spm9155 2014 2015

**Assignment 4: Exploratory Model Analysis**

Matthijs Brouns 1505157

Hélène van Heijningen 4076850

*Question 1*

*Analyze the model you have built carefully or choose to use the 2-area epidemic model supplied. What are the main uncertain factors (7 to 15 factors) in this model? What are justifiable ranges/values for these factors? Use the sensitivity analysis tools in Vensim to explore the behavior of your model over these ranges. Analyze the results and discuss the policy implications of your findings.*

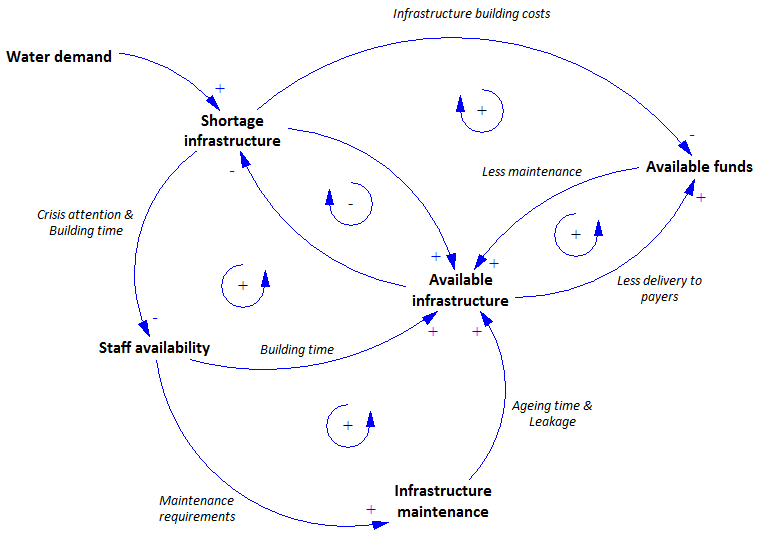
*In your write up of the assignment, clarify which uncertain factors you have chosen and why, what ranges you have used and why, motivate your sampling technique, and your analysis of the results. The write up should be such that your analysis can be replicated by me, in theory.*

*Tips:*

* *You can export the data from the sensitivity analysis to tab separated or comma separated files for further analysis. Read the Vensim manual for details*
* *Many of the discussed techniques (PRIM, CART) are available in open source implementations. You could even use the code of the exploratory modeling workbench if desired.*
* *Ranges should be plausible and grounded in information.*
* *Reference the theory material.*

Our approach for this assignment is as following:

* Start broad: x uncertain factors
* Determine ranges
* multivariate analysis
* Run sensitivity analysis
* Classify model behaviors
* Determine combination of factors which lead to desired behavior (narrow down)
* Draft policy,
* Put policy in model to test



Factors:

#### Staff hired / time period

0 – 30

UNIFORM

#### Staff service time

5 – 25

UNIFORM

#### Attention rate for primary / secondary staff activities

Divide staff by ratio between maintenance and planning+construction

Do all maintenance tasks first

Do all planning+construction tasks first

VECTOR

#### Attention rate between planning and construction tasks

0 planning 1 construction -> 1 planning, 0 construction

UNIFORM

#### Staff required for tankers

0 – 1

VECTOR

#### Annual financial bail-out

0 – 200

UNIFORM

#### Percentage of income used for maintenance

0.2 – 0.6

UNIFORM

#### Infrastructure aging time

20 – 40

UNIFORM

#### Max infrastructure pushing

0 – 0.5

UNIFORM

#### Target percentage infrastructure maintenance

5 - 25

UNIFORM

Number of runs: 2000

Say we want at least 20 levels per variable

20 ^10 -> 1,024 e13 runs for full factorial -> completely unrealistic

Latin Hypercube needs less samples since more samples are not needed for more input dimensions. Max number of combinations for a cube can be determined using following formula with M divisions and N dimensions.

\left(\prod_{n=0}^{M-1} (M-n)\right)^{N-1} = (M!)^{N-1}

This still results in 1e7 possibilities. To keep analysis feasible we run 10.00 runs.

## Model behavior classification

How do we define a crisis: maybe simply whether there is infrastructure shortage?

#### Classify behavior on type of crisis which occurs

Problems:

Crises happen in pairs and follow up on eachother soon

No distinguishable difference between runs where crisis is never solved or crisis is solved

#### Classify behavior on timeslot where crisis first occurs

Problem: no feedback on whether crisis gets solved

#### Classify behavior on number of years of crisis

No feedback on type of crisis, however, not a problem because crises happen together mostly

Problem: No idea when crisis occurs, only how long the crisis occurs

No idea on severity of the crisis

#### Classify behavior on number of years of crisis + timeslot where crisis occurs

Solves problem with previous

However: much more classification categories (previous category \* number of timeslots) -> decreased interpretability and increased curse of dimensionality (requires a lot more runs to statistically make sense)

#### Classify behavior on infrastructure shortage \* period (area under the curve)

Gives very accurate, high resolution result

Lower interpretability

To increase interpretability we can divide the infrastructure shortage (in ML / year) by the average household water consumption to get household\*years of shortage which is a more relatable measure.

Keep in mind that this means that a low number of households with a long period of shortage is classified the same as many households with short shortages. This may or may not be desirable for all purposes but we think it’s fine for this.

