Agent-based model of Mexico City

ODD Protocol

Andres Baeza

Postdoctoral Researcher

School of Sustainability

Arizona State University

Table of Contents

[Introduction 4](#_Toc497308407)

[Agents, actions, and scale (with notation) 5](#_Toc497308408)

[Process overview and scheduling 6](#_Toc497308409)

[Design concepts 7](#_Toc497308410)

[Basic principles 7](#_Toc497308411)

[Emergence 7](#_Toc497308412)

[Observation 7](#_Toc497308413)

[Adaptation 7](#_Toc497308414)

[Interactions 8](#_Toc497308415)

[Stochasticity 8](#_Toc497308416)

[Details 8](#_Toc497308417)

[Initialization 8](#_Toc497308418)

[Input data 9](#_Toc497308419)

[Sub-models and components 9](#_Toc497308420)

[Water authority decision-making procedures 9](#_Toc497308421)

[Priority assessment 9](#_Toc497308422)

[Site selection 10](#_Toc497308423)

[Normalization functions 10](#_Toc497308424)

[Exposure models 11](#_Toc497308425)

[Exposure to water supply disruption 12](#_Toc497308426)

[Exposure to flooding 13](#_Toc497308427)

[Exposure to gastrointestinal diseases 13](#_Toc497308428)

[Water authority actions and changes to census block attributes 14](#_Toc497308429)

[Maintenance 14](#_Toc497308430)

[New infrastructure 14](#_Toc497308431)

[Water distribution 15](#_Toc497308432)

[Water extraction 15](#_Toc497308433)

[Resident actions 16](#_Toc497308434)

[House modification 16](#_Toc497308435)

[Protests 16](#_Toc497308436)

[Criteria for decisions - Water authority 16](#_Toc497308437)

[Capacity 16](#_Toc497308438)

[Failures 16](#_Toc497308439)

[Lack of infrastructure 17](#_Toc497308440)

[Amount 17](#_Toc497308441)

[Social pressure 18](#_Toc497308442)

[Age of infrastructure 18](#_Toc497308443)

[Criteria for decisions - Residents 18](#_Toc497308444)

[Urbanization 18](#_Toc497308445)

[Clogged drainage 19](#_Toc497308446)

[Insufficient infrastructure 19](#_Toc497308447)

[Water scarcity 19](#_Toc497308448)

[Flooding 19](#_Toc497308449)

[Policy scenarios 20](#_Toc497308450)

[Indicators 20](#_Toc497308451)

[City average age of infrastructure system 20](#_Toc497308452)

C[ity average exposure to flooding and scarcity . 20](#_Toc497308453)

[City average level of socio-political pressure 21](#_Toc497308454)

[Vulnerability index 21](#_Toc497308455)

[Inequality in exposure 22](#_Toc497308456)

[Sensitivity to policy changes 22](#_Toc497308457)

[Appendix 22](#_Toc497308458)

[Extensions 22](#_Toc497308459)

[Generating classes of census blocks using a k-mean clustering classification 22](#_Toc497308460)

[GIS pre-processing information 22](#_Toc497308461)

[Instructions for creating new types of agents and actions 22](#_Toc497308462)

[Tables 24](#_Toc497308463)

# Introduction

This document describes the implementation of the agent-based model of the MEGADAPT project (adaptation in a megacity). The full MEGADAPT model simulates the coupling between biophysical processes and the decisions of residents and the water authority of Mexico City. The aim of the model is to investigate the consequences of this coupling for the spatial distribution of socio-hydrological vulnerability in Mexico City. The agent-based model presented here simulates decisions regarding investments in infrastructure by water authority agents, and the decision of resident to adaptation and demanding response to the authority agents. The decision-making process is done by evaluating the condition of the landscape across the census blocks. These investments in turn trigger actions in selected census blocks. These actions then influence the conditions of the socio-political and physical attributes in the landscape, which in turn modify the biophysical conditions, affecting risk of infrastructure hazards and exposure to flooding, disruption of water supply, and gastrointestinal diseases. The current version of the model incorporates stochastic simulations of weekly water supply, annual flooding events, and annual cases of gastrointestinal incidences. The model was constructed using available empirical observations of these hazard events With the model we aim to investigate the dilemmas and spatial patterns of vulnerability to infrastructure hazards that emerge due to the actions and decision-making processes of the agents.

The decision-making process of the water authorities involves the identification and selection of census blocks that require investment. This prioritization of census blocks for investment is rooted in computing a multi-criteria metric in which the water authority evaluates the landscape based on a set of criteria that are weighted differently based on the priorities of the actors. To compute the multi-criteria metric the model relies heavily on empirical data derived from mental models and translated into a multi-criteria decision framework, which are inputs to the model. The decision-making process of the residents is also done via a multi-criteria decision metric, but in this case the evaluation is only of their local census block and the action taken modifies only those local attributes. Figure 1 shows an overview of the basic linkages between the actors and the procedures included in the current version of the model.

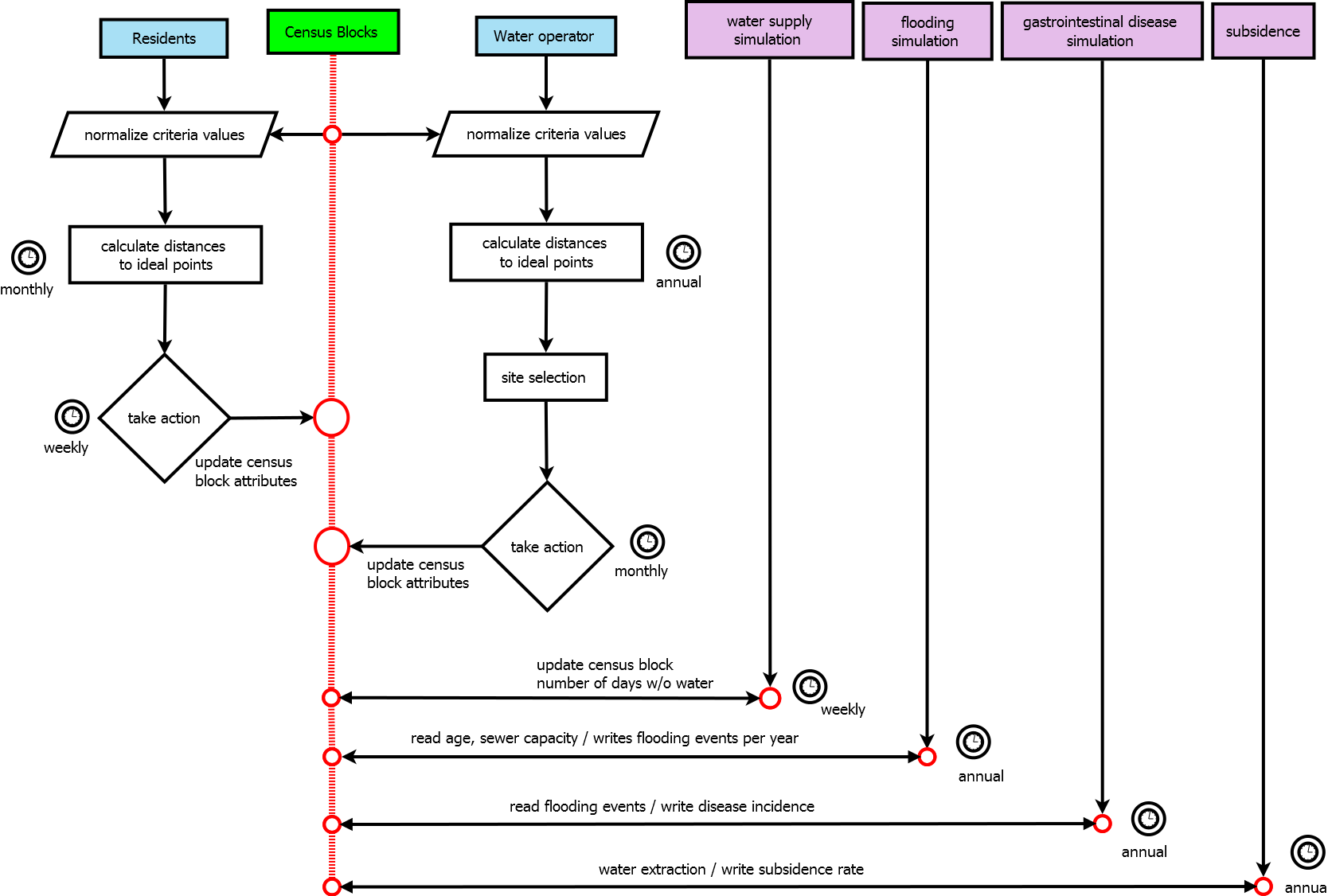


Figure 1: An overview of the processes and sub-models included in the current agent-based model

The current version of the model was built in NetLogo 6.2.1. The advantage of NetLogo is its versatility in that it permits agents to be developed as objects, allowing modelers to define actors, their actions, and infrastructure systems as separate computational entities. This provide modularity to develop new procedures, including new agents, actions, infrastructure systems and policy scenarios.

# Agents, actions, and scale (with notation)

Three types of agents are incorporated in the model: Water operators, residents, and infrastructure systems. Residents are located in census blocks. Each census block is indexed using the symbol , and the total number of census blocks is indicated by , such that . Each census block contains a set of attributes, . Infrastructure systems are also considered agents when they are defined as objects, with properties and methods similar to those of other agents.

Water operators make investments in different infrastructure systems, , in select census blocks. These investments are related to the set of actions : 1) repairs and maintenance,, 2) construction of new infrastructure, , 3) distribution of water, , and 4) extraction of water, . Residents can take different actions as well. These actions are: 1) house modification, , 2) collective actions, , and 3) protesting, .

The decisions made by the water operators to invest and by residents to modify the local environment are evaluated based on a multi-criteria metric. This metric is calculated based on a set of criteria, and each criterion is weighted with values specifically related to the actor and the actions they take. This information and the actual actions of the agents is derived from empirical data from the actors’ mental models and additional information collected during focus groups, interviews, and workshops with the managers and operators of the water system and residents. This information was pre-processed and transformed into an analytical network process (ANP) to derive the criteria weights and the weighting of the actions.

Infrastructure systems are labeled with the symbol . The current version of the model includes the following as infrastructure systems: 1) the wells for water extraction, , 2) the system of pipes for potable water distribution, , 3) the sewer system, , and 4) the system for water distribution (trucks), . In the current version the system of pipes and sewer system are represented in each census block as a set of attributes, such as coverage, average age, condition, and capacity. The system of wells, on the other hand, is represented as a set of agents with particular attributes. The model associates the set of wells with its corresponding census block based on the location of both the wells and the census blocks.

The model also contains procedures for representing the stochastic production of hazard events related to the risk of flooding, ; disruptions of water supply, ; and gastro intestinal disease incidences, . These models are defined based on empirical information about past events in Mexico City. In all these submodels, the frequency of the hazards is assumed to be related to failures in the provision of services by a particular infrastructure system. These failures are in turn associated with the local conditions of the system, the local biophysical conditions in the census blocks and, most importantly, the actions of the water operators and residents.

The actions the agents take modify the attributes of the census blocks, which in turn will influence the stochastic production of hazards. These changes in events then inform subsequent decisions made by the actors. Therefore, the model simulates the feedback loop between actors’ decisions and biophysical changes in a dynamic and spatially-explicit platform.

The spatial resolution of the model is the census block, represented in the model by census blocks of Mexico City and the state of Mexico. The spatial scale of the model engulfs all the census blocks in these two states. Each census block is also identified by the municipality to which it belongs.

The temporal resolution of the model is one week. Depending on the actions and agents, different periods in a year can represent different evaluations of priorities. We call this period a decision cycle, denoted by *T*. A year in the model represents a decision cycle , while a month would be a decision cycle . For the water operators making decisions about maintenance and new infrastructure actions, for each decision cycle , the agent re-evaluates the census blocks selected for investment including maintenance, new infrastructure, and extraction. For water distribution actions, the decision cycle is . For the residents the decision cycle is represented as as well. Actions are independent from the evaluation in the sense that they can occur in different time periods. These are cycles of actions, . The actions of the water operators occur in monthly cycles for maintenance, extraction, and new infrastructure, and weekly for water distribution. For residents, cycles of action are weekly for protests, and monthly for water storage and house modification.

# Process overview and scheduling

The model is currently composed of three files. A setup file (“setup.nls”), a value-function file (“value-functions.nls”), and the code file (“ABM-Empirical-MexicoCity\_6.nlgo”). The setup file defines and set values for global variables, and for the agents and their attributes. It also contain procedures for reading the information that will define the decision-making processes of the agents, namely their actions, the criteria set, criteria weights, and action weights. Finally, in setup, the GIS layers are loaded from which the value census block attributes are defined.

The value-function file contains the procedures for updating the information needed for the actors to calculate the multi-criteria decision metric. Specifically, these procedures update the standardized values of the attributes of census blocks that are criteria for decisions, and they also calculate the distance metric.

The code file contains a procedure called “GO”. Procedure “GO” involves the suitability assessment and site selection procedures, which define the census blocks that will be selected for investment by the water authorities. Once the census blocks have been selected, the model invokes the actions of the different agents, which will in turn modify the attributes of the landscape. Inside the code file, there are also procedures to calculate indicators of performance at the scales of the census block, municipality, and city.

To run the a single simulation of the model, the observer needs to invoke the procedure “GO”. This procedure will trigger a set of low-level processes related to the stochastic simulation of water allocation, flooding and health models, the decision-making processes of the actors, the actions of the actors, and the changes to the attributes of the census blocks. The decision-making process of the water authorities agents includes procedures for site selection and suitability of census blocks to investments. In the case of the resident agents, they invoke a selection of actions. Within each decision cycle, the actions will modify the attributes of the landscape associated with each action in each census block. Details about the processes are provided in the “sub-models” section below.

# Design concepts

## Basic principles

The model aims to represent the decisions of important actors in the Mexico City water management system in terms of investments in infrastructure systems using a geo-spatial and dynamic multi-agent system. The decision-making process of the actors involves a multi-criteria decision procedure in which the actors evaluate a set of criteria they consider to be important when making decisions (criteria), assigning them different levels of importance (criteria weights). In the model each criterion is associated with an attribute of the landscape. Therefore, the agents evaluate the attributes in the landscape when making decisions related to investment, in the case of the water authority, and adaptation, in the case of residents. The attributes in the landscape change over time according to the stochastic variability generated by the statistical processes of hazard production and, most importantly, due to the changes in the attributes caused by the actors’ actions.

Emergence

From the interaction between the agents’ actions and modifications to the attributes of the census blocks, we expect an emergence of spatial patterns in investments that in turn should influence the spatial pattern of vulnerability of census blocks to infrastructure-related hazards.

### Observation

The agents observe the attributes of the landscape that are considered to be criteria in the decision-making process. The observer can evaluate the vulnerability of the city using different indicators of performance and a metric of vulnerability. We compute the average age of the infrastructure, the numbers of day in a year without water, and the number of actions taken by the water operators. These observations are computed by census block and aggregated by municipality.

Adaptation

Residents in each census block can modify their local environment to reduce their sensitivity to exposure to flood and scarcity hazards. Thus the action “modification of house” reduces the sensitivity of the census block to flooding, whereas the action “water storage” reduces the sensitivity of a census block to suffering exposure to water scarcity (see Details). These processes over time can be considered as adaptations to the local environment.

Interactions

The water operator and resident agents interact with the census blocks and the infrastructure systems by modifying their attributes. The attributes of the census blocks in turn influence the risk of infrastructure failures, which will influence subsequent decisions made by water operators and residents. There is also interaction among actors that influence decisions. Specifically, the protest actions of the residents generate social pressure, which is one of the criteria that is used by the water authorities when engaging in the decision-making process.

Stochasticity

The model has procedures that generate stochastic events of hazards related to flooding, scarcity, and health. Currently, the model incorporates two different methods for generating flooding events. The first is a linear regression model that generates an expected number of events per year using a set of infrastructure-related, biophysical independent predictors from each census block. The second method is based on a contingency table using Bayes rules, where flood outcomes of a certain magnitude are cross-classified according to a finite number of states that are related to the condition of the infrastructure.

Water distribution by pipes is simulated using a Poisson distribution model that generates the number of days that census blocks do not received water via the system of pipes, P. The model assumes that the mean of the Poisson process is driven by the condition of the infrastructure and its propensity to fail in the delivery of water by the pipe system.

## Details

Initialization

The model is initiated by invoking the procedure “SETUP”, which triggers the procedures inside the setup.nls file. These procedures are: 1) “define-global-variables”, 2) “define-type-of-agents”, 3) “define-agents-attributes”, 4) “load-GIS-data”, 5) “define-census-blocks”, 6) “define-ActionsCriteria”, 7) “define-infrastructure-systems”, 8) read”-statistical-data”, 9) “set-initial-values-globals”, 10) “set\_maximum”, and 11) “set-matrix-contj”.

Procedure 1) defines the global variables needed (Table 1). In 2) the model sets the type of agents, or “breeds” in NetLogo terminology, and in 3) their attributes are set. Procedure 4) contains methods for reading layers of geo-spatial information using the GIS-NetLogo extension. In procedure 5) this GIS layers are used to fill up the initial condition of the attributes of the census blocks, including location. Procedure 6) contains methods for reading the data to define the actions and to define the information making up the multi-criteria decision mechanism involved in the actors’ decision-making. In 7) are the algorithms for defining the infrastructure systems, . In 8) are methods for reading data needed to simulate stochastic realizations of flooding and water supply. Procedure 9) defines the initial values for the global variables, and 10) sets the maximum values of the attributes of the census blocks. This is needed to define the range of variability of the attributes, which in turn will inform the computation of the decisions. Finally, procedure 11) defines the census blocks that are associated with each other based on the contiguity matrix. This defines the local network of interaction between census blocks that influence each other, which is needed to simulate the health model.

Input data

The setup procedure invokes a procedure that loads layers of geographic information to fill the attributes of each census block that will define the decision-making criteria. Table 1 shows the attributes of each census block. The model also loads \*.csv files to define the actions, the criteria, criteria weights, and action weights using MCDA weighed and limit matrix (Fig. 3). This information is read using .csv files in the “SETUP” procedure using the .cvs-extension capability of NetLogo. Another type of input data is the average number of days without water per municipality, which is needed to simulate the days with water provided by the system of pipes, , using a Poisson process. Finally, the model contain procedures for reading data to represent the probability of flooding using Bayesian contingency matrix analysis () and regression analysis of health.

#### Social data to define Resident’s decisions

We obtain multi-criteria models parametrized differently by 5 different neighborhoods in Mexico City that suffer from either flooding scarcity or both. These are located in three borrows at the southbound of Mexico City. Magdalena Contreras in the southwest, and Xochimilco and Iztapalapa to the east (figure ##) We assigned a multi-criteria model to the rest of the neighborhoods of Mexico City. This is done by defining a priory clusterization of the neighborhoods in five classes according to socio-economic indicators, using a principal component, k-mean approach (appendix). We then, when possible, assigned a resident multi-criteria model to all the neighborhoods from the same cluster. As more information about other neighborhoods is obtained, more resident agents can be defined.

# Sub-models and components

## Water authority decision-making procedures

### Site Suitability

An assessment of the census blocks that are prioritized for investment is obtained through multicriteria evaluation of the distance of each census block from an “ideal point,” or utopian state, defined as a set of decision-making criteria and the relative importance of each criterion for the decision makers (Bojórquez-Tapia *et al* 2011). Formally, we calculate a distance , such that:

(1)

where is the distance to the ideal point of census block with respect to decision and system ; is the criterion weight of criterion related to system , is the alternative weight of action with respect to system (Fig. 2); is the normalized value in a census block of the attribute corresponding to criterion , with respect to infrastructure system and decision ; is the departure of an alternative from the ideal point for a criterion. This variable is the standardized score , which represents a judgment about the importance of an observable stimulus (census block attribute value) in the water authority’s decision; , , , , and are indices for criteria, census blocks, action, time, and infrastructure system, respectively. Finally, In Equation 1 is the compensatory parameter that define the metric to compute the distance. With , the metric is Euclidian.

Given that the variables representing the criteria are continuous and interval- and ratio-scaled, these scores are obtained by means of value functions (Beinat 1997), which transform the natural scale of a criterion to a [0, 1] value scale (1 represents the most undesirable state and 0 the most desirable state).

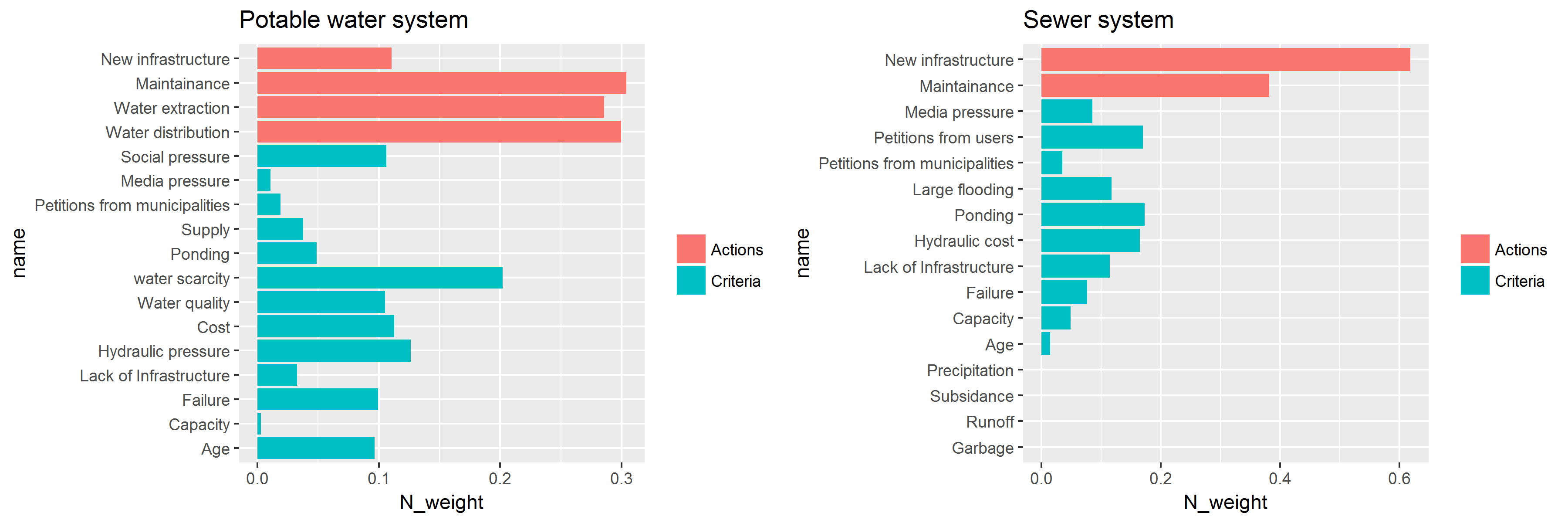


Figure 2: Criteria, actions, and criteria and action weights. The figures shows the input data to define water operators’ policy scenario. This input data correspond to the criteria weights and action weights using the limit of a weighted matrix T obtained from Analytic network (ANP). In red the actions and in blue the criteria. The value of w and are represented by the length of the bar.

### Site selection

Every year site selection is invoked by the water authorities for choosing investments , in action in system ,in a specific number of census blocks, established by budgetary constraints . Formally, this involves using a 0-1 (or binary) programming model (Dykstra 1984) in which the objective function maximizes . In this way, the model simulates a preference for investing in the census blocks where investment in infrastructure of system is most needed. Formally:

(2)

Subject to

where is the number of census-blocks where investment related to action can take place; is the 0-1 decision variable for action for system in census block at time (, if census block is selected for investment, or 0 otherwise). An implementation of other forms of maximization with more complete set of constrains is underway.

Budget represents a total capital or resources divided among census blocks. Thus when there is a limited among of resources to be divided in the city. The budget can be aggregated or disaggregated according to institutional rules. For instance in Mexico City the budget for drainage is separated from the budget for potable water system, and within each system the budget is divided in actions. The optimization of the objective functions were done using a non-dominant sorting assessment (NDSA) that find solutions near the Pareto frontier. The results from the NDSA was then used to constrain a genetic algorithm that find the best alternative in for investments each spatial unit (See Apendix optimization).

### Standardized score functions

This procedure standardizes each criterion using normalization functions needed to evaluate the distance to the ideal point of each census block related to each action and system . This procedure is called in every decision cycle to update in a standardized scale the value of the criteria.

##### Different procedures are implemented to capture different functional forms. Generally, the procedures take the following notation:

(3)

where is the perceived magnitude of stimulus defined by the value of attribute in census block at time , and a set of control parameters to ensure that , given the scale of the attributes and it relevance to each agent.

In the Netlogo version, the code also includes a procedure to compute using step-functions that take as argument the value of the attribute and a set of cut-off values, such that

(4)

where are cut-off values that follow a progression for increasing and decreasing functions. Parameter represents the maximum value of the attribute , which will set the range of the value function.

## Criteria for decisions and standardization function - Water authority

### Capacity

This refers to the perception of the capacity of the sewer system and the pipe network system . Formally, the standardized value of the criterion “capacity”, , is constructed using:

### 

Figure 3: This plot represents the relationship between the capacity of the sewer system and the standardized score .

### Failures

The criterion “failure”, represents the authority’s perception of the propensity of the infrastructure system to fail. It is represented by the condition of the infrastructure such that

(27)

### Lack of infrastructure

Lack of infrastructure, , is a criterion that is represented by an increasing function, that takes as arguments the residual of the attribute “infrastructure coverage” and a set of specific cut-off values for action, ,. That is:

(28)

Thus, census blocks with more unconnected residents will be prioritized.

### 

Water scarcity

This criteria Water Scarcity is define differently according to the action. Specifically scarcity for the action of distributing water uses monthly days without water to calculate the weekly distribution of water. However to conduct maintenance and new infrastructure provision the water authority uses the annual number of days without water. Both variables, days in a month and in a year without water, are represented by the accumulation of disruption of the water supply by systems and , evaluated monthly and yearly as part of the decision cycle :

where

represents the number of days in a year census block did not receive water by pipes nor by trucks. are the days without piped water in a week, and are the days that census block did not receive water in a week via system .

Figure 4: This plot represents the relationship between days in a month without water , and the standardized score .

Where, and are the value of the function evaluated at and , respectively, which are the limit of the range of possible monthly values in the model (28=4\*7, 4 weeks per month and 7 days per week) (Fig. 4).

Flooding

The criterion “flooding” is represented by the variable number of flooding per year per census block, . Thus,

. (30)

Ponding

Where, and are the value of the function evaluated at and , respectively, which are the limits of the range of possible annual number of ponding in a year (Fig. 4). The data suggests that .

Figure 5: This plot represents the relationship between number of ponding events and the standardized score .

Health

Health is represented by the annual number of incidences per year, such that

(31)

Where and are the value for the function evaluated at and respectivelly. Parameters and are control parameters to ensure .

Figure 6: This plot represents the relationship between number of incidences of gastro intestinal diseases (GID) and the standardized score .

Lack of supply

The criterion “lack of supply” represents the perception of the authority and the resident aboyt the lack of water related infrastructure. This is assumed to be related to the lack of infrastructure. Operationally we implemented a value associated to the proportion of the population connected to infrastructure systems, that is using variable . Thus for all infrastructure system we have the value functions

Where and are the value for the function evaluated at and respectivelly. Parameters and are control parameters to ensure .

Figure 7. This plot represents the relationship between the lack of infrastructure , and the standardized score.

Social pressure

Social pressure, , is a measure that results from the accumulation of protests in a census block over a period of time . Formally, it is represented as follows:

(33)

(34)

where are the total protests in census block , accumulated over the past yearly decision cycle .

Figure 8: The relationship between the number of protests accumulated in a year , and the standardized score .

Age of infrastructure

The criterion “age of infrastructure” represents the water operator’s perception of the age of the system. This is represented by the age,, such that

Where and are the value for the function evaluated at and respectivelly. Parameters and are control parameters to ensure.

Figure 8: The relationship between the average age of infrastructure systems in a census block, , and the standardized score .

Layers needed to be created and defined are the “Amount” and “Water quality”

## Criteria for decisions - Residents

Urbanization

Urbanization is the perception that green areas are being replaced by new urban areas. This is represented by the percentage of area urbanized in a census block at the time of the re-evaluation of priorities. Formally,

. (36)

Clogged drainage

Clogged drainage is a criterion that represents the perception of the residents that the problem of flooding is related to the accumulation of garbage and the lack of maintenance of the drainages. To represent this criterion geographically, we use the variable garbage, as a proxy, such that

Where and are the value for the function evaluated at and respectivelly. Parameters and are control parameters to ensure.

Figure #: the relationship between and index of garbage production , and the standardized score .

Insufficient infrastructure

This criterion is represented by the lack of coverage in the census block. Depending on the action, this criterion may refer to one of a number of different infrastructure systems. Therefore:

(38)

Water scarcity

The criterion “scarcity”, , is the residents’ perception of lack of water, represented by the number of days without water in the last week.

Where and are the value for the functionevaluated at and respectivelly. Parameters and are control parameters to ensure.

Figure # the relationship between the number of days without water in a census block, , and the standardized score for the residents.

Flooding

Similar to what was described for the water authority, the residents’ criterion “flooding” is represented by the variable :

(40)

Health

The criterion “health” is represented by the incidence of gastrointestinal diseases, :

(41)

Water deviation

This criterion refers to the perception by indigenous communities that water from local sources is being distributed to other census blocks, and it is represented by the “water deviated” attribute, :

(42)

## Policy scenarios

We define a policy scenario as the set of parameters and variables associated with each actor’s actions. These include the criteria names, the biophysical variables associated with each criterion, the set of criteria weights, the set of alternative weights, and the cut-offs of each value function. Currently the model includes the policy scenarios of the two water operators. Figure 3 and Tables 2 3 and 4 summarize the information.

## Water authority actions and changes to census block attributes

Once the model computes the distance metric for each census block and the selection procedure is activated, a set of actions are invoked. These actions change specific attributes of the landscape. Here we explain the actions and their consequences on the census block attributes. Table 2 summarizes these actions and the original explanations contained in the matrix input files.

### Maintenance

Maintenance reduces the age of infrastructure system , at a rate proportional to its effectiveness:

(15)

where is the effectiveness of maintenance.

### New infrastructure

The provision of new infrastructure influences the proportion of the population in census block covered with infrastructure system , such that:

(16)

where is the effectiveness of the action “new-infrastructure” in providing system to the proportion of houses that lack coverage, .

In addition, we assumed that when the action “new-infrastructure” is invoked, the wells in census block , , with , are replaced by new ones such that

(17).

The new well is located in the same census block and inherits the same extraction rate as the old one.

### Water extraction

The action “water extraction”, , represents the yearly increment in the system of wells . We assume that in each year census blocks can receive a maximum of one new well. Thus, when the procedure “water extraction” in invoked, the number of wells per census block, , increases according to:

(20).

Subsequently, the subsidence rate, , is influenced by water extraction pressure. This is implemented by assuming that the rate of subsidence, , changes over time with:

(21)

where and are the total number of wells inside aquifer , at time , and time 1 respectively, formally calculated using:

(22)

where is the number of wells in census block at time , and identifies the aquifer to which the census block belongs.

### Water distribution

The action “water distribution”, is the action of supplying water to neighborhood by the system pipes . When the water distribution procedure is invoked, a probabilistic model simulates the number of days a census block receive water in a week. We simulate the number of days with water per week per census block using a Poisson process, truncated between 0 and 7:

(10).

Where .

Thus, the amount of water delivered by the pipe system to census block is

(11)

where is the volume of water supplied to census block by the pipe system and w is the volume of water per day delivered to each census block. A condition that is related to the number of people in the census block and the , and the proportion of population connected to the system of pipes , .

, defined by the site selection procedure and the budget , are scheduled to receive water by the system of “trucks”, , such that the days with water delivered by trucks,, is formally defined as

(18).

Another way to represent this would be:

.

Thus, when the action “water distribution”,, is taken in a census block, that is , the number of days water was delivered by trucks is set to the difference between the days with water by pipes and the maximum number of days in a week (7).

If water is then distributed to this census-block (), then the total water delivered to the census block by this means is then:

(19).

## Resident actions

Residents can invoke actions that either influence the local infrastructure of the census block or change the socio-political landscape via protesting.

### House modification

House modifications influence the sensitivity of the house to hazard events. We define the sensitivity of the house as a change of the perception of the magnitude of hazards, assuming that modifications are cumulative, but saturated, meaning that there is a limit to adaptation and a rate at which these changes influence the sensitivity. Formally, we assume that sensitivity is defined by

(23)

where

(24)

is the accumulated number of times that the action “house modification” was invoked by the residents of census block .

### Protests

The procedure “protest” in a census block, defined by the symbol , is triggered when the action “protest” is perceived as a more needed action compared to “house modification”. Only triggered when the tolerance of local inhabitants to water scarcity among other criteria is higher than other hazards. Accordingly:

(25)

where is the distance to the ideal point related to the protest action .

## Exposure models

Exposure of census blocks to infrastructure hazards is related to the average condition of infrastructure system that create risk of water supply disruption and flooding. In addition to floding and scarcity residents suffer the exposure to waterborne pathogens represented by the number of incidence as a function of . The risk associated with these hazards depends on the condition of clean water and sewer systems, *c*. Formally (ten Veldhuis *et al* 2011):

(5)

where is the rate of decline of the infrastructure system . is the infrastructure´s age at time in weeks. The function assumes an exponential decay in condition related to the aging process of the infrastructure system . refers to the effect of subsidence on the condition of the infrastructure associated with the rate of subsidence in census block , with:

, (6)

where is the subsidence rate [mm/year] in each census block, and is the effect of subsidence on system . It is a conversion parameter that must be parametrized to ensure that .

#### Age of infrastructure

One of the attributes included tracks the age of infrastructure systems as they decline over time in a census block . We assumed that the age of the infrastructure system in census block , , changes weekly accordingly to:

(7)

where is a single weekly time-step.

### Exposure to water supply disruption

A census block can have weekly disruptions in water supply. These disruptions are assumed to be caused by failures of the infrastructure system , n addition to the risk associated with the location of the census block; that is, exposure to water scarcity increases with altitude. We also assumed that the water supply can only be delivered within each municipality by either the pipe system or by distribution via mobile sources , such as trucks. Formally, we define the weekly supply of water to census block using system as

Accordingly, the risk of exposure to water supply disruption from the network of pipes *P* is assumed to be associated with the condition of the system. Formally, we assumed that the average number of days in a week that a census block is without clean water service from pipes is represented by:

(8)

where is the mean number of days in a week without piped water, and is the estimated average number of days without piped water in a municipality (a parameter estimated using available survey data). Parameter represents the local correction factor for altitude differences of census block from the mean altitude of the municipality, such that:

(9)

where is the altitude in census block, and is the mean altitude of municipality.

Parameter represents the additional number of days without piped water due to specific disruptions associated with the condition of the pipes. To represent a specific disruption we use variable with , such that

.

Thus, when the condition of infrastructure system is lower than a random number drawn from a uniform distribution, there are extra days census block will suffer from disruption.

#### Infrastructure coverage

We define infrastructure coverage as the percentage of houses in a census block with connection to infrastructure system ,. Thus, when all the houses in census block are connected to system , and if none are.

### Exposure to flooding

This sub-model simulates the number of yearly events of flooding in each neighborhood, given a probability. Three initial prototypes have been explored thus far to obtain these probabilities. The first is a contingency Bayesian matrix approach based on calculating the partial probabilities of the number of events, given the amount of rainfall. The second model uses the same Bayesian contingency table approach, but divides the city in terms of age and defines contingency tables for an older city and a newer city. Instead of rainfall, these tables use the capacity of the sewer system as a predictor. In this case the model calculates the partial probability that a census block will have a number of events in intervals, conditional on the capacity of the sewer system. Both variables show a relationship, and therefore a new model that combines the effect of all variables was also included.

#### Bayesian contingency matrix model using age and sewer capacity

This model assumes that the risk of flooding associated with the malfunctioning of the sewer system in a census block depends on the condition of the sewage infrastructure system, which is influenced by age, capacity, and the effect of subsidence. The yearly number of flooding events was simulated using a contingency matrix and Bayesian rules to calculate posterior marginal probabilities of the number of events, conditional to the condition of the sewer system. The first step then was to separate the city between old and new sectors. Thus we define the set of old census blocks as the probability of having more than *f* flooding events, which is calculated using

.

This is the predicted posterior distribution of the number of events of magnitude *m*, given the observation of *F* events in a year, given the condition *C* and the observations, that is the likelihood, and is the prior capacity information, given the observations in the data. Whenever the value of the risk associated with age and capacity is higher than a random number generated from a uniform distribution :

The model also includes other procedures for simulating flooding events based on a linear regression model:

(13)

where is the expected number of flood events in census block at year . ,, and are the regressor parameters associated with the independent variables age , capasity , and hydraulic cost of the sewer system , respectively, all of them evaluated at week .

### Exposure to gastrointestinal diseases

The health model is implemented as two separate regression models that simulate the expected number of incidences of gastrointestinal diseases in the lowlands and in the highlands of Mexico City.

For the lowlands, a regression model of the form

(14)

was used to incorporate the full set of predictors and the spatial dependency observed in the data. is a vector of observations of the dependent variable, with one observation for every census block, is the number of flooding events in census block , and is a parameter that relates the number of flooding events to the risk of gastrointestinal diseases. is a vector of disturbance terms, where is assumed to be independently and identically distributed for all , with zero mean and variance of . In order to capture the spatial dependency observed in the incidence data, the model incorporates an additional regressor in the form of a spatially-lagged variable,(Anselin, 2001). This variable captures cross-section dependencies, in which a covariance structure exists in different locations derived from the geographic space (Anselin, 1998; Anselin, 2001). The term is the unknown spatial lag coefficient, and *W* is the contiguity matrix. This equation was estimated empirically in prior a prior study (*Baeza et al. In review*).

For the highlands…the

## Indicators of Vulnerability

The indicators obtained at the end of the simulation period are described in the sections below.

City average age of infrastructure system

This indicator corresponds to the average age of the infrastructure in the city over the last years of the simulation. formally,

where is the average age of infrastructure system , and the total number of census blocks in the urban landscape.

City average exposure to flooding and scarcity

These indicators are calculated using

for scarcity, and

for flooding, where and are the annual number of water disruption and flooding events, respectively, in census block at year . is the final time-step of the simulation and . is the total number of census blocks.

Census block average exposure was measured using

The total number of events in the 10 years of simulation in each census block was represented as

City average level of socio-political pressure

This index is calculated using the number of accumulated protests over the last time-steps of the simulation, divided by the total number of census blocks :

Vulnerability index

The vulnerability of a census block is calculated using the “surface of vulnerability” definition by (Luers 2005). In this framework, the vulnerability index is summarized as the ration between the product of exposure *E* and sensitivity *S*, and the adaptive capacity of the census block. Formally:

where is the vulnerability in census block at time . is the exposure, defined as the level of flooding or scarcity of water. represents the sensitivity of census block to hazard events . We measure this by keeping track of the number of decisions that involve house modifications and water storage. The more these actions accumulate in a census block, the less sensitive it would be to the exposure. Parameter represents the adaptive capacity of the census block. We assume that

(48)

where is the income index of census block . Thus, we explicitly assumed that census blocks with more resources have higher adaptive capacity than poor census blocks. That is, wealthy areas are less vulnerable because they have more access to resources to take action. We use purchase power as an indicator of adaptive capacity.

Inequality in exposure

This index is obtained by calculating the Gini coefficient. The Gini coefficient is a measure of dispersion, often used to measure dispersion in income and wealth in a population. We use it here to evaluate the dispersion in combined exposure to flooding and scarcity. The Gini coefficient is an index between 0, completely equal, to 1, completely unequal. Thus, the larger the value of the index, the higher the inequality in exposure. The Gini is effectively calculating by

(49)

where is the population of census blocks.

is the sum of exposure to flooding and scarcity at time t, where , are the values of the vulnerability index of census blocks, indexed in increasing order (). The inequality index is then the average of the Gini coefficient over the last yearly decision cycle:

(50)

Sensitivity to policy changes

To evaluate the sensitivity of each census block to changes in the policy scenarios, we calculate the coefficient of variation in exposure. The coefficient of variation is a measure of the variance in a sample relative to its mean, as it is calculated.

,

where is the mean in exposure to indicator in census block at the end of a set of simulations, and is the standard deviation of the sample of simulations. Thus, if census block *a* has a higher coefficient of variation than census block *b*, we say that census block *a* is more sensitive to changes in policy.

## Appendix

### Extensions

One of the extensions needed to run a set of decisions is the gis extension, which allows for the loading of the geographic information. Also, the csv-extension is needed to allow the model to read the comma-separated values and translate them into NetLogo variables. The Matrix-extension is used to update the priorities (weights) of census blocks when HNP are used.

### Generating classes of census blocks using a k-mean clustering classification

### Bayesian contingency tables

### Capacity index using attributes of sewer systems

### GIS pre-processing information

### Instructions for creating new types of agents and actions

#### Creating actions as agents

#### Setting value variables for new agents

## Tables

Table 1: Attributes of the census blocks of Mexico City.

|  |  |  |  |
| --- | --- | --- | --- |
| Variable | Description | Symbol | name variable in netlogo |
| Census block ID | Numerical key to identify each census block |  | ID |
| CVEGEO | Unique national identifier | - | CVEGEO |
| Estate | Administrative units |  | CV\_estado |
| Municipality | Units |  | CV\_mu |
| Aquifer | The aquifer that is below the census block |  | Aquifer |
| Aquifer zone | The zone of the aquifer below the census block |  | zona\_aquifera |
| Altitude | Meters above sea level |  | altitude |
| Group cluster ID | A classification of the census blocks of MC based on socio-economic and environmental similarities (appendix) |  | group\_kmean |
| Weekly water supply | Days in a week water was supplied by system |  | scarcity |
| Flooding | Number of flooding events per year |  | flooding |
| Protests | 1 if a protest occurs in a given week; 0 if otherwise |  | protests |
| Health | Annual incidences of gastrointestinal diseases |  | salud |
| Social pressure | Number of protests per year |  | Presion\_social |
| Media pressure |  |  |  |
| Wells | Set of wells within a census block |  | Pozos\_neighborhood |
| Infrastructure coverage | % houses connected to infrastructure system |  | Houses\_with\_*v* |
| Age infrastructure | Age of infrastructure system |  | Antiguedad-infra\_v |
| Condition | An index of the state of the infrastructure system in census block |  | condicion |
| Hydraulic load | Average volume of water per unit of time received by the sewer system in census block in a year |  | Gastro-hydraulico |
| Capacity | Index of the capacity of the pipes of system |  | capasidad |
| Rainfall | Total annual rainfall in census block |  | precipitation |
| Subsidence | Rate of subsidence per year |  | hundimientos |
| Income index | The purchase power by census block |  | Income-index |
| Potable water | Volume of water supplied to the census block by system at time |  | Water-in |
| Garbage | Garbage produced in each census block |  | garbage |
| Water quality | Index of the quality of potable water |  | water\_quality |
| Urban growth | Percentage of census blocks considered to be urbanized |  | urban\_growth |
| Water deviated | Number of wells multiplied by the days without water in census blocks with indigenous communities |  | desviacion\_agua |

Table 2: Actions of actors water authority and residents.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | Actions | Symbol | Definition | Attribute changed | Decision cycle | Action weight |
| Water Authority | Distribution |  | Refers to the water distributed by government trucks to areas without connection to the supply network |  | Weekly |  |
| Extraction |  | Increased extraction results from increasing number of wells |  | Annual |  |
| Maintenance |  | Repairing infrastructure system *v* |  | Monthly |  |
| New infrastructure |  | Refers to the action of providing new infrastructure in census blocks that lack coverage |  | Monthly |  |
| Residents | Collective action |  | Refers to the action of organizing to demand changes from the authority or to internally generate change locally | TBD | TBD |  |
| Water storage |  | Refers to the capture of water using storage devices |  | TBD |  |
| Purchase of water | TBD | Refer to the action of buying water from private sources | TBD | TBD |  |
| House modification |  | Action taken to modify the local condition of dwellings to reduce damage from flooding |  | Annual |  |
| Mobilization/protests |  | To express dissatisfaction with the public services of water delivery and sewage |  | Monthly |  |
| Water re-use |  | The action of recycling and re-using water | TDB | TDB |  |

Table 3: Criteria used in residents’ decision-making.

|  |  |  |
| --- | --- | --- |
|  | |  |
| Criteria | Definition | Attribute associated |
| Urbanization | Percentage of area urbanized  (driver of change from simulations of urban growth model) |  |
| Waste of water | Dummy |  |
| Water deviación | The perception of local people that live close to wells that water is being distributed to other census-blocks |  |
| Service efficiency | related to efficiency of infrastructure system |  |
| insuficient infrastructure | Represented by the percentage of population in each census-block not connected to sewer system |  |
| Contaminación de agua/water quality |  |  |
| drainage system cloGged | accumulation of garbage connected to the failure of the sewer system (in the layers folder says it is still dummy) |  |
| water scarcity | Number of days of water disruption |  |
| Flooding | Number of flooding events per year |  |
| Health risk | Number of incidences per pear |  |

Table 4: Criteria used in water supply operators’ decision-making.

|  |  |  |  |
| --- | --- | --- | --- |
| Criteria for calculating THE decision metric OF THE water supply operator | | | |
|  |  | Definition | unit |
| Infrastructure | Age of infrastructure | Average age of infrastructure per census block | years |
| Capacity | Capacity [in length of pipes] of the infrastructure to supply water or to discharge | Mts/area |
| Failure | An index of the number of infrastructure-related hazards per year (e.g., pipes break) | [events/year] |
| Lack of infrastructure | The lack of connection to potable water ans sewer systems | % houses not connected to infrastructure system |
| Hydraulic pressure | Pressure in pipes TBD | ? |
| Budget | Amount | The importance of the budget received from central authority | ? |
| Risks to the population | Water quality | TBD | ? |
| Water scarcity | Number of weeks in a year without water supply by system or | [weeks/year] |
| Flooding | Represented by the number of flooding events per year | [events/year] |
|  | Health | Represented by the number of gastrointestinal incidences per year | [events/year] |
| Socio-institutional | Supply | Represented by the requirements of the population on the infrastructure system | [pop \* need/pop.] |
| Petitions from census blocks | Demand of population funneled by politicians at the level of the municipality, representing the collective level of response from each municipality. |  |
| Social pressure | Represented by the number of protests per year | [pop \* need/person] |