Agent-based model Mexico City
ODD Protocol

Andres Baeza
Postdoctoral researcher
School of Sustainability
Arizona State University

Table of Contents

Introduction

This document describes the implementation of the Agent based model of the MEGADAPT project. The 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 on the spatial distribution of socio-hydrological vulnerability in Mexico City. The model simulates decisions on investments in infrastructure by the water authority agents, which is done by evaluating the condition of the landscape across the neighborhoods. These investments in turn trigger actions in selected neighborhoods. These actions then influence the condition of the socio-political and physical attributes in the landscape that in turn modify the biophysical condition that influence the risk to infrastructure hazards and the exposure to flooding, disruption of water supply, and gastrointestinal diseases.

The decision making process of the water authorities involves the identification and selection of neighborhoods that need investments. This prioritization of neighborhoods for investments 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 relay 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 in their local neighborhood and the action taken modifies only those local attributes. Figure 1 shows an overview of the basic processes included in the model.

The current version incorporates stochastic simulations of weekly water supply, annual production of floods and annual cases of gastro intestinal incidences. The model were 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 process of the agents.

The current version of the model is constructed in Netlogo V 6.2.1. The advantage of Netlogo is its versatility to develop agents as objects, which allow to define actors their actions and the infrastructure systems as separate computational entities.

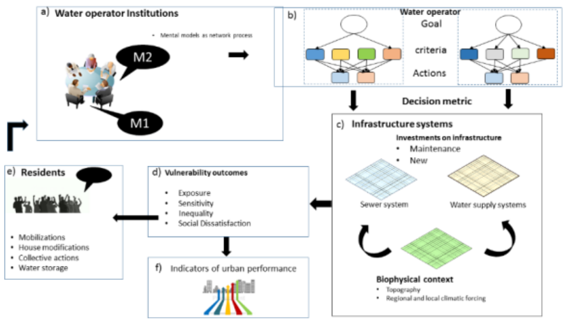

Figure 1: Overview of the Agent-based model

Agents, actions, scale (and notation)

Three types of agents are incorporated in the model: Water operators, residents and infrastructure systems. Resident are located in neighborhoods. Each neighborhood is indexed using symbol j and the total number of neighborhoods by symbol J. Each neighborhood contains a set of attributes, A_j. Infrastructure systems are also considered as agents when they are defined as object, with properties and methods similar to those of other agents.

Water operators make investments on different infrastructure systems, v in selected neighborhoods. These investments are related to the set of actions, k: 1) repairs and maintenance,k_R, 2) construction of new-infrastructure k_N, 3) distribution of water k_W , extraction of water, k_E. Residents, can take different actions as well. These actions are: 1) house modification k_M, 2) collective actions, k_C, and 3) protesting, k_∂.

The decisions to invest by the water operators and to modify the local environment by residents, are evaluated based on a multi-criteria metric that its calculation requires a set of criteria, criteria weight, alternative weights each specifically related to the actor and the actions they take. This information and the actual actions of the agents is obtained from empirical data from the actor's mental models, and obtained from 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).

Infrastructure systems are labeled with the symbol v. The current version of the model includes as infrastructure systems the wells for water extraction, v_Q, the system of pipes for potable water distribution, v_P, the sewer system v_S, the system for water distribution (trucks) v_T. In the current version the system of pipes and sewer system are represented as a set of attributes in each neighborhood, such as coverage, average age, condition, and capacity. The system of wells on the other hand, are defined as agents, with attributes specific for them. However neighborhoods also storage as a tributes the set of wells that bellow to each neighborhood.

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

The actions the agents take modify the attributes of the neighborhoods that in turn will influence the stochastic production of hazards. These changes in events in turn inform subsequent decisions by the actors. Therefore the model simulates the feedback loop between actor's decisions and biophysical changes in a dynamics and spatial explicit platform.

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

The temporal resolution of the model is 1 week. Depending on the action and agent different period in a year can represent different evaluation of the priorities. We call this period a cycle of decision *T*. A year in the model represent a cycle of decision T_y^D while a month would be cycle of decision T^D. for the water operators for action maintenance and new-infrastructure. Each cycle of decision T_y^D this agent re-evaluate the neighborhoods selected for investments for action maintenance, new-infrastructure, and extraction. For the action water distribution the cycle of decision is T_m^D. For the residents the cycle of decision is T_y^D as well. Actions are independent from the evaluation in the sense that they can occur in different periods. These are cycles of actions, T^A. 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 the action protest, and monthly for the actions water storage and house modification.

Process Overview and scheduling

The model currently is 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 the agents and their attributes. It also contain procedures to read the information that will define the decision making process 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 neighborhood attributes are defined.

The file value-function contains the procedures to update the information needed for the actors to calculate the multi-criteria decision metric. Specifically, these procedures update the standardize values of the attributes of neighborhoods that are criteria for decision, and calculate the distance metric.

The code file contains the procedure call "GO". Procedure "GO" involves the suitability assessment and site selection procedures, which define the neighborhoods that will be selected for investments by the water authorities. Once the neighborhoods 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 scale of the neighborhood, municipality and city.

To run the dynamic part of a single simulation, the observer needs to invoke the procedure "GO", which will trigger a set of processes related to the stochastic simulation of water allocation, flood and health models, the decision-making process of actors, the action of the actors, and the changes that those actions caused to the attributes of the neighborhoods. The decision making process of the water authorities includes procedures to a site selection process and suitability of neighborhoods In the case of the residents this involve a selection of actions. Within in each cycle of decision, the actions will modify the attributes of the landscape associated to each action in each neighborhood. Details about these processes are provided in the section "sub-models".

Design concepts

Basic principles

The model aims to represent the decisions of important actors in Mexico City water management system to invest in infrastructure systems in a geo-spatial and dynamical multy-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 as important for decision (criteria), and they consider them with different importance (criteria weights). In the model each criterion is associated to an attribute of the landscape. Therefore the agents evaluate the attributes in the landscape to make decisions related to investments 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 important, due to the changes in the attributes made by the actors' action.

Emergence

From the interaction between agent's actions and the modification to the attributes of the neighborhoods, we expect the emergence of spatial patterns of investments that in turn should influence the spatial pattern of vulnerability of neighborhoods to infrastructure related hazards.

Observation

The agents observe the attributes of the landscape that are considered as 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 made by the water operators. These observations are computed by each neighborhood and aggregated by each municipality.

Adaptation

Neighborhood can modify their local environment to reduce their sensitivity to the exposure to flood and scarcity hazards. Thus the action "modification of house" reduces the sensitivity of the neighborhood to flooding, whereas the action water storage reduces the sensitivity of a neighborhood to suffer exposure to water scarcity. These processes over time can be consider as adaptation to their local environment.

Interactions

The water operators and resident agents interact with the neighborhoods and the infrastructure systems by modifying their attributes. The attributes of the neighborhoods in turn influence the risk of infrastructure failures which will influence subsequence decisions of the water operator and the residents. There also interaction between actors that influence their decisions. Specifically, the action of the residents to protest generate social pressure which is one of the criterion that is used in the decision-making process of the water authorities.

Stochasticity

The model has procedures that generate stochastic events of hazards related to flooding, scarcity and health. Currently, the model incorporate two different methods to generate flooding events: A linear regression model that generate expected number of events per year using a set of infrastructure-related and biophysical independent predictors from each neighborhood. The second method is based on contingency table using Bayes rules, where flood outcomes of 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 process which generates number of days that neighborhoods received water by 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-neighborhoods" and 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 #) in 2) the model sets the type of agents or "breeds" in netlogo terminology, and in 3) their attributes. Procedure 4 contains methods to read 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 neighborhoods, including location. Procedure 6) contains methods to read data to define the actions and the information to define the multi-criteria decision mechanism involve in the decisions of the actors. In 7) are the algorithms to define the infrastructure systems, v. In 8) are methods to read 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 neighborhoods. This is needed to define the range of variability in the attributes, which in turn will inform the computation of the decisions. Finally, procedure 11) defines the neighborhoods that are associated with each other based on the contiguity matrix. This defines the local network of interaction between neighborhoods that influenced each other. Which is needed to simulate the health model.

Input data

The setup procedure invoke a procedure that loads layers of geographic information to fill the attributes of each neighborhood that will define the criteria for decisions. Table 1 shows the attributes of each neighborhood. The model also loads data to define the actions, the criteria, criteria weights and action weights. This information is read csv files in procedure SETUP using the cvs-extension capability of Netlogo. Another type of input data is the average number of days without water by municipality, which is needed to simulate the days with water by the system of pipes, v_P, using a Poisson process. Finally, the model contain procedures to read data to represent the probabilities of flood using Bayesian contingency matrix analysis () and regression analysis of health.

Sub-models and components

Water authority decision procedure

Priority assessment

An assessment of the prioritarian neighborhoods for investment is obtained through multicriteria evaluation of the distance of each neighborhood from an “ideal point” or utopian state, defined a set of decision criteria and how relatively important those criteria are for the decition makers. Formally we calculate a distance d_{jvt}^{k}, such that:

$$d_{jvt}^{k} = \sqrt[1/\rho]{\alpha \sum_i^I {w_{iv}}^{\rho} \, {x_{ijvt}^{k}}^{\rho}} \qquad (1)$$

where d_{jvt}^{k} is the distance to the ideal point of neighborhood j with respect to decision k and system v; w_{iv} is the criterion weight of criterion i related to system v; x_{ijvt}^{k} is the standardized score in a neighborhood j of the attribute corresponding to criterion i with respect to infrastructure system v and decision k; x_{ijvt}^{k} is the departure of an alternative from the ideal point for a criterion; i, j, k, t and v are indices for criteria, neighborhoods, action, time and infrastructure system, respectively. In Equation 1, the standardized score, x_{ijvt}^{k}, represents a judgment about the importance of an observable stimulus (neighborhood attribute value) for the water authority’s decision.

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).

Site selection

Every year site selection is invoked by the water authorities for choosing a single investment on on system v,y, in a specific number of neighborhoods that is established by budgetary constraints, B_k. Formally, this involves using 0-1 (or binary) programming model (Dykstra 1984) in which the objective function maximizes d_{jvt}^{k}. In this way, the model simulates a preference for investing in the neighborhoods where investment on infrastructure of system v is most needed; formally:

$$maximize\ F(y) = \sum_{j}^{J} d_{jvt}^{k} y_{jvt}^{k} \qquad (2)$$

Subject to

$$\sum_{j}^{J} y_{jvt}^{k} < B_k$$

where B_k is the number of neighborhoods where investment related to action k can take place; y_{jvt}^{k} is the 0-1 decision variable for action k for system v in neighborhood j at time t ($y_{jvt}^{k} = 1$, if neighborhood j is selected for investment, or 0 otherwise);

Value Functions

This procedure updates the information needed to evaluate distance to the ideal point of each neighborhood related to each action k, and system v. This procedure is called every cycle of decision. The information will define the vectors of criteria and will update their representation in a standardize scale using the procedure report "value function". This steps is critical to quantify relationships between condition of the attributes in the landscape (e.g., age, capacity, etc) and the perceived response by agents. Formally this procedure takes the following formulation:

$$x_{ijvt}^{k} = f(A_{ijt}, \varrho), \quad (3)$$

Where x_{ijvt}^{k} is the perceived magnitude of stimulus defined by the state of attribute i in neighborhood j at time t, A_{ijt}. Parameter ϱ refers to the constant fraction (ref). Function $f()$ is often represented by a logarithmic function:

$$x_{ijvt}^{k} = \varrho \log A_{ijt}$$

However, in the current version of this model, the function $f()$ is implemented using a set of cutoff ϱ_{ivc}^{k}, such that

$$x_{ijvt}^{k} = \begin{cases} 1 & \text{if} \quad A_{ijt} > \varrho_{iv4}^{k}\widetilde{A}_{\iota} \\ 0.8 & \text{if} \quad \varrho_{iv4}^{k}\widetilde{A}_{\iota} > A_{ijt} > \varrho_{iv3}^{k}\widetilde{A}_{\iota} \\ 0.6 & \text{if} \quad \varrho_{iv3}^{k}\widetilde{A}_{\iota} > A_{ijt} > \varrho_{iv2}^{k}\widetilde{A}_{\iota} \\ 0.4 & \text{if} \quad \varrho_{iv2}^{k}\widetilde{A}_{\iota} > A_{ijt} > \varrho_{iv1}^{k}\widetilde{A}_{\iota} \\ 0.2 & \text{if} \quad \varrho_{iv1}^{k}\widetilde{A}_{\iota} < A_{ijt} \end{cases} \quad (4)$$

Where $\varrho_z$are canonical cut-off, that follow the Weber-Fechner progression {0.5,0.25 0.125,0.0625} for increasing functions and {0.937,0.875 0.725,0.5} for decreasing functions. Parameter $\widetilde{A}_{\iota}$ represents the maximum value of the attribute, A_{ijt}, which will set the range of the value function.

In the current version of the model, this procedure is called in the context of actions, which are called in the context of the neighborhood.

Exposure models

Exposure of neighborhoods to infrastructure hazards is assumed to be related to the average risk of water supply disruption and flooding. Floods also influence the exposure of the population to waterborne pathogens.

In addition to the landscape (that is, exposure to water scarcity increases), the risk associated with these hazards depends on the condition of clean water and sewer systems, *c*; formally (ten Veldhuis, Clemens, and van Gelder 2011):

$$c_{jvt} = \frac{\Psi_{jvt} + \left(1 - e^{-\alpha\lambda_{jdvt}}\right)}{2} \quad (5)$$

where α is the rate of decline of the infrastructure system v. λ_{jvt} is the infrastructure´s age at time t in weeks. The function $e^{-\alpha\lambda_{jvt}}$ assumed a exponential decay in the condition related to the aging process of the infrastructure system, v. Ψ_{jvt} refers to the effect of subsidence on the condition of the infrastructure associate to the rate of subsidence in neighborhood j, with:

$$\Psi_{jvt} = \xi_v \psi_{jt}, \quad (6)$$

Where ψ_{jt} is the subsidence rate [mm/year] in each neighborhood, and ξ_v is the effect of subsidence to system v. It is a conversion parameter that must be parametrize to ensure that $\Psi_{jvt} \in [0,1]$.

Age of infrastructure

It is an attribute that tracks the age of infrastructure systems, v, as they decline over time in the a neighborhood, j; We assumed that the age of the infrastructure system v, in neighborhood j, λ_{jvt}, changes weekly accordingly to:

$$\lambda_{jvt} = \lambda_{jvt-1} + week \quad (7)$$

where t is a single weekly time-step.

Exposure to water supply disruption

A neighborhood j can have weekly disruptions in water supply. These disruptions are assumed to be caused by failures of the infrastructure system, v. We also assumed that water supply can only be delivered within each municipality by either the pipe system, v_P, or by distribution by mobile sources, v_T, such as trucks. Formally, we define the weekly supply of water to neighborhood j using system v_P as, s_{jv_Pt}.

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

$$\varkappa_{jt} = \breve{\varkappa}_{\mathcal{M}} + b_1(\hbar_j) + b_2\varphi \text{ (8)}$$

Where $\varkappa_{jt}$ is the mean number of days in a week without piped water, $\breve{\varkappa}_{\mathcal{M}}$ is the estimated average number of days without piped water in a municipality $\mathcal{M}$ (A parameter estimated using available survey data). Parameter b_1 represent the local correction factor for altitude differences of neighborhood j from the mean altitude of the municipality $\mathcal{M}$, such that:

$$\hbar_j = h_j - \bar{h}_{\mathcal{M}} \quad , \quad \forall\, j: \mathcal{M}_j = \mathcal{M} \text{ (9)}$$

Where, $h_j$is the altitude in neighborhood j, and $\bar{h}_{\mathcal{M}}$ is the mean altitude of municipality $\mathcal{M}$.

Parameter b2 represent 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 φ with $\varphi = \{0,1\}$ such that

$$\varphi = \begin{cases} 1 \text{ if } (1 - c_{jvt}) > X \sim U([0,1]) \\ 0 \qquad\qquad \text{otherwise} \end{cases}$$

Thus, when the condition of infrastructure system v is lower than a random number drawn from a uniform distribution, there are b_2 extra days neighborhood j will suffer from disruption.

We simulate stochastic realization of days with water per week per neighborhood using a Poisson process, truncated between 0 and 7:

$$s_{jv_Pt} = 7 - pois(\varkappa_{jt}) \text{ (10)}$$

Thus, the amount of water delivered by the pipe system to neighborhood j is

$$W_{jv_Pt} = s_{jv_Pt} P_j \eta_{jv_Pt} w \quad \text{(11)}$$

Where W_{jv_Pt} is the volume of water supplied to neighborhood j by the pipe system, v_P. P_j is the number of people in neighborhood j,and η_{jv_Pt} is the proportion of population connected to the system of pipes v_P. Parameter w is the consumptive use of water per person, in units of volume, and parametrized using minimum water requirements per person. Therefore it is assumed that water is delivered by pipes is proportional to the population usage and the coverage of supply infrastructure system v_P.

Infrastructure coverage

We define infrastructure coverage as the percentage of houses in a neighborhood with connection to infrastructure system v, η_{jvt}. Thus, when $\eta_{jvt} = 1$ all the houses in neighborhood j are connected to system v, and $\eta_{jvt} = 0$ if none.

Exposure to flooding

Commented [AB1]: Feedback from Yosune here. Thanks

Another assumption is that the risk of flooding associated with malfunctions of sewer system in a neighborhood depends on the condition of the sewer infrastructure system, which is influenced by age, capacity, and the effect of subsidence. Yearly number of flooding events were simulated using contingency matrix and Bayes rules, to calculate posterior marginal probabilities of the number of events, conditional to the condition of the sewer system. The first then was to separate the city between old and new. Thus we define the set of old neighborhoods as

The probability of having more than f event of flooding is calculated using

$$p(f_m|F) = \int_{\theta} p(f|C,F)p(C|X,a)\,d\theta$$

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 $p(C|X,a)$ is the prior information of the capacity, given the observations. Whenever the value of the risk associated with age and capacity is higher than a random number generated from a uniform distribution, U :

$$f_{jT} = \begin{cases} f_m \text{ if } U > p(f_m|F) > U \\ \end{cases} \quad (12)$$

The model also includes other procedures to simulate flooding events based on a linear regression model:

$$E\left[f_{jT}\right] = \beta_0 + \beta_1 \lambda_{jv_st} + \beta_2 \varepsilon_{jv_s} + \beta_3 o_j + \xi(0,1) \quad (13)$$

Where $E\left[f_{jT}\right]$ is the expected number of flood events in neighborhood j at year T. $\beta_1, \beta_2, \beta_3$ are the regressor parameters associated to the independent variables age λ_{jv_st}, capasity ε_{jv_s}, and hydraulic cost, o_j, of the sewer system, v_s, respectively all of them evaluated at week t.

Exposure to gastro intestinal diseases

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

$$g_{jT} = \rho W G_T + \boldsymbol{\beta} f_{jT} + \varepsilon_j \quad (14)$$

was used to incorporate the full set of predictors and the spatial dependency observed in the data. G_T is an $J \times 1$ vector of observations of the dependent variable, with one observation for every neighborhood, f_{jT} is the number of flooding in neighborhood j, $\boldsymbol{\beta}$ is a parameter that relates the number of flooding to the risk of gastrointestinal diseases. ε_j is a $J \times 1$ vector of disturbance terms, where ε_j is assume to be independently and identically distributed for all j, with zero mean and variance σ^2. 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, WG_T(Anselin, 2001). This variable captures cross-section dependencies, in which exist a covariance structure in different locations derives from the geographic space (Anselin 1998, Anselin, 2001). The term ρ is the unknown spatial lag coefficient, and *W* is the $J \times J$ contiguity matrix. This equation was estimated empirically in (*Baeza et al. In review*).

For the highlands…TBC

Water authority actions and changes on neighborhoods' attributes

Once the model computes the distance metric for each neighborhood 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 its consequences on the neighborhood attributes. Table 2 summarize these actions and the original explanation contained in the matrix input files.

Maintenance

Maintenance reduces the age of infrastructure system, v, proportionally to its effectiveness, formally:

$$\lambda_{jvt} = \lambda_{jvt-1} + \phi\lambda_{jvt-1}, \quad (15)$$

where ϕ is the effectiveness of maintenance.

New Infrastructure

The provision of new infrastructure influences the proportion of the population in neighborhood j covered with infrastructure system v, η_{jvt}, such that:

$$\eta_{jvt} = \eta_{jvt} + \varsigma\,(1 - \eta_{jvt}) \text{ if } \; y_{jvt}^{k_N} = 1 \quad (16)$$

where ς is the effectiveness of the action "new-infrastructure" in providing system v to the proportion of houses that lack of coverage, $(1 - \eta_{jvt})$.

In addition, we assumed that when the action new-infrastructure is invoked, the wells in neighborhood j, q_{jt}, with $\lambda_{jv_Qt} > 40$ years, are replaced by new ones such that

$$\lambda_{jv_Qt+1} = 0 \quad \text{if} \quad y_{jvt}^{k_N} = 1 \quad (17)$$

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

Water distribution

The action water distribution, k_Q, is the action of supplying water to areas that had a disruption in the supply of water by the system pipes, v_P. When the water distribution procedure is invoked, a set of neighborhoods, defined by the site selection procedure and the budget B_T, are scheduled to receive water by system of “trucks”, v_T, such that the days with water delivered by trucks, s_{jQt}, is formally defined as

$$s_{jv_Tt} = \begin{cases} (7 - s_{jPt}) y_{jv_Tt}^{k_W}, \; y_{jv_Tt}^{k_W} = 1 \\ 0, \; y_{jv_Tt}^{k_W} = 0 \end{cases} \quad (18)$$

Another way to represent this would be:

$$s_{jv_Tt} = (7 - s_{jPt}) y_{jv_Tt}^{k_W}$$

Thus, when the action water distribution, k_W, is taken in a neighborhood, that is $y_{jv_Tt}^{k_W} = 1$, the number of days water was delivered by trucks is set to the difference between the days with water by pipes and the max number of days in a week (7).

IF water is then distributed to this neighborhood ($y_{jv_Tt}^{k_W} = 1$), then the total water delivered to the neighborhood by this mean is then:

$$W_{jv_Pt} = s_{jv_Tt} P_j w \quad (19)$$

Water extraction

The action water extraction, k_E, represents the yearly increment in the system of wells, E. We assume that in each year neighborhoods can receive a new well. Thus when the procedure water extraction in invoked, the number of wells per neighborhood, q_{jt}, increases according to:

$$q_{jt+1} = q_{jt} + 1 \text{ if } y_{jv_Qt}^{k_E} = 1 \quad (20)$$

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

$$\psi_{jt} = \psi_{jt} \frac{Q_{Zt}}{Q_{Z1}}, \quad (21)$$

where Q_{Zt} and Q_{Z1} are the total number of wells inside aquifer Z, at time t, and time 1 respectively, and are formally calculated using:

$$Q_{Zt} = \sum_{\forall j: Z_j \in Z} q_{jt} \quad (22)$$

Where q_{jt}, is the number of wells in neighborhood j at time t, and Z_j identifies the aquifer in which the neighborhood j belongs to.

Resident actions

Residents invoke actions that influence the local infrastructure of the neighborhood or change the socio-political landscape by protesting.

House modification

House modifications influence the sensitivity of the house to suffer hazard events. We define the sensitivity of the house as the modulation of the perception of the magnitude of hazards. We assumed also that modifications are cumulative, but it saturates, meaning that there is a limit to adaptation and a rate at which these changes influence the sensitivity. Formally we assume that sensitivity S is defined by

$$S_{\mathcal{E}} = 1 - \frac{Y_{jt}^{k_M}}{Y_{jt}^{k_M} + \varsigma} \text{ (23)}$$

Where

$$Y_{jt}^{k_M} = \sum_{t=1}^{t} y_{jvt}^{k_M} \text{ (24)}$$

Is the accumulated times that the action house modification was invoke by the residents of neighborhood j

Protests

When the procedure protests in a neighborhood, ∂_{jvtq} , are triggered when the valuation of the action protest exposure to infrastructure hazards, the response of the water authority, and the tolerance of local inhabitants to those hazards; accordingly:

$$\partial_{jt} = \begin{cases} 1 \; if \; d_{jt}^{k_\partial} > d_{jt}^{k}, \; \forall \; k \neq k_\partial \\ 0, \qquad \text{otherwise} \end{cases} \text{ (25)}$$

where $d_{jd}^{k_\partial}$ is the distance to the ideal point related to action protest, k_∂.

Criteria for decision -Water authority

Capacity

This refers to the perception of the capacity of the sewer system v_S and the pipes network system, v_P. Formally the criteria “capacity”, $x_{i_\varepsilon jvt}^{k}$, is constructed by standardizing the attribute capacity, using $f()$, such that:

$$x_{i_\varepsilon jvt}^{k} = f\left(\varepsilon_{jv}, \left\{\varrho_{i_\varepsilon vc}^{k}\right\}\right) \text{(26)}$$

Failures

The criterion failures, i_C, represents the perception of the authority about the propensity of the infrastructure system to failures. It is represented by the condition of the infrastructure c_{jvt}, such that

$$x_{i_C jvt}^{k} = f\left(c_{jvt}, \left\{\varrho_{i_C vc}^{k}\right\}\right). \quad (27)$$

Lack of infrastructure

Lack of infrastructure is a criteria, i_η, that is represented by an increasing function, $x_{i_\eta jvt}^{k}$, that takes as arguments the residual of the attribute "infrastructure coverage", and a set of specific cut-off values for action k, $\varrho_{i_\eta vc}^{k}$. That is

$$x_{i_\eta jvt}^{k} = f\left(\left(1-\eta_{jvt}\right), \varrho_{i_\eta vc}^{k}\right). \quad (28)$$

Thus, neighborhoods with more unconnected residents will be prioritized.

Amount

This need to be define. (dummy layer now)

Water quality

This need to be included. (dummy layer now)

Water scarcity

This criteria is represented by the accumulation of disruption of water supply by systems v_P and v_T evaluated at yearly for cycle of decision, T_y^D:

$$x_{i_S jvt}^{k} = f\left(S_{jT}, \varrho_{i_S vc}^{k}\right) \quad (29)$$

Where

$$S_{jT} = \sum_{t}^{t-T_y^D} \left(7 - s_{jv_P t}\right) + \left(1 - y_{jv_T t}^{k_W}\right) s_{jv_T t}$$

Represent the number of days in a year neighborhood j did not received water neither by pipes nor by trucks. $\left(7 - s_{jv_P t}\right)$ are the days without piped water in a week, and $\left(1 - y_{jv_T t}^{k_W}\right) s_{jv_T t}$ the days that neighborhood j did not received water in a week by system v_T, either.

Flooding

The criteria flooding is represented by the variable with the same name, f_{jt}.Thus

$$x^k_{i_F jvt} = f\left(f_{jt}, \varrho^k_{i_F vc}\right) \quad (30)$$

Health

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

$$x^k_{i_G jvt} = f\left(g_{jt}, \varrho^k_{i_G vc}\right) \quad (31)$$

Supply

The criteria lack of supply represent the people without piped water infrastructure

$$x^k_{i_P jvt} = f\left((1 - \eta_{jv_P t}), \varrho^k_{i_G vc}\right) (32)$$

Social pressure

Social pressure, $\theta_{j\Delta T}$, is a measure that results from the accumulation of protests in a neighborhood in a period of time T, formally:

$$x^k_{i_\theta jvt} = f\left(\theta_{j\Delta T}, \varrho^k_{i_\theta vc}\right) \quad (33)$$

$$\theta_{j\Delta T} = \sum_t^{t-T_y^D} \partial_{jt}, \quad (34)$$

Where, $\theta_{j\Delta T}$are the total protests ∂_{jt} in neighborhood j and accumulated over the past yearly cycle of decision T_y^D.

Age of infrastructure

The criteria age of infrastructure represent the perception of the water operator about the age of the system. This is represented by the age, λ, such that

$$x^k_{i_\lambda jvt} = f\left(\lambda_{jvt}, \varrho^k_{i_\lambda vc}\right). \quad (35)$$

Criteria for decision -Residents

Urbanization

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

$$x^k_{i_\varpi jvt} = f\left(\varpi_{jt}, \varrho^k_{i_\varpi vc}\right) (36)$$

Drainage clogged

Drainage clogged represents a criterion about the perception of the residents that the problem of flooding relates to the lack of maintenance on drainage that get clogged with garbage. To represent this criteria geographically, we use in this implementation, the amount of garbage produced in each neighborhood, θ_j, such that

$$x^k_{i_\theta jvt} = f\left(\theta_j, \varrho^k_{i_\theta vc}\right) (37)$$

Insufficient infrastructure

This criteria is represented by the lack of coverage in the neighborhood. Depending on the action this criteria refers to different infrastructure systems, therefore:

$$x^k_{i_\eta jvt} = f\left(\left(1 - \eta_{jvt}\right), \varrho^k_{i_\eta vc}\right) (38)$$

For instance, action "storage water occur water

Which the same function than for insufficient infrastructure from the perspective of the water authority actions.

Water scarcity

The criteria scarcity residents, i_{SR}, is the perception of lack of water by the residents and is represented by the days without water in the last week.

$$x^k_{i_{SR} jvt} = f\left(\left(\left(7 - s_{jv_P t}\right) + (1 - y^{k_W}_{jv_T t}) s_{jv_T t}\right), \varrho^k_{i_{SR} vc}\right) (39)$$

Flooding

Similar to the water authority, the criteria flooding is represented by the variable f_{jt}:

$$x^k_{i_F jvt} = f\left(f_{jt}, \varrho^k_{i_F vc}\right) (40)$$

Health

The criteria health is represented by the variable incidence of gastrointestinal diseases g_{jt}:

$$x^k_{i_G jvt} = f\left(g_{jt}, \varrho^k_{i_G vc}\right) \text{ (41)}$$

Water deviation

This criteria refers to the perception by indigenous communities that water from local sources is being distributed to other neighborhoods, and is represented by the attribute water deviated, ζ_{jt}:

$$x^k_{i_\zeta jvt} = f\left(\zeta_{jt}, \varrho^k_{i_\zeta vc}\right) \text{ (42)}$$

Petitions from Neighborhoods

TBD

Policy scenarios

We define a policy scenarios as the set of parameters and variables associated to each actor's action. These include the criteria name, the biophysical variables associated to each criteria, the set of criteria weights, the set of alternative weights, and the cutoffs of each value function. Currently the model includes the policy scenarios from the two water operators. Table # summarize the information.

Indicators

At the end of the period of simulation, the following indicators were obtained:

City-average age of infrastructure system v

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

$$\Gamma_v = \sum_{t=T_{Max}-\Delta T}^{T_{Max}} \frac{1}{J} \sum_{j=1}^{J} \lambda_{jvt} \qquad (43),$$

where Γ_v is the average age of infrastructure system v, and J the total number of neighborhoods in the urban landscape.

The city-average exposure to flooding E_F and scarcity E_S.

This indicators is calculated using

$$E_S = \frac{1}{J}\sum_{j=1}^{J}\frac{1}{\Delta T}\sum_{T=T_{Max}-\Delta T}^{T_{Max}} S_{jT} \quad (44)$$

For scarcity, and

$$E_F = \frac{1}{J}\sum_{j=1}^{J}\frac{1}{\Delta T}\sum_{T=T_{Max}-\Delta T}^{T_{Max}} f_{jT}$$

for flooding. Where S_{jT} and f_{jT} are the annual number of events water disruptions and to flooding in neighborhood j at year T respectivelly. T_{Max} is the final time-step of the simulation and $\Delta T = 10$ years. J is the number of neighborhoods.

Neighborhood-average exposure

$$E_{St} = \sum_{t=T_{Max}-\Delta T}^{T_{Max}} S_{jT} \qquad (45)$$

This is the number of events in the 10 years of simulation in each neighborhood:

$$E_{St} = \sum_{t=T_{Max}-\Delta T}^{T_{Max}} f_{jT} \qquad (46)$$

City-average level of socio-political pressure

This index is calculated using the accumulated protests over the last Δt time-steps of the simulation and then divided by the total number of neighborhoods J:

$$\Sigma = \frac{1}{\mathrm{J}}\sum_{j=1}^{\mathrm{J}}\sum_{t=T_{Max}-\Delta T}^{T_{Max}} \partial_{jt} \quad (47)$$

Vulnerability index

The vulnerability of a neighborhood is calculated using the definition of "surface of vulnerability" 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 neighborhood. Formally

$$V_{jt}^{\mathrm{H}} = \frac{E_{jt}^{\mathrm{H}} * S_{jt}^{\mathrm{H}}}{\tau_{jt}} \ (48)$$

Where $V_{n,t}$ is the vulnerability in neighborhood j at time t. E_{jt}^{H} is the exposure, defined as the level of flooding or scarcity of water. S_{jt}^{H} represents the sensitivity of neighborhood j to hazard events H. 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 neighborhood, the less sensitive it would be to the exposure. Parameter $\tau_{j,t}$ represents the adaptive capacity of the neighborhood. We assume that

$$\tau_{jt} = \left(1 - I_{jt}\right) \text{ (48)}$$

Where $I_{j,t}$ is the income-index of neighborhood j. Thus, we explicitly assumed that neighborhoods with more resources have higher adaptive capacity than poor neighborhoods. that wealthy areas are less vulnerable because they can have more access to resources to take actions. We use the purchase power as 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, complete equal, to 1 completely unequal. Thus the larger the value of the index the higher the inequality in exposure we day. The Gini is effectively calculating by

$$gini_t = \frac{1}{J}\left(J + 1 - 2\frac{\sum_{\iota=1}^{J}(J+1-\iota)V_{\iota t}}{\sum_{\iota=1}^{J} V_{\iota t}}\right) \text{ (49)}$$

Where j is the population of neighborhoods.

$V_{\iota t} = \sum_v e_{\iota v t}$ is the sum of exposure to flood and scarcity at time t, where V_ι , $\iota = 1$ to J are the values of the vulnerability index of neighborhoods, indexed in increasing order ($V_\iota < V_{\iota+1}$). The inequality index is then the average of the gini coefficient over the last yearly cycles of decision:

$$GINI = \sum_{t=T_{Max}-T_y^D}^{T_y^D} gini_t \text{(50)}$$

Sensitivity to policy changes

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

$$CV_{jT_{max}} = \frac{\sigma_{jT_{max}}}{\mu_{jT_{max}}}$$

Where $\mu_j$is the mean in exposure indicator E_ε in neighborhood j at the end of a set of simulations, and σ_j is the standard deviation of the sampler of simulations. Thus, if neighborhood a) has higher coefficient of variation than neighborhood b, we say that neighborhood a) is more sensible to changes in policy.

Appendix

Extensions

The extensions needed to run a run of decisions are the gis extension, which allow to load the geo-information. Also, the csv-extension is needed to allow the read of coma separated values into Netlogo variables. Matrix-extension is used to update the priorities (weights) of neighborhoods when HNP are used.

Generating classes of neighborhoods using a k-mean clustering classification

GIS pre-processing information

Instructions to create new type of agents and actions

Create actions as agents

Set value variables new agents

Tables

Table 1: Attributes of the neighborhoods of Mexico City.

Variable	description	Symbol	name variable in netlogo
Neighborhood ID	Numerical key to identify each neighborhood	j	ID
CVEGEO	Unique National identifier	-	CVEGEO
Estate	Administrative units	$\wp_j$	CV_estado
municipality	units	$\mathcal{M}_j$	CV_mu
aquifer	The aquifer that is below the neighborhood	Z_j	Aquifer
Aquifer Zone	The zone of the aquifer below the neighborhood	ζ_j	zona_aquifera
Altitude	Meters above see level	h_{jv}	altitude
group cluster ID	A classification of the neighborhoods of MC based on socio-economic and environmental similarities (appendix)	κ_j	group_kmean
Weekly water supply	Days in a week water was supplied by system v	s_{jvt}	scarcity
Flooding	Number of flood events per year	f_{jT}	flooding
Protests	1 if a protest occur in a week; 0 otherwise	∂_{jt}	protests
Health	Annual incidences of gastrointestinal diseases	g_{jT}	salud
Social pressure	Number of protests per year	θ_{jT}	Presion_social
Media pressure			
Wells	Set of wells within a neighborhood	q_{jt}	Pozos_neighborhood
Infrastructure coverage	% houses connected to infrastructure system v	η_{jvt}	Houses_with_v
Age Infrastructure	Age infrastructure system v	λ_{jvt}	Antiguedad-infra_v
Condition	An index of the state of the infrastructure system v in neighborhood j	c_{jvt}	condicion
Hydraulic cost	Average volume of water per unit of time received by the sewer system in neighborhood j in a year	o_j	Gastro-hydraulico
Capacity	Index of the capacity of the pipes of system v	ε_{jv}	capasidad
Rainfall	A total annual rainfall in neighborhood j	r_{jT}	precipitation
Subsidence	Rate of subsidence per year	ψ_{jT}	hundimientos
Income-Index	The purchase power by neighborhood	I_j	Income-index
Potable water	Volume of water supplied to the neighborhood j by system v at time t	$W_{j v_p t}$	Water-in

Commented [AB2]: Better name for this. In Spanish is gasto hydraulico.

Garbage	Garbage produced in each neighborhood	θ_j	garbage
Water quality	Index of the quality of potable water	φ_j	water_quality
Urban growth	Percentage of neighborhood considered as urbanized	ϖ_{jt}	urban_growth
Water deviated	Number of wells multiplied by the days without water in neighborhoods with indigenous communities	ζ_{jt}	desviacion_agua

Table 2 Actions of actors water authority and residents

	Actions	*symbol*	*Definition*	*Attribute changed*	*Decision cycle*	*Action weight*
Water Authority	Distribution	k_W	Refer to the water distributed by government trucks to areas without connection to the supply network	s_{jt}	Weekly	
	Extraction	k_E	Increase extraction by increasing number of wells	q_{jt}	Annual	
	Maintenance	k_R	Repairing infrastructure system *v*	λ_{jvt}	Monthly	
	New infrastructure	k_N	Refer to the action of providing new infrastructure in neighborhoods that lack of coverage	η_{jvt}	Monthly	
Residents	Collective action	k_C	Refer to the action of organizing to demand changes from the authority or to internally generate change locally	TBD	TBD	
	Water-storage	k_r	refer to the capture of water using storage devises	S_s	TBD	
	Purchase of water	TBD	refer to the action of buying water from private sources	TBD	TBD	
	House modification	k_M	Action to modify local condition of dwelling to reduce damage from flooding	S_f	Annual	
	Mobilizations/protests	k_∂	To express dissatisfaction with the public services of water delivery and sewage	∂_{jt}	Monthly	
	Water Re-uso		the action of recycling and re-using water	TDB	TDB	

Table 3: Criteria residents

CRITERIA FOR CALCULATING DECISION METRIC OF RESIDENTS

CRITERIA	Description	Attribute associated

URBANIZATION	Percentage of area urbanized (driver of change from simulations of urban growth model)	ϖ_{jt}
WASTE OF WATER	dummy	
WATER DEVIACIÓN	The perception of local people that live close to wells that water is being distributed to other neighborhoods	
SERVICE EFFICIENCY	related to efficiency of infrastructure system	
INSUFICIENT INFRASTRUCTURE	Represented by the percentage of population in each neighborhood not connected to sewer system	$(1-\eta_{jvt})$
CONTAMINACIÓN DE AGUA/WATER QUALITY		φ_j
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	s_{jt}
FLOODING	Number of flooding events per year	f_{jt}
HEALTH RISK	Number of incidences per pear	g_{jt}

Commented [AB3]: Pueblos originarios?

Table 4 criteria water operator supply

Criteria for calculating decision metric water supply operator			
		definition	unit
Infraestructura	Age of infrastructure	Average age of infrastructure per neighborhood	years
	Capacity	Capacity [in length of pipes] of the infra to supply water or to discharge	Mts/area
	Failures	An index of the number of infrastructure related hazards per year (e.g. pipes break)	[events/year]
	Lack of infrastructure	The lack of supply and discharge connections	% houses not connected to infra 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 v_P or v_T,	[weeks/year]
	Flooding	Represented by the number of flood events per year	[events/year]
	Health	Represented by the number of incidences per year	[events/year]
Socio-institucional	Supply	Represented by the requirements of the population to infrastructure system	[pop * need/pop.]
	Petitions from Neighborhoods	Demand of population funneled by politician at the level of the municipality. It represents the collective level of response from each municipality	
	Social pressure	Represented by the number of protests per year	[pop * need/person]