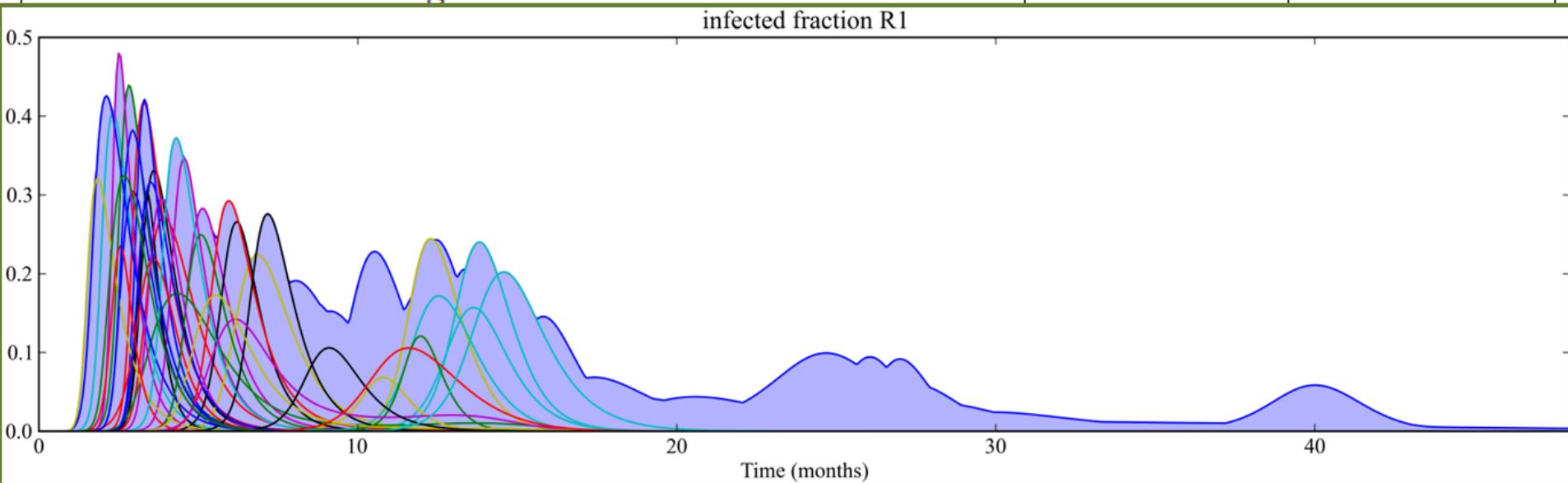
Example 1: 2009 A(H1N1)v

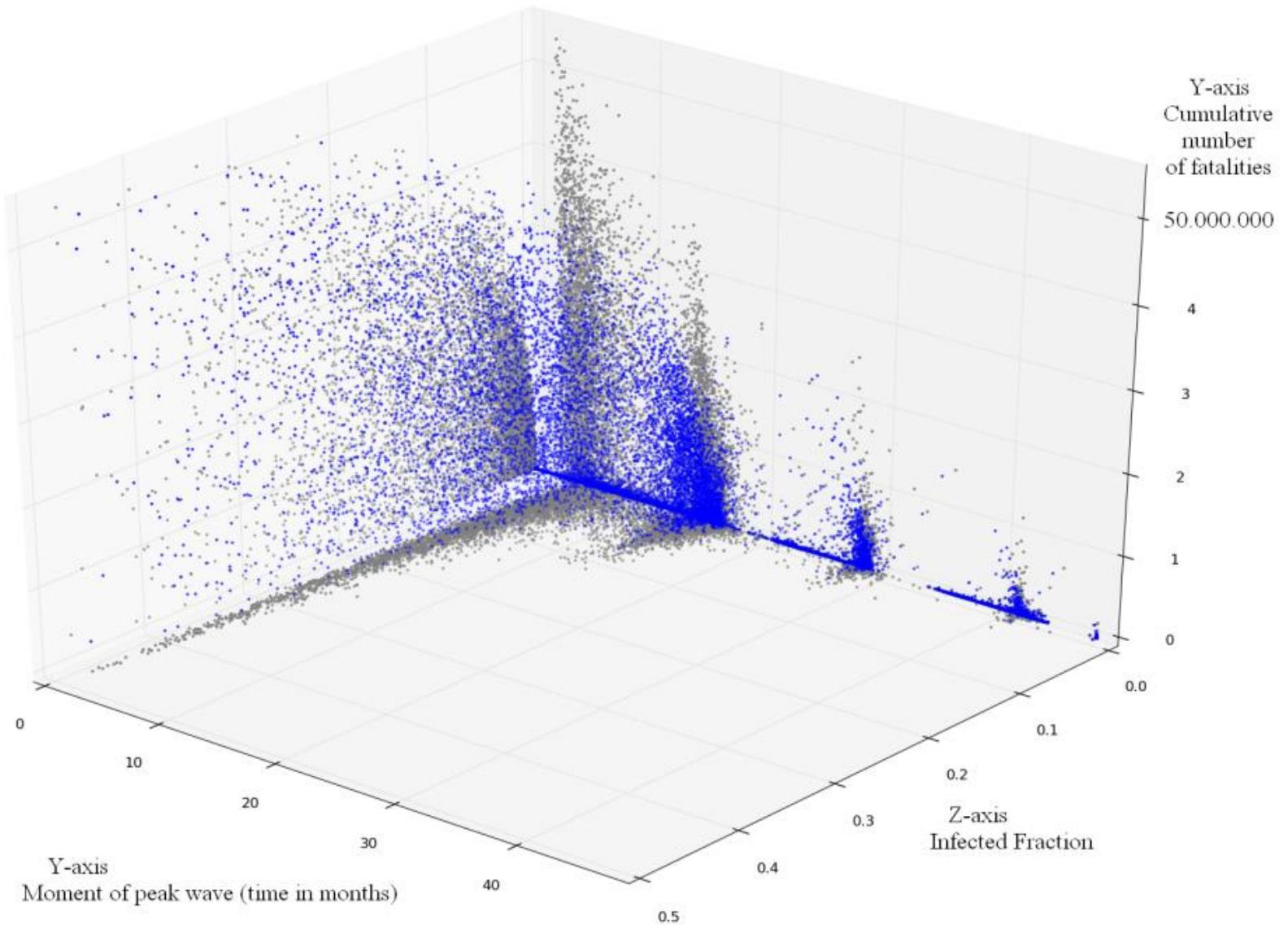
- April 2009: imminent epidemic/pandemic of new flu strand
- **Unknowns!**

Date	24 April	30 April	08 May	20 May	12 June	20 July	21 August
Infectivity Ro	unknown unknown	unknown unknown	unknown 1–2; prob. 1.4–1.9	unknown 1–2; prob. 1.4–1.6	unknown —	unknown —	unknown [R up to 2]
Immunity	unknown	unknown	indications (elderly)	idem	idem	idem	idem
Virulence	unknown	unknown	unknown	unknown	unknown	mild and self-limiting	idem
Incubation period	unknown	unknown	long tail? (up to 8d)	—	median 3-4d range 1-7d	idem	idem
CFR (Mexico)	17%?	—	4%?	2%?	0.4–1.8%?	—	—
CFR (USA)	unknown	unknown	0.1%?	0.1%?	0.2%?	0.4%?	—
CFR (UK)	unknown	unknown	unknown	unknown	unknown	0.3%(-1%)?	0.1–0.2%? (0.35% ↘ ex.)
Age distribution	unknown	unknown	older people less affected?	—	skewed tow. younger	idem	idem
Antiviral suscep.	unknown	possible	indications	—	—	—	—
% asymptomatic	unknown	unknown	unknown	unknown	unknown	indications	33–50% (ass.)
Future severity?	unknown	unknown	unknown	unknown	unknown	unknown	unknown
Respective sources	(ECDC 2009g)	(ECDC 2009f)	(ECDC 2009b)	(ECDC 2009c)	(ECDC 2009a)	(ECDC 2009d)	(ECDC 2009e)



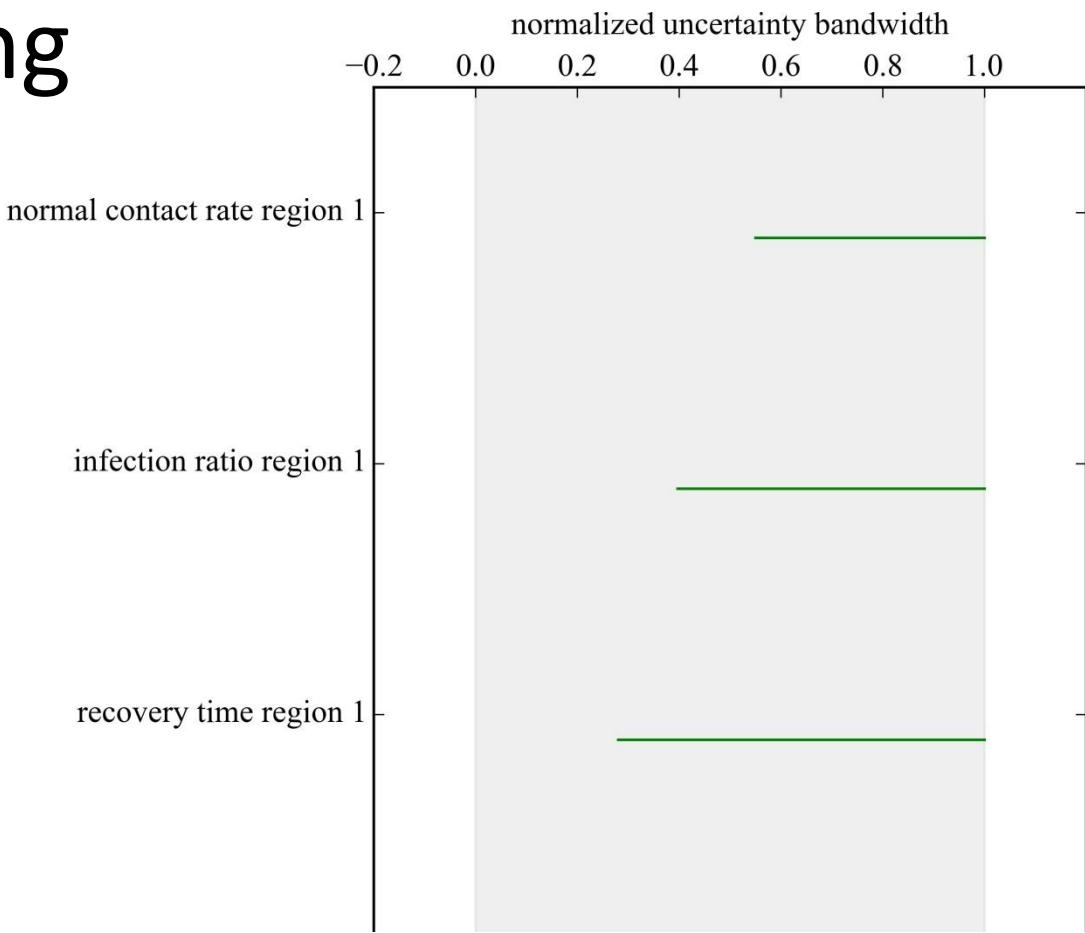
Parameter	Lower Limit	Upper Limit
additional seasonal immune population fraction region 1	0.0	0.5
additional seasonal immune population fraction region 2	0.0	0.5
fatality ratio region 1	0.0001	0.1
fatality ratio region 2	0.0001	0.1
initial immune fraction of the population of region 1	0.0	0.5
initial immune fraction of the population of region 2	0.0	0.5
normal interregional contact rate	0.0	0.9
permanent immune population fraction region 1	0.0	0.5
permanent immune population fraction region 2	0.0	0.5
recovery time region 1	0.2	0.8
recovery time region 2	0.2	0.8
root contact rate region 1	1.0	10.0
root contact rate region 2	1.0	10.0
infection ratio region 1	0.0	0.15
infection ratio region 2	0.0	0.15
normal contact rate region 1	10	100
normal contact rate region 2	10	200



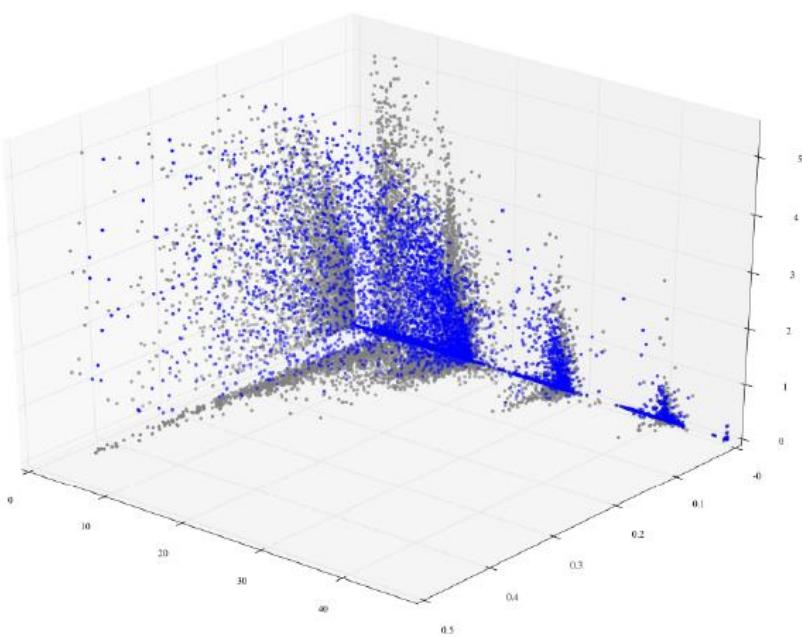


Machine Learning Algorithms

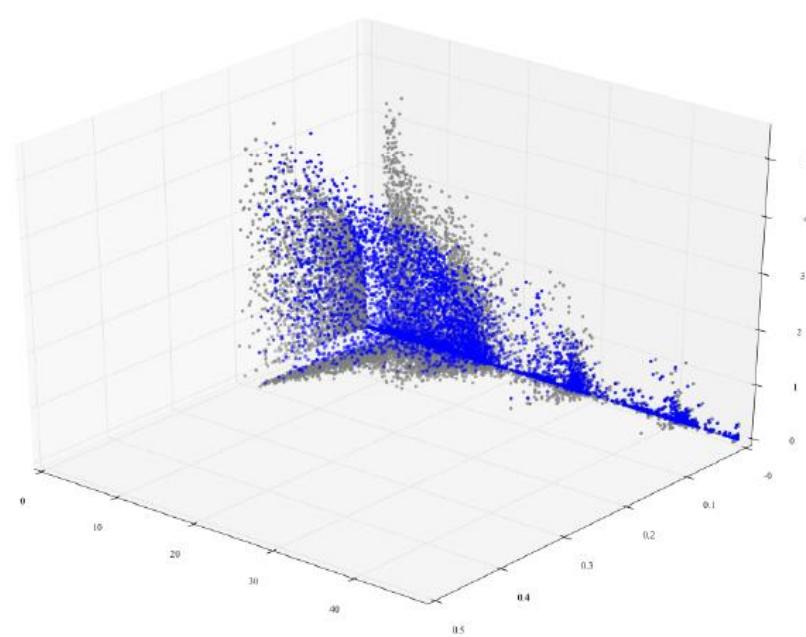
- Combined causes of ° of undes. runs
- Not just parametric
- Dynamics



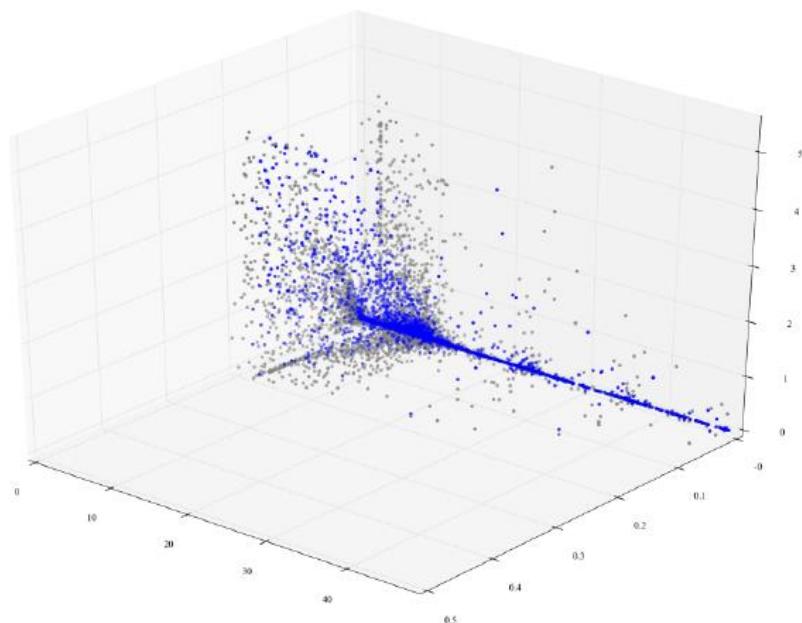
PRIM box bounding uncertainties	box 1: min:	max:	rest box min:	max:
normal contact rate region 1	59.490	99.986	10.003	99.986
infection ratio region 1	0.060	0.150	0.000	0.150
recovery time region 1	0.282	0.750	0.100	0.750
additional seasonal immune pop fraction R1	0.023	0.470	0.000	0.500
fatality ratio region 1	0.010	0.100	0.000	0.100
infection rate region 2	0.014	0.150	0.000	0.150
root contact rate region 1	0.012	4.693	0.014	4.998
permanent immune population fraction R1	0.000	0.478	0.000	0.500



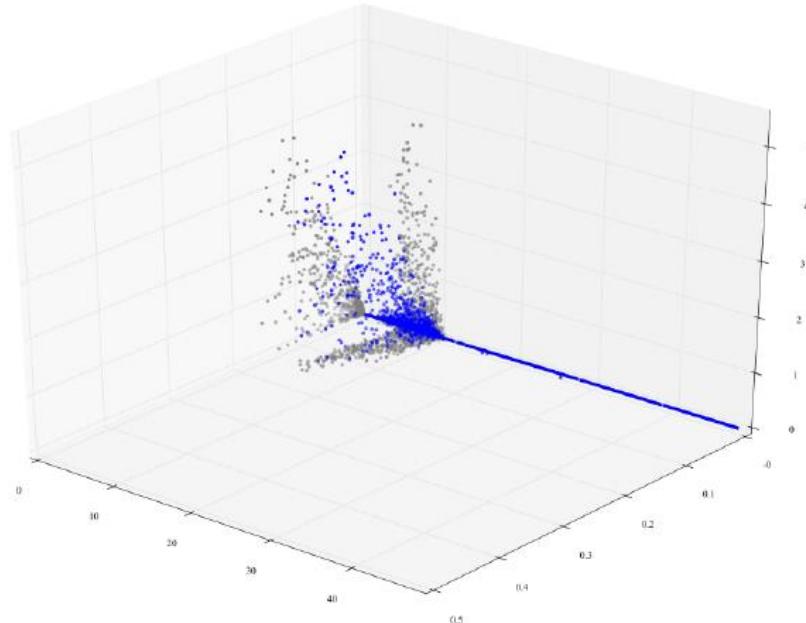
(a) No Policy



(b) Basic Policy

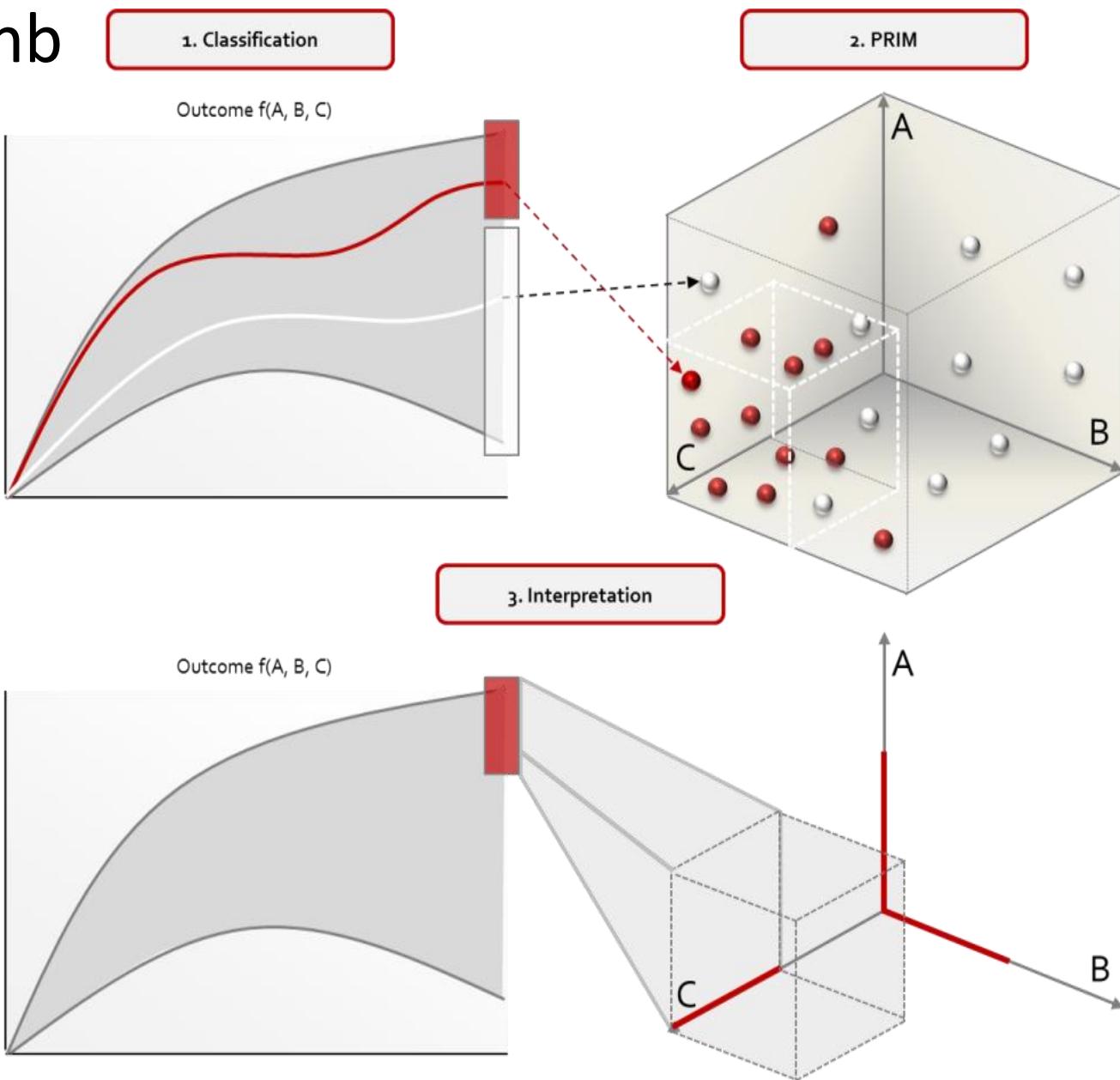


(c) Adaptive Policy

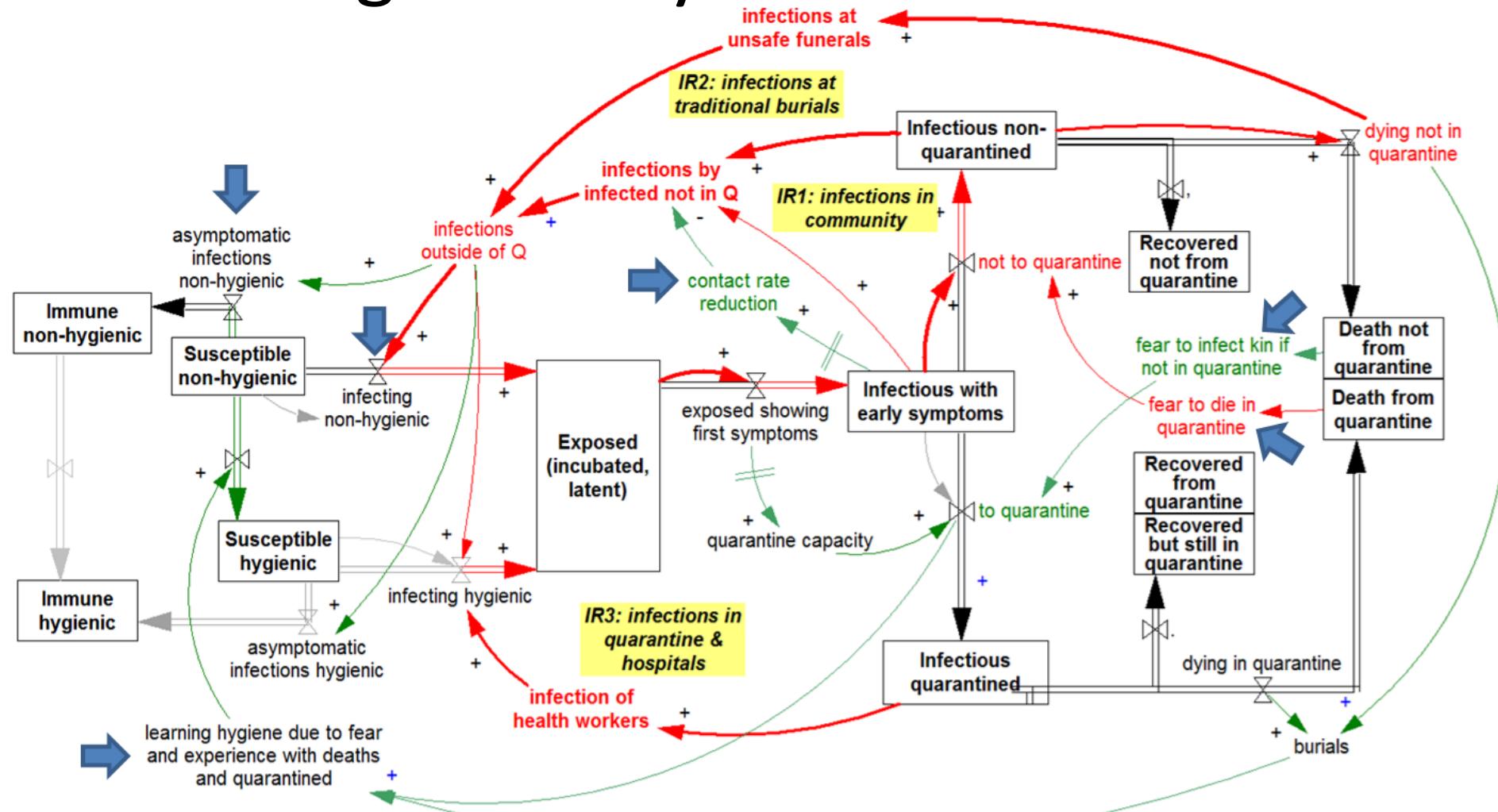


(d) Optimized Adaptive Policy

- ExampleFluWithPoliciesV.ipynb
- PrimFlu.ipynb

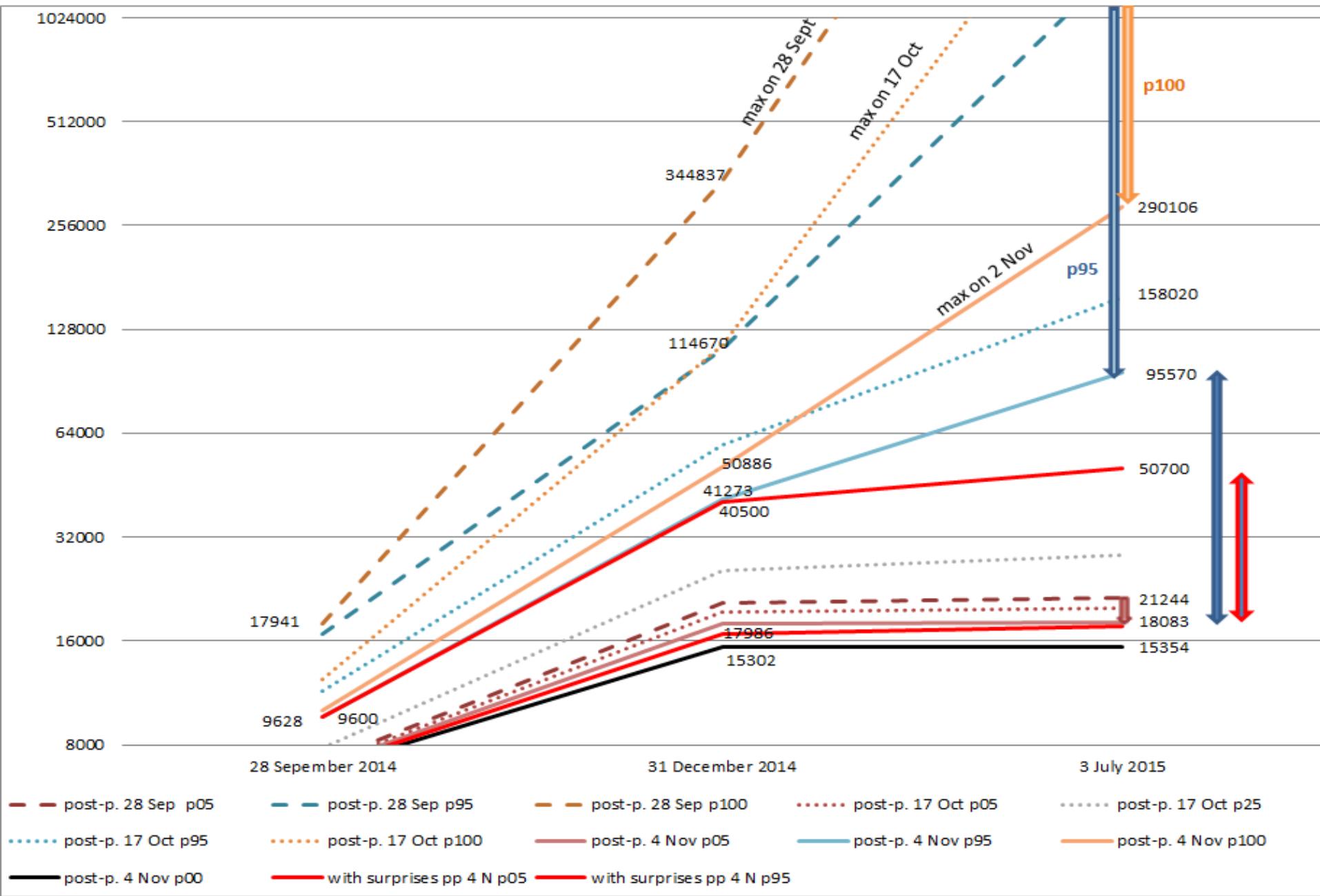


Example 2: Ensemble Forecasting and Testing of Policy Robustness - Ebola



For some of our Ebola studies, see: Pruyt *et al.*, 2015. *Systems Research & Beh. Science*

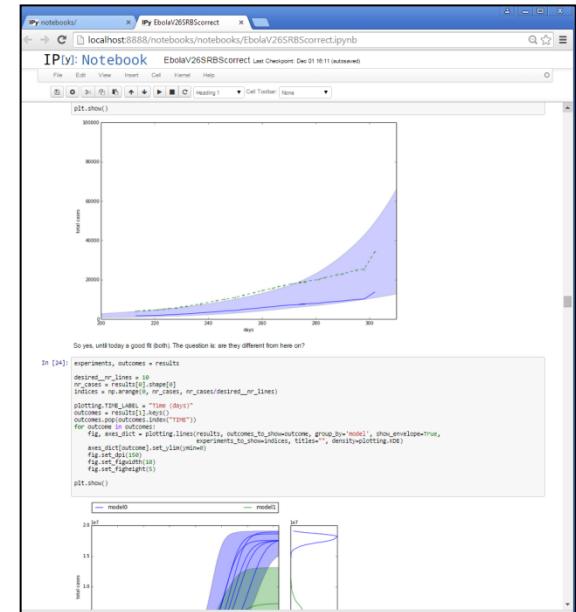
As well as: Auping *et al.*, 2016. *International Journal of Systems Science*



- EbolaV26Models123456wPOLSprimmedV03_overtime
D291UncRed.py

The screenshot shows the Eclipse IDE interface with the following details:

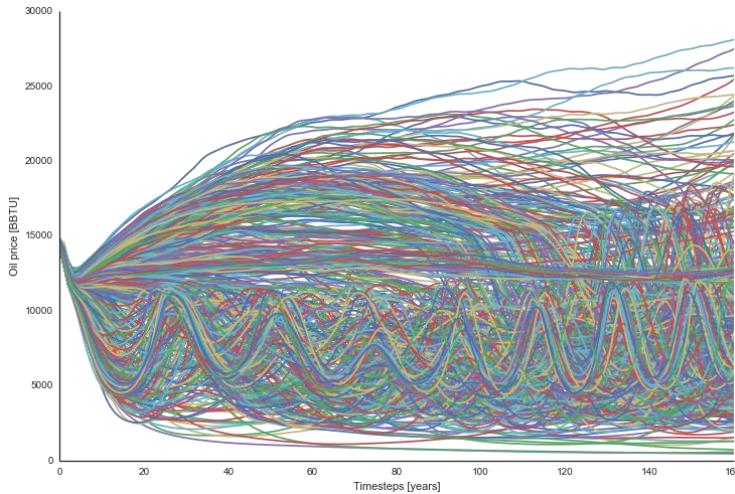
- File menu:** File, Edit, Source, Refactoring, Navigate, Search, Project, PyDev, Run, Window, Help.
- PyDev perspective:** Python, Java, SVN Repository Exploring, Team Synchronizing.
- Package Explorer:** Shows the project structure with packages like flu, fluMultipot, fluModel, fluUvensimExample, envelope_lines, and others.
- Code Editor:** Displays the fluModel class definition, which includes methods for creating a model, setting parameters, and running simulations. It also includes sections for uncertainty handling and logging.
- Outline View:** Shows the class hierarchy and member details.
- Problems View:** Shows no errors or warnings.
- Status Bar:** Shows the file is saved (fluModel.pydev).



- + jupyter notebook
 - Soft calibration with PRIM
 - Postprocessing
 - Plotting of ensembles (different policies)
 - PRIM to identify what works / does not work
 - ...

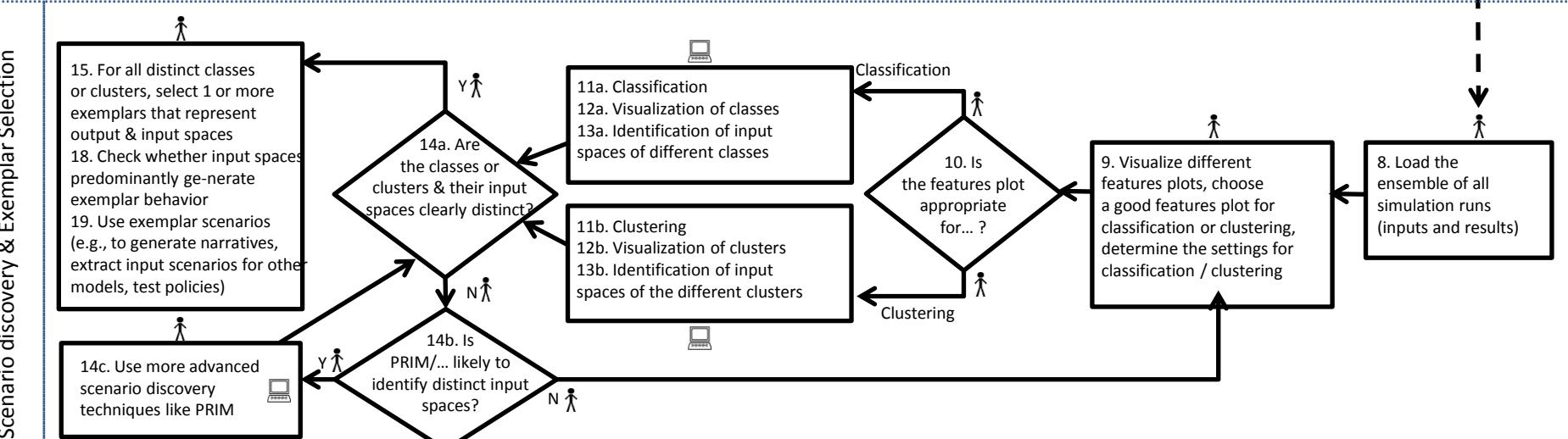
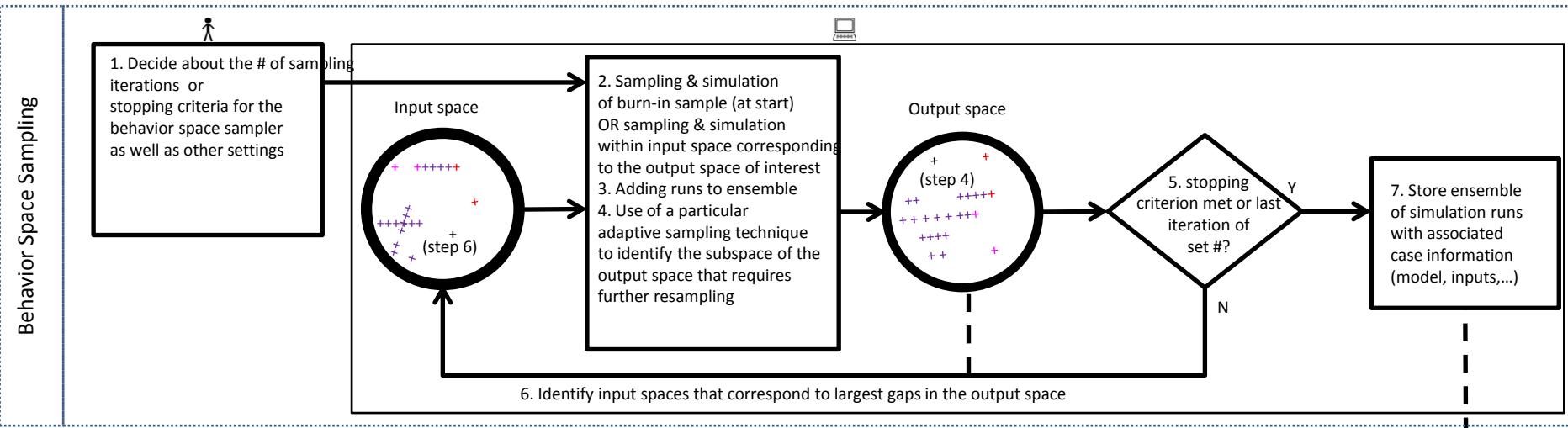
Adaptive Sampling: Why? What?

- What-if: really non-linear?
- Monte-Carlo? Latin Hyp?
- Sample, map, ID, resample
- For:
 - Scenario selection and narrative building
 - Scanning the spectrum of behaviours
 - Scanning the spectrum of policies without resistance
 - Identifying robust policy portfolios
 - Discovering surprising behaviours & corresponding input space

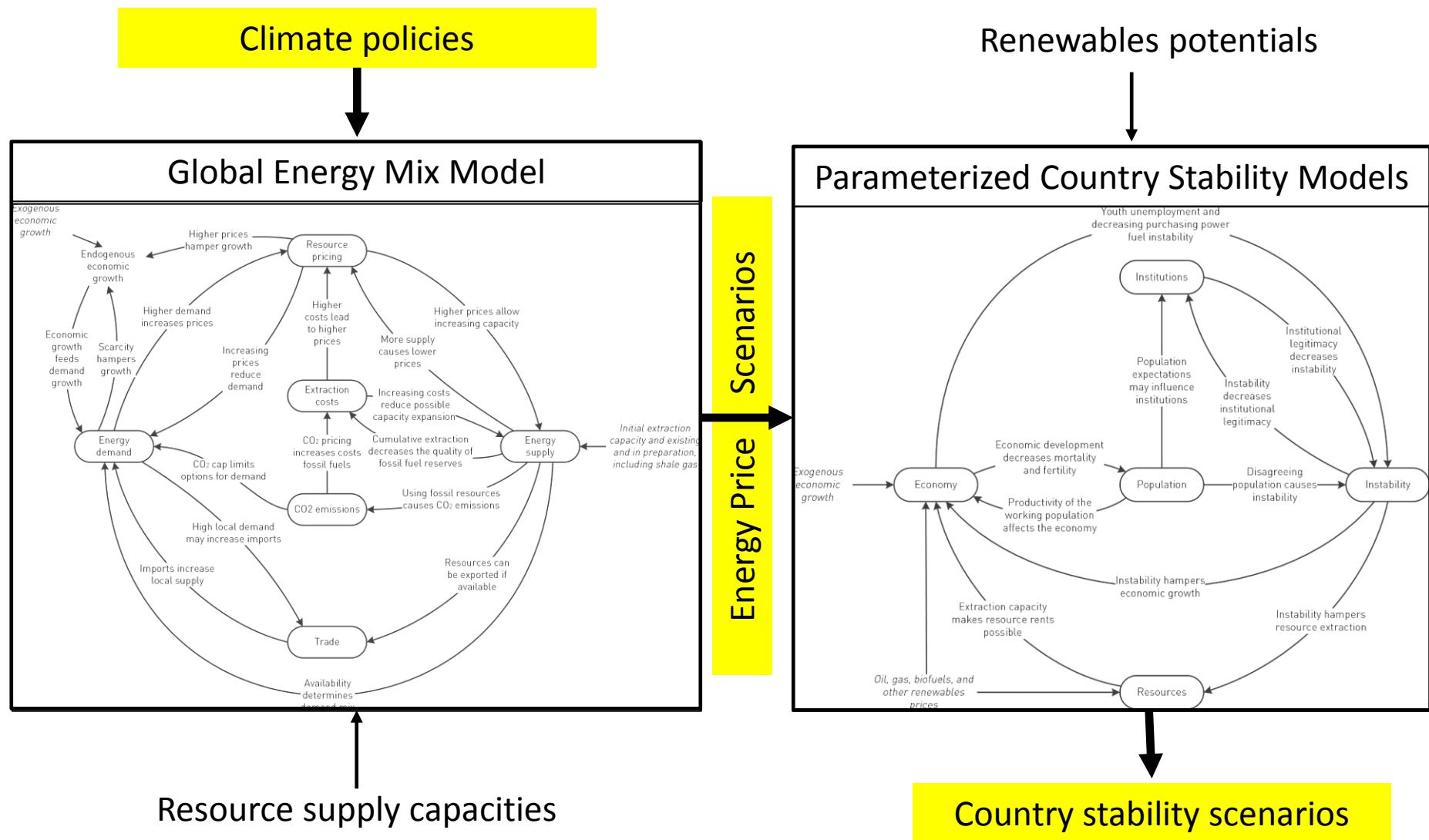


How? Identifying (Surprising) Behaviours?

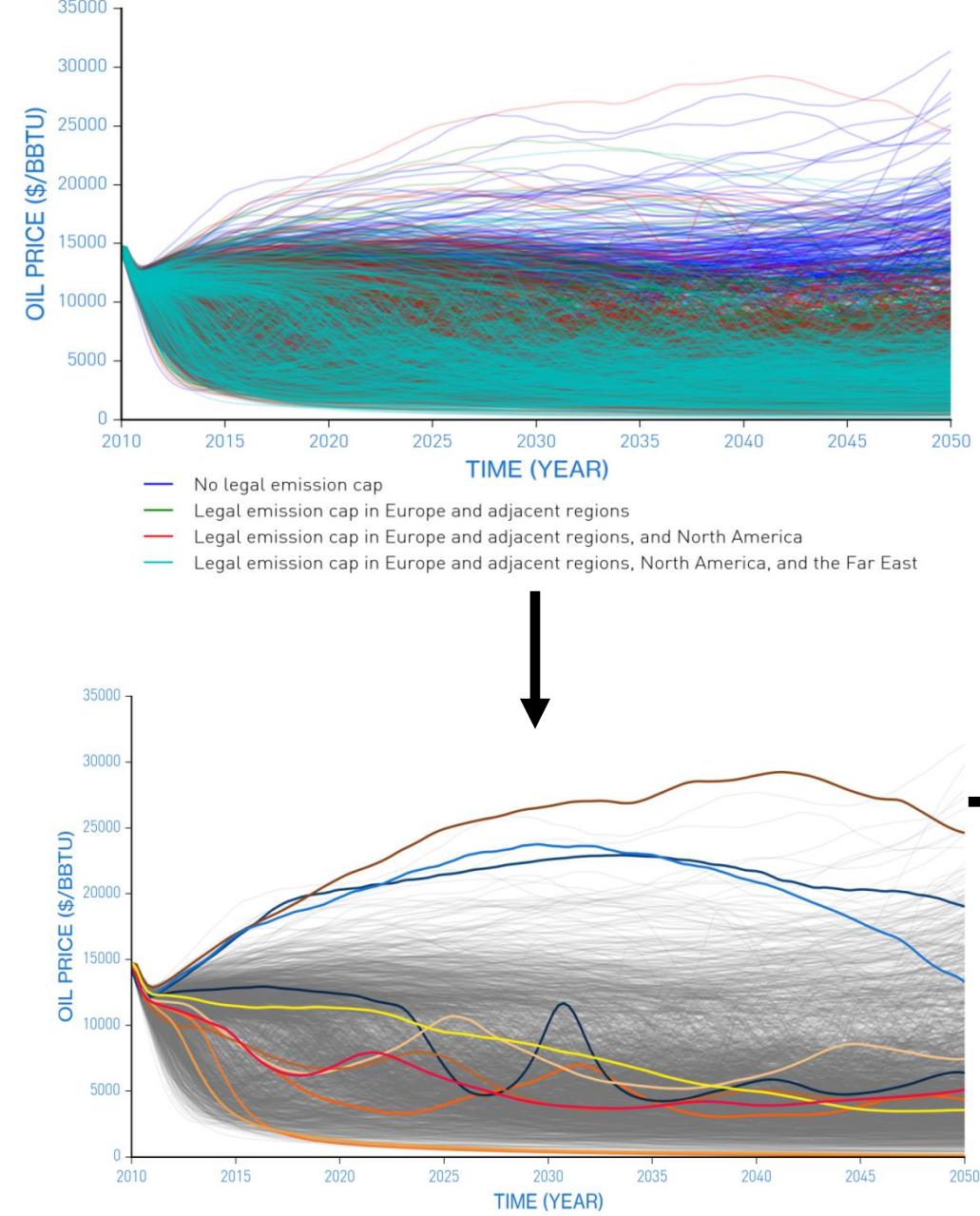
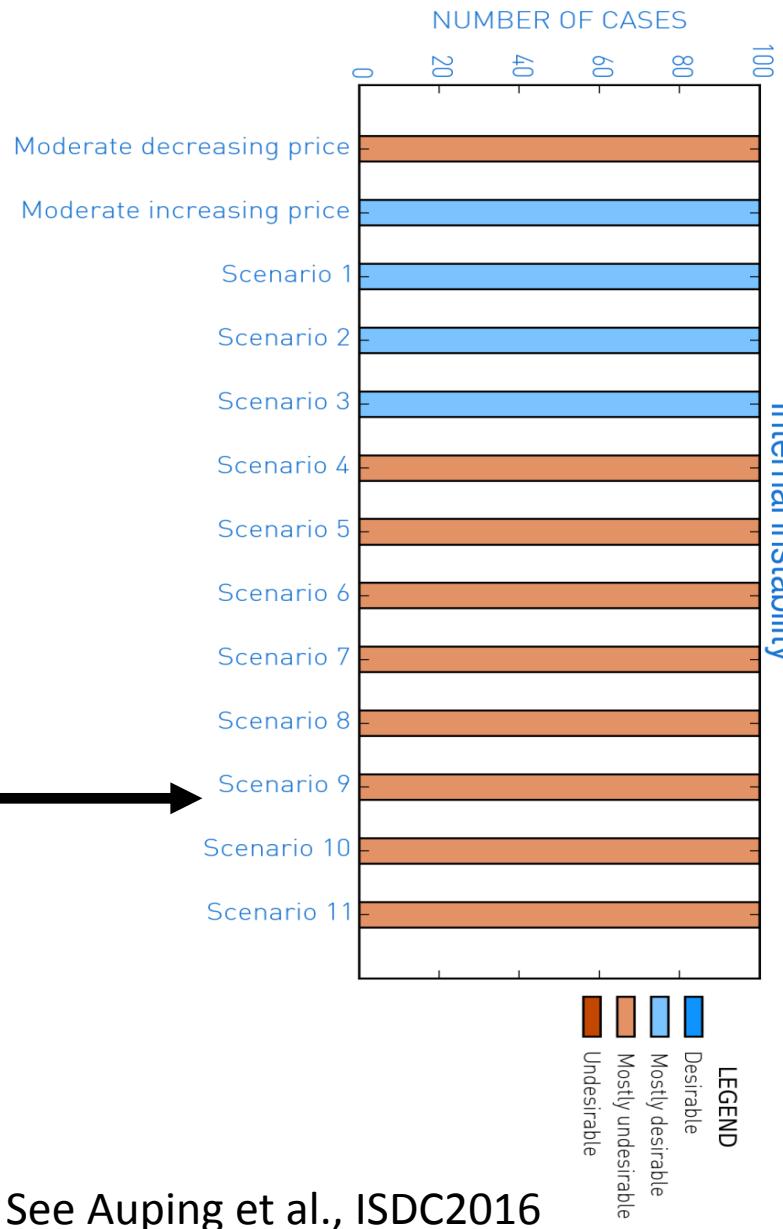
Behaviour Space Sampling + Behaviour-based classification/clustering + Input space identification & exemplar selection (Pruyt Islam 2016 SDR)

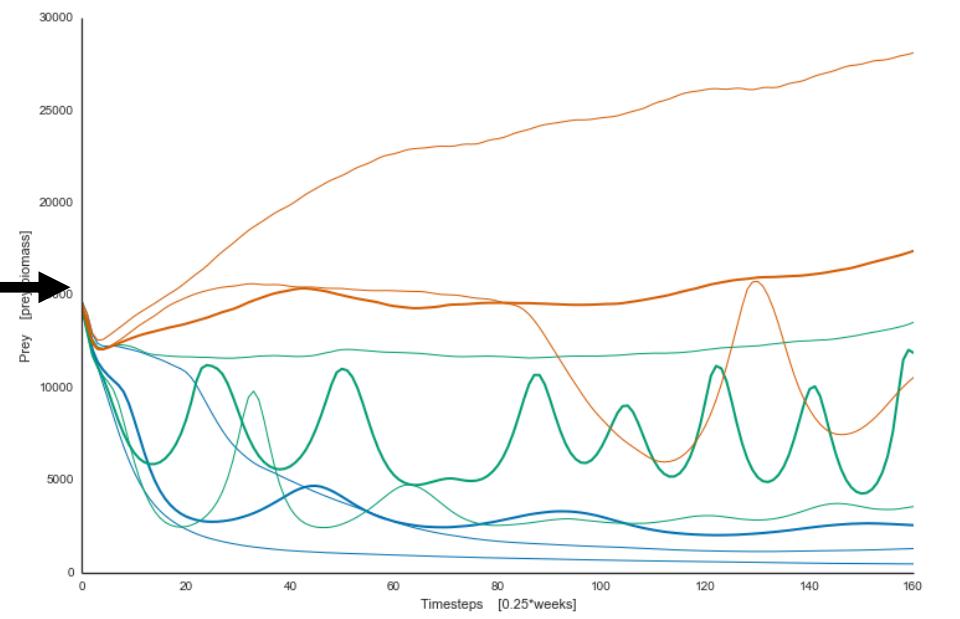
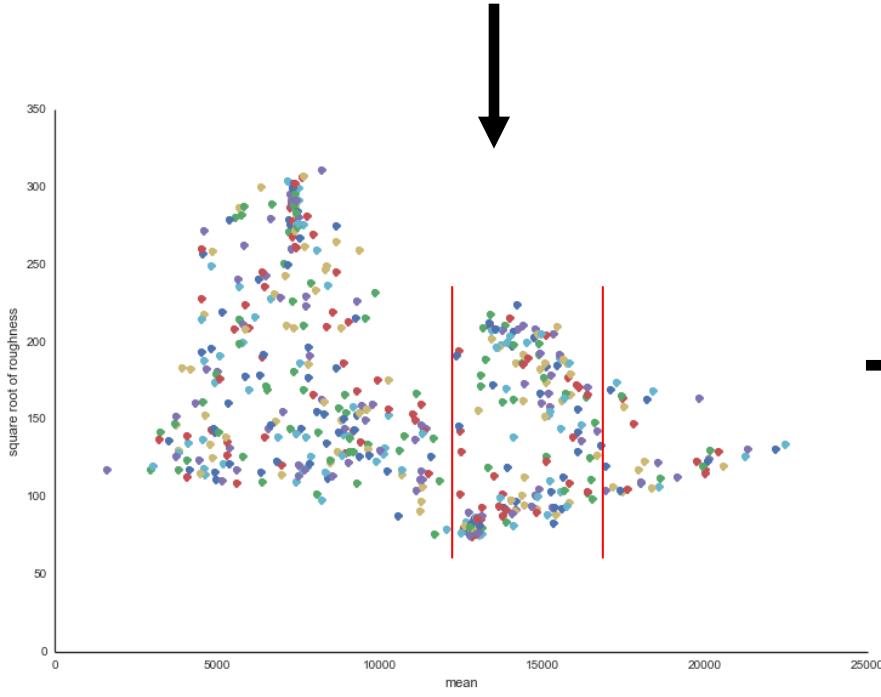
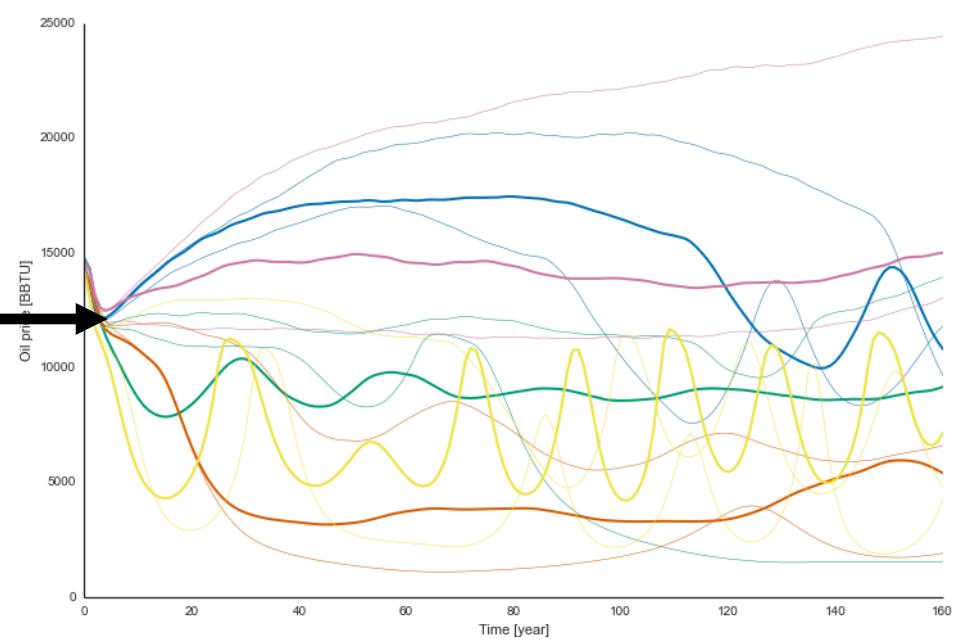
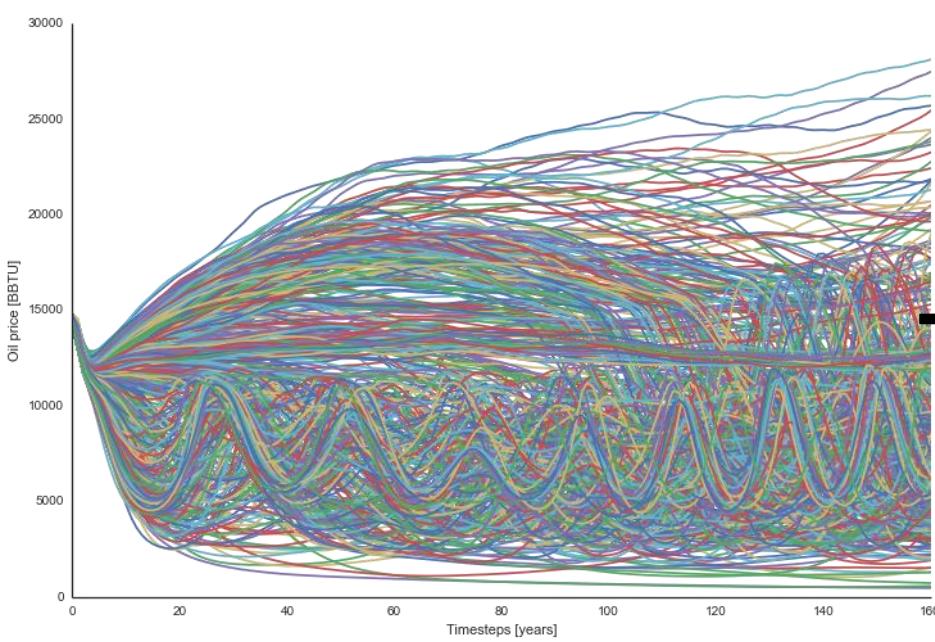


Example 1: Subset of Scenarios/Runs from Model 1 to Model 2, under Deep U.



SCENARIO EFFECTS FOR ALGERIA: Internal instability





Example 2: Burnout

- [BSSv02b_Burnout01short_v01.ipynb](#)
- adp_july.py
- adp_anlys.py

3) PySD: From Vensim models to Python +..., Simulation & Analysis in Python

- Vensim: 2 make models, but fitting/plotting/...
- pySD: wrap Vensim model to python version
 - All python libraries become applicable
 - No need for Vensim DSS
 - No need for expensive computer: Raspberry Pi is enough
- Download pySD frm James Houghton's Git Rep
- Source>analyses>
 - Getting_started
 - Fitting
 - ...
 - Geo

4. From Aggregated to Disaggregated and Hybrid Modelling & Simulation

From 1 (subscripted) model as of old to:

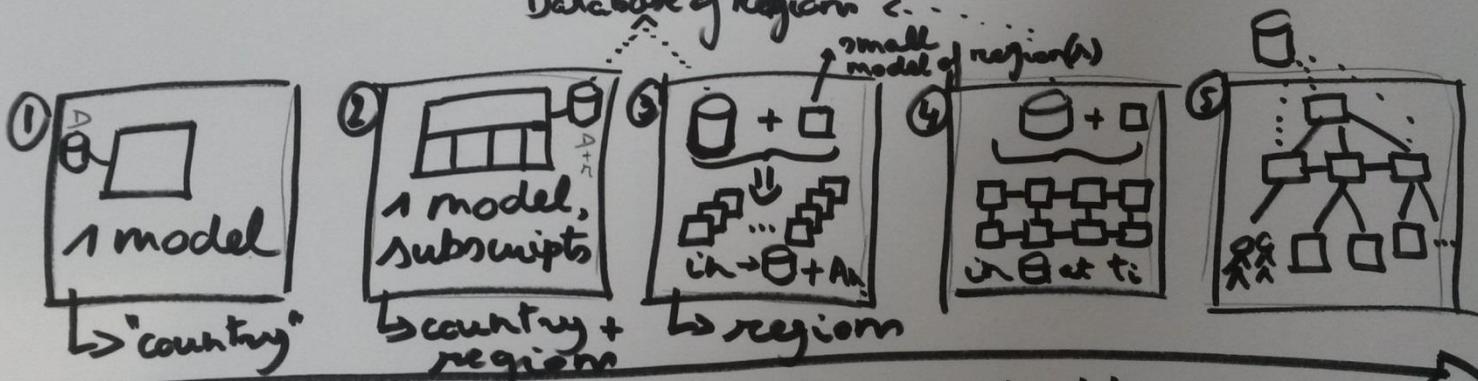
1. Simultaneous simulation of alternative models uU
See: Pruyt & Kwakkel 2014 SDR; Pruyt et al 2015 SRBS;...
2. Subset of scen. runs from 1st model into 2nd model uU
[Auping et al. ISDC2016, (JEPO)], (Pruyt Islam, '16 SDR)
3. Database + model(s) parameterized for many regions
Without interaction btwn regions [see food study]
4. Database + parameterized model + interaction regions
[see Schwarz Pruyt ISDC2016]
5. Layered geo-spatial model components x SD/AB/DES/
[see Hesan Pruyt ISDC2016]

Or with a drawing

DIFFERENT ARCHITECTURES FOR
DEALING WITH REGIONS & COUNTRY

Database of regions c...

small
model of region(s)

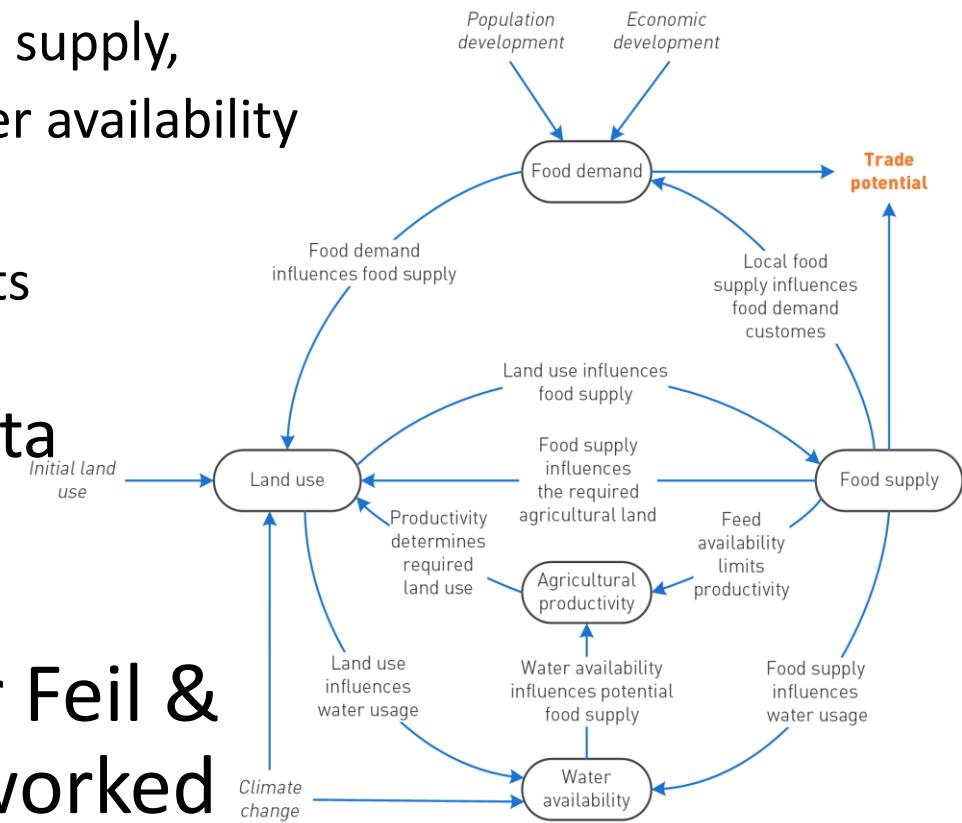


more innovative, better f levels
+ linking between regions & levels.

.....
Difficult + β-tality

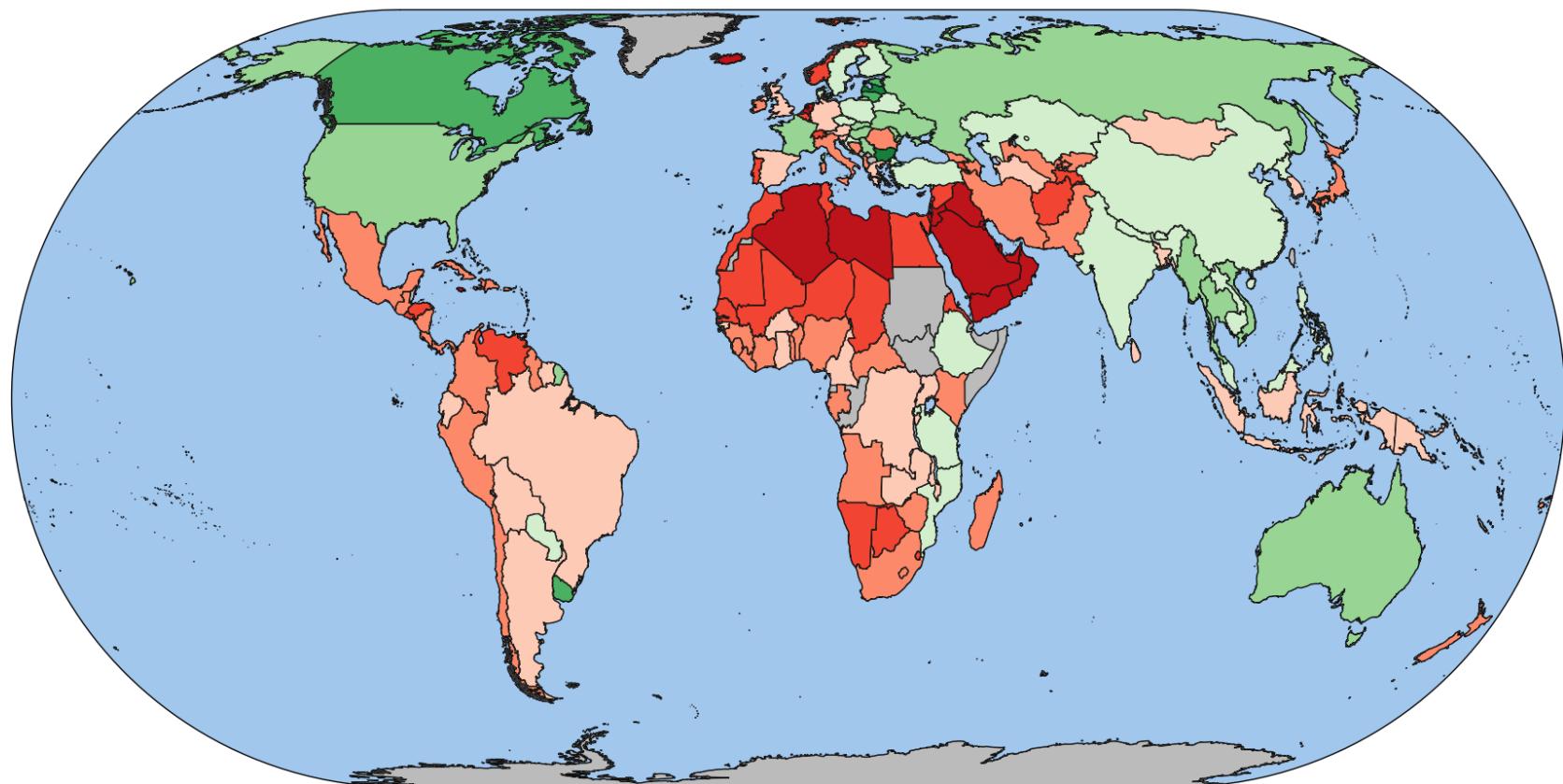
(3) Database + model parameterized for many regions (wo interaction)

- 1 model for food supply & demand per country
 - Five sub-models for
 - Food demand, Land use, Food supply,
 - Agricultural productivity, Water availability
 - 11 different food types, focus on 3: cereals, meat, & roots and tubers)
 - Large amount of input data
 - Data from FAOSTAT, UN pop. dept., IPCC,...
 - For questions: Wouter Feil & Laurens de Kok have worked on this project (well done guys!)



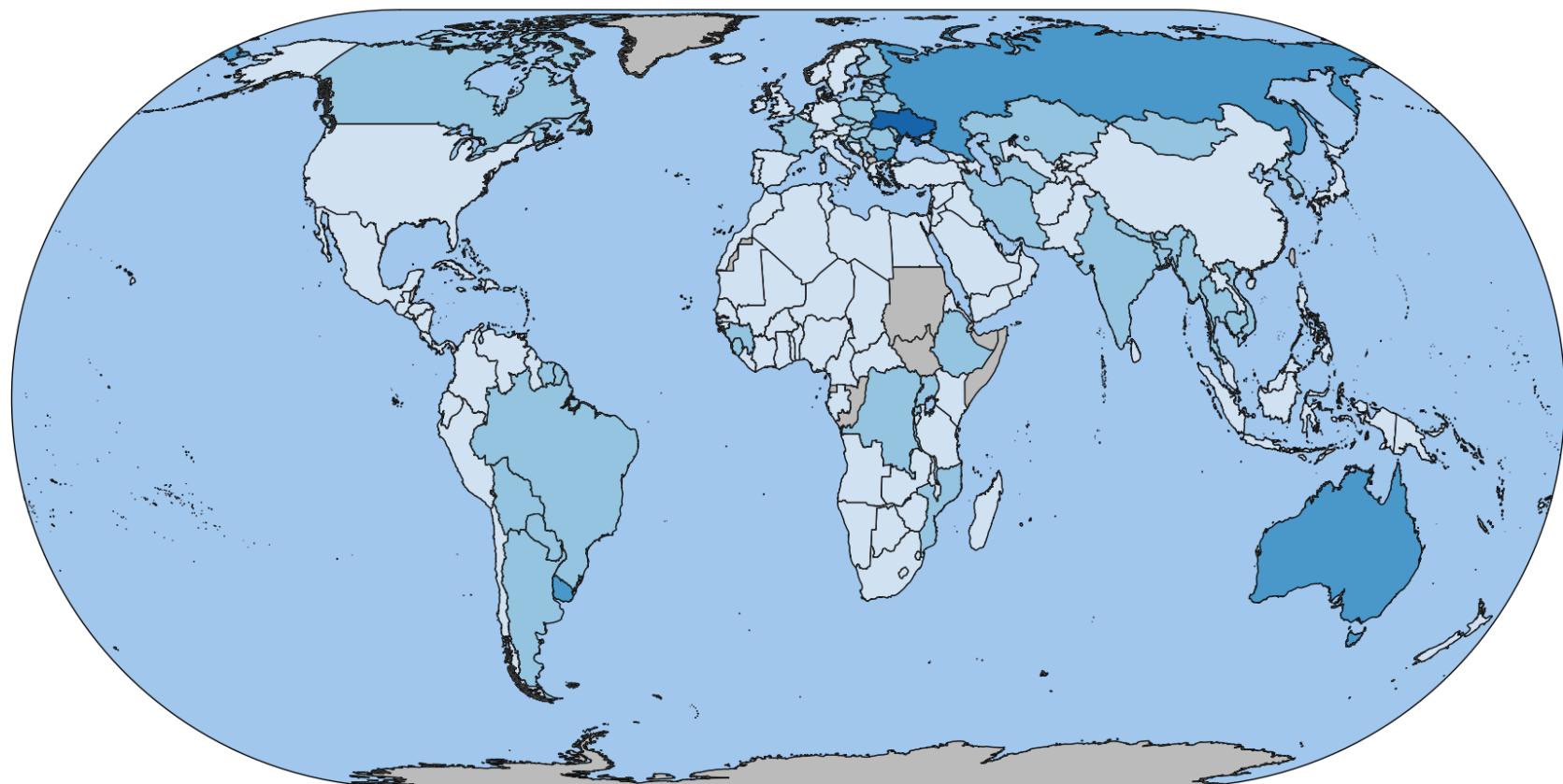
Descriptive statistics

RELATIVE TRADE POTENTIAL CEREALS 2031



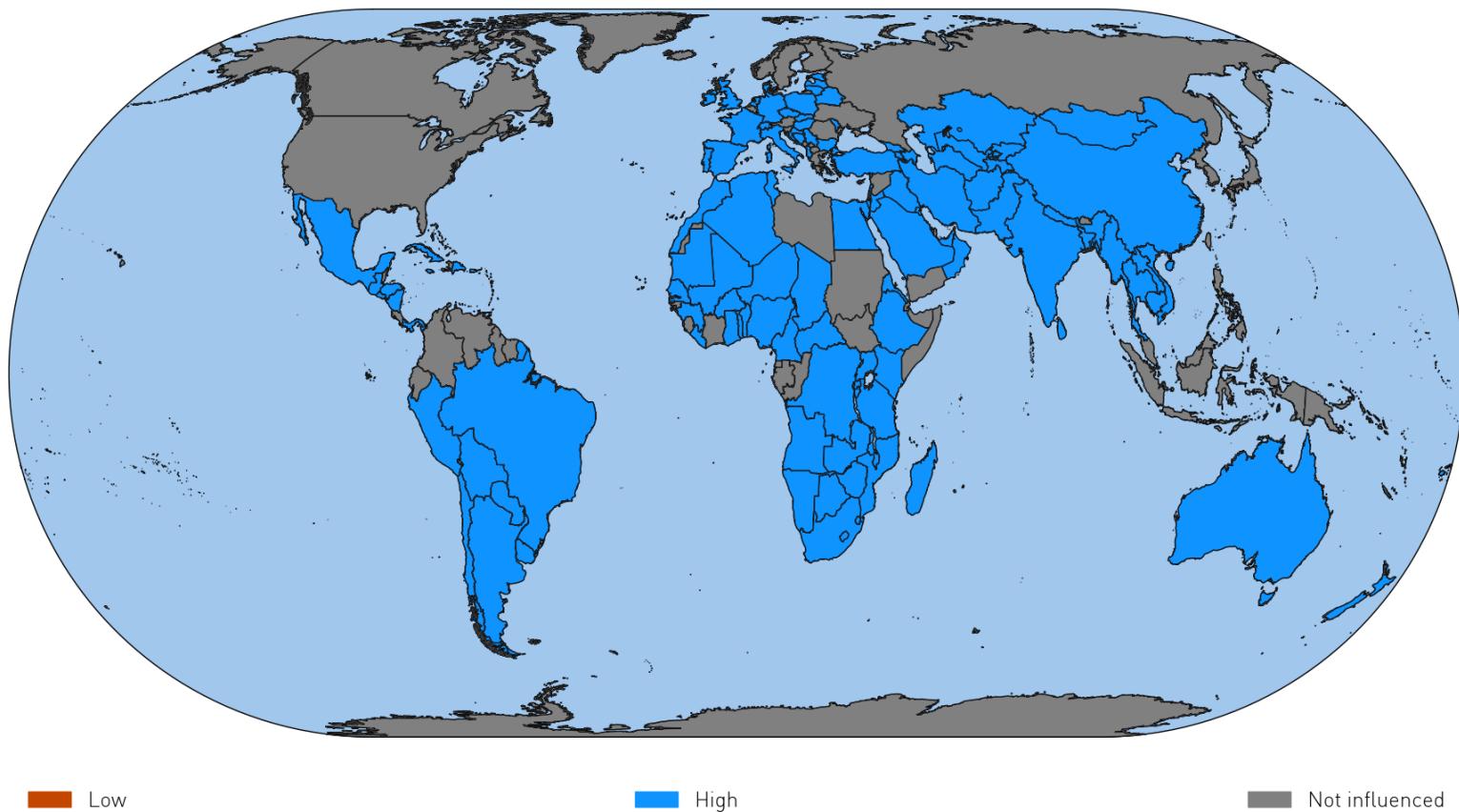
Uncertainty in outcomes

UNCERTAINTY IN RELATIVE TRADE POTENTIAL CEREALS 2031



PRIM: LAND DEGRADATION

NATURAL LAND DEGRADATION RESTORATION TIME AFFECTING
RELATIVE TRADE POTENTIAL CEREALS IN 2031



Low

High

Not influenced

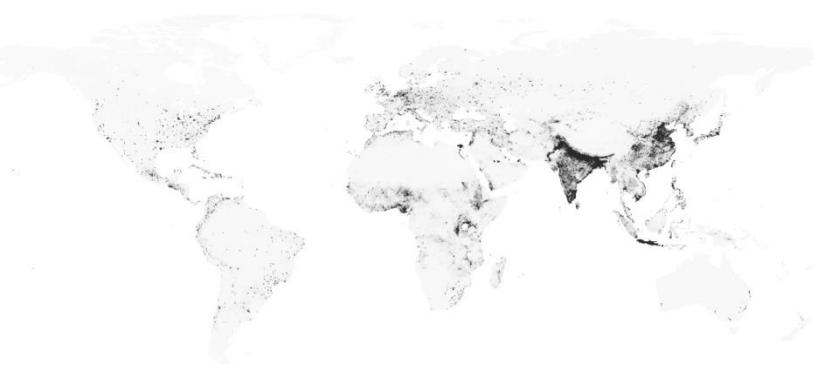
(4) MSc thesis of Philipp Schwarz

Collecting & using geo-referenced big data

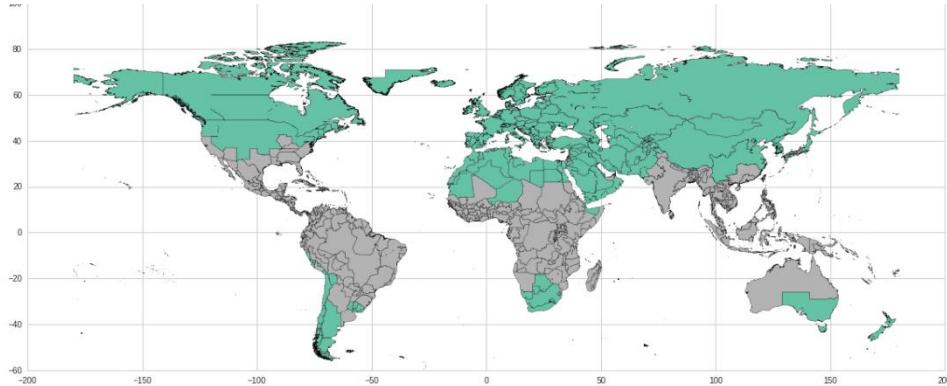
Databases provided by Intergovernmental organization



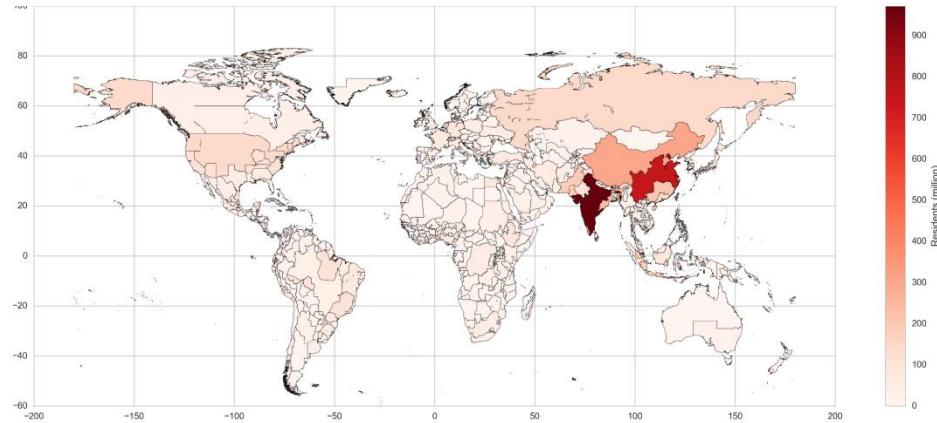
High-resolution gridded maps



Data requirements on higher resolution than national-level

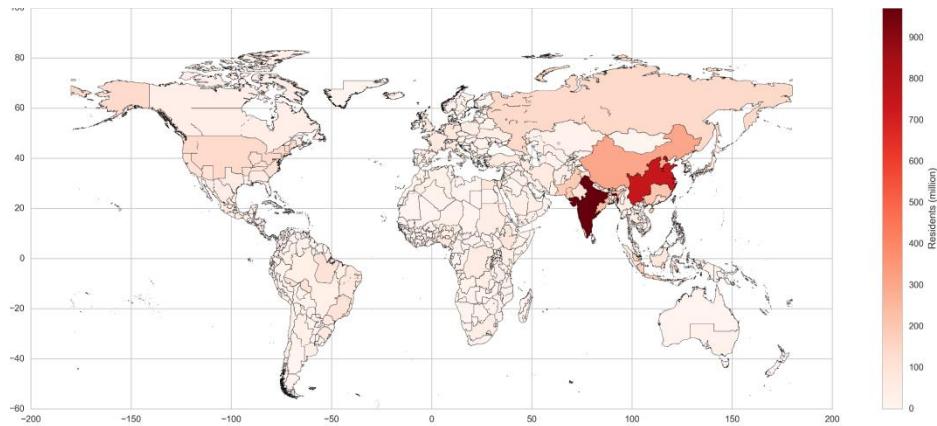


Aggregate Data to level of modelling



CIESIN. 2015. "Gridded Population of the World, Version 4 (GPWv4): Population Count Adjusted to Match 2015 Revision of UN WPP Country Totals." Center for International Earth Science Information Network - CIESIN - Columbia University. NASA Socioeconomic Data and Applications Center (SEDAC). <http://dx.doi.org/10.7927/H4PR7SX1>.

Data on the level of modelling



Model of global interconnectedness

Gravity Model of monthly global passenger air travel.
Based on socio-econ. of two places and their distance

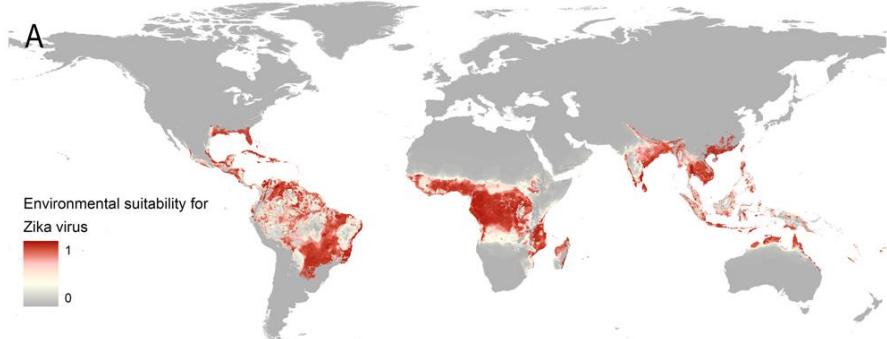


Mao, L., Wu, X., Huang, Z., Tatem, A.J., 2015. Modeling monthly flows of global air travel passengers: An open-access data resource. *J. Transp. Geogr.* 48, 52–60. doi:10.1016/j.jtrangeo.2015.08.017

+ data from other modeling studies:

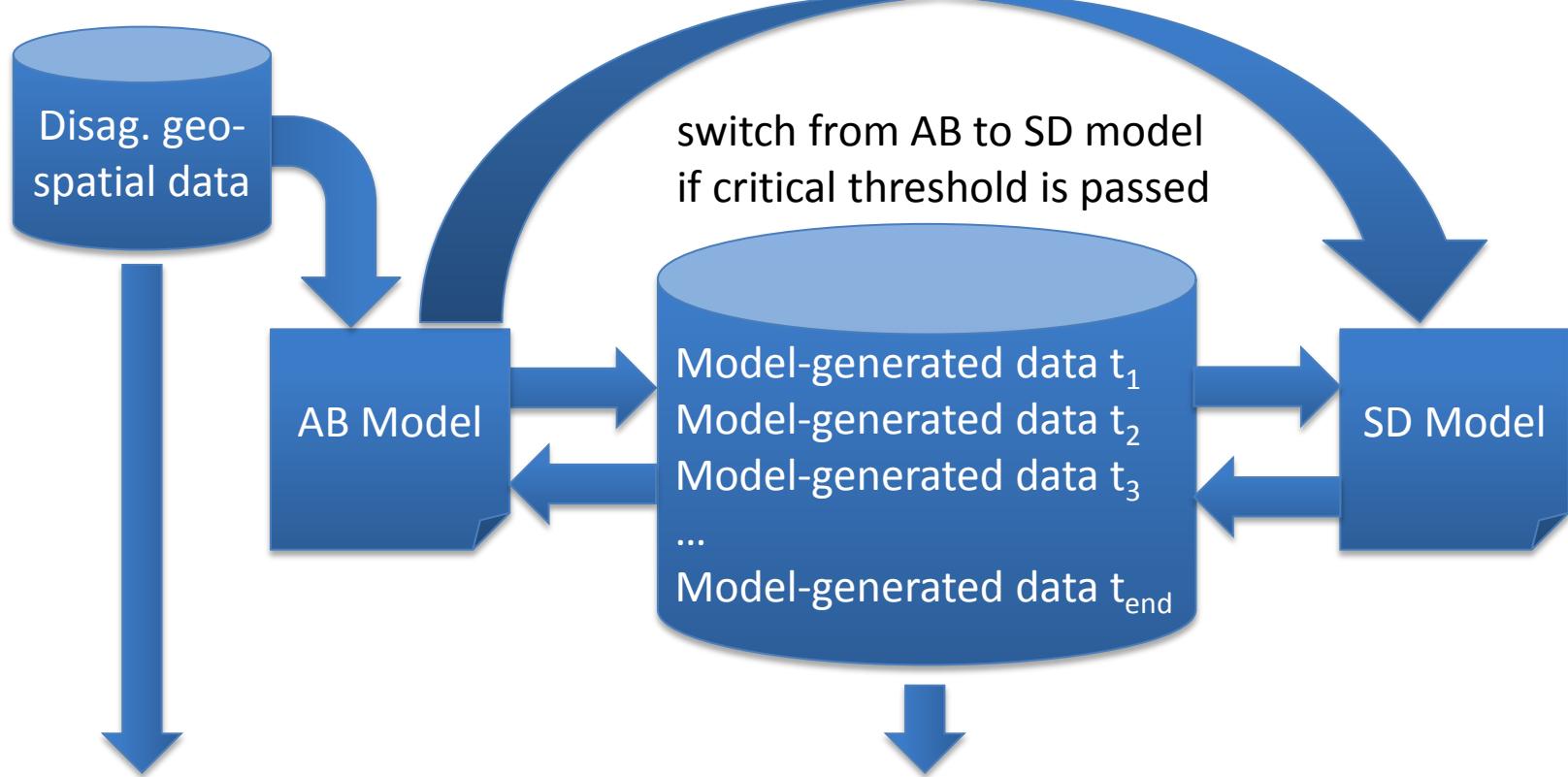
Species Distribution GIS Model

Complex multivariate spatial relation between observations of species occurrence and environmental conditions



Messina, J. P. et al. (2016). Mapping global environmental suitability for Zika virus. *eLife* 5, e15272 [Data set] Requested from authors. <https://elifesciences.org/content/5/e15272>

Under deep uncertainty
with EMA workbench



Geo-spatial animation of runs: Cumulative Zika cases % of population

Setting up database-driven simulations for a lower level of aggregation

- Raspberry Pi
- Philipp Schwarz's Master Thesis (2016)
- Python code:
 - For making databases
 - Data handling and stacking
 - Switching between ABM and SD
 - Simulation under deep uncertainty
 - Analysis of results

(5) Hierarchical geo-spat. SD-ABM-DES M&S in Anysim [beta-version]

HybSim -- 4.Sim

File Edit Help View

Elements Palette

- Continuous SD
 - Stock
 - Flow
 - Converter
- + Discrete SD
- ABM
 - Agent
 - Space
 - Link
- Sub_system
- Plotter

Canvas X 2.Sim X 4.Sim X

Lng: -106.87500, Lat: 34.59704, zoomLevel: 1

Fit bound Reserve

Leaflet | Map data (c) OpenStreetMap contributors

Go to Panel

Python Shell

```
100
Traceback (most recent call last):
  File "mapfile.py", line 242, in addElement
AttributeError: 'float' object has no attribute 'iteritems'
100
Traceback (most recent call last):
  File "mapfile.py", line 242, in addElement
AttributeError: 'float' object has no attribute 'iteritems'
this is the 9459
9459
this is the new way of 5435
this is the new way of 1696
this is the new way of 5435
```

Setting Palette

General setting Code

Elements

Name3
Name0
Name0

```
self.Name1 = self.Name0/10
```

apply

0%

The screenshot shows the HybSim software interface. The main canvas displays a simulation model with nodes labeled Name0, Name1, Name2, Name3, Name4, Name5, and Name8, connected by various links. To the right of the canvas is a world map with a blue marker indicating a specific location. Below the canvas is a Python Shell window showing several error messages related to 'addElement' and 'iteritems' attributes. On the left, there's an Elements Palette with categories like Continuous SD, Discrete SD, ABM, and Sub_system. At the bottom, there's a Setting Palette with tabs for General setting and Code, and a list of elements (Name3, Name0, Name0). A code editor in the Setting Palette contains the line `self.Name1 = self.Name0/10`. The status bar at the bottom right shows '0%'.

(6) The Full Monty [Python]?

- Complex models: Everything in python?
- See Bas Keijser's Master Thesis:
 - SD x ABM x Oligopolistic market handling
 - Statistical analyses
- “Best graduate” competition on Nov 17!

Conclusions

- Python greatly enhances modelling and simulation!
 - ⇒ Database structures (into and out of)
 - ⇒ Scripting to force programs what to do
 - ⇒ Analytical tech (eg BS Sampling, Clustering, ML,...)
 - ⇒ Visualization of the results: on spot plotting on a map
- Many recent and near-future innovations [needed]
 - Modelling under deep uncertainty + analysis
 - Hybrid and systems-of-systems modelling
 - Database-driven / geo-spatial simulation
 - Visualizations and animations: “dynamic & on the map”
 - + under uncertainty and on the spot in workshop settings