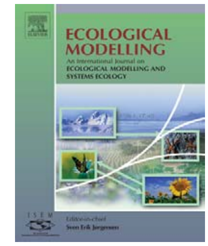
ECOLOGICAL MODELLING 195 (2006) 114–128

available at www.sciencedirect.com

SCIENCE DIRECT®

journal homepage: www.elsevier.com/locate/ecolmodel



Modelling population and community dynamics with qualitative reasoning

Paulo Salles[a,*], *Bert Bredeweg*[b]

[a] *University of Brasília, Biological Sciences Institute, Campus Darcy Ribeiro, Brasilia-DF 70.904-970, Brazil*
[b] *University of Amsterdam, Human Computer Studies Laboratory, Kruislaan 419 (matrix I), 1098 VA Amsterdam, The Netherlands*

ARTICLE INFO

Article history:
Published on line 6 January 2006

Keywords:
Ecological modelling
Qualitative reasoning
Population and community dynamics

ABSTRACT

Ecological knowledge has been characterised as incomplete, fuzzy, uncertain, sparse, empirical, and non-formalised. It is often expressed in qualitative terms, verbally or diagrammatically. There is a need for new and efficient computer-based tools for making this knowledge explicit, well organised, processable, and integrated with quantitative knowledge. Qualitative reasoning is an area of artificial intelligence that creates representations for continuous aspects of the world to support reasoning with little information. Particularly relevant for our work are qualitative representations of differential equations, (in-)equality reasoning and the explicit representation of causal relationships between quantities. We present qualitative models and simulations of population and community dynamics in the Brazilian Cerrado vegetation. The models are organised in clusters of predictive simulation models. The first cluster implements a general theory of population dynamics, with the explicit representation of processes such as natality, mortality, immigration, emigration, colonisation, and population growth. These models are the basis for more complex community models. The second cluster represents interactions between two populations, such as symbiosis, competition, and predation. The third cluster represents the Cerrado succession hypothesis, a commonsense theory of succession in the Cerrado vegetation. It is assumed that fire frequency influences a number of environmental factors. When fire frequency increases succession leads to open grasslands and when fire frequency decreases the vegetation becomes woody and denser. This article shows the potential of qualitative reasoning for ecological modelling, particularly for answering questions of interest and making scientifically valid predictions using only qualitative terms.

1. Introduction

Computer-based simulations are becoming more important for ecological modelling, as computational resources are becoming widely available. A number of textbooks published recently (Gillman and Hails, 1997; Haefner, 1996; Jørgensen and Bendoricchio, 2001) demonstrate that ecological modelling has thus far been almost synonymous with building mathematical models. However, this approach may not be suitable for representing much of the available ecological knowledge. Ecological knowledge can been characterised as incomplete, fuzzy, uncertain, sparse, empirical, and non-formalised. It is often expressed in *qualitative* terms, verbally or diagrammatically.

Qualitative reasoning (QR) is an area of artificial intelligence (AI) that is concerned with the construction of *knowledge*

* *Corresponding author*. Fax: +55 61 3272 1497.
E-mail addresses: psalles@unb.br (P. Salles), bredeweg@science.uva.nl (B. Bredeweg).
0304-3800/$ – see front matter © 2005 Elsevier B.V. All rights reserved.
doi:10.1016/j.ecolmodel.2005.11.014

models that capture insights that domain experts have of systems and their behaviour. Most of the research on qualitative reasoning deals with physics and engineering (Bredeweg and Struss, 2003; Weld and de Kleer, 1990). Qualitative reasoning claims to provide alternative approaches to modelling system behaviour and may be of interest for a number of reasons. One reason is that quantitative data may be missing, which may hamper the applicability of a pure mathematical modelling approach. Another reason is that a qualitative interpretation of system behaviour is essential in order to know what kind of equations to apply. Yet another reason is that mathematical models do not relate well to notions of causality and do not explicitly capture the structural details of the system being modelled. This makes it difficult to interpret a model, explain it or compare alternative models.

It is interesting to note that researchers from the domain of ecology refer to similar issues. For instance, Rykiel (1989) points out that knowledge expressed in qualitative terms may be useful for many purposes. Many questions of interest in ecology (especially for decision-makers) can be answered in qualitative terms (for example, using notions such as worse/better, sooner/later, etc.). Also scientifically valid qualitative predictions can be made when quantitative predictions cannot. Rykiel concludes that there is a need for new and efficient computer-based tools for making this knowledge explicit, well organised, processable on computers, and (the main challenge) integrated with quantitative knowledge. Gillman and Hails (1997: p.4), when talking about population dynamics, argues that ecological modellers produce 'realistic' models "in the sense that they closely mimic what we believe is happening in the field. However, we do not really know why these models produce the answer they do (...). We are often no nearer to the mechanism(s) driving the dynamics."

Following the claims from the qualitative reasoning community, we argue that pure mathematical models fail to capture important aspects of a system's behaviour. These models do not represent knowledge about: (a) objects and associated quantities that describe their characteristics; (b) situations in which these objects occur; (c) descriptions of mechanisms of change grounded in causal relations between the quantities; (d) the assumptions the modeller adopted in order to provide a coherent perspective of the phenomena being modelled. Qualitative models, on the other hand, do capture this *conceptual* knowledge. Being explicitly represented in models, this knowledge can be inspected, possibly modified, by users and by other modellers. The construction of such *articulate* models is of particular interest for education and training, because they facilitate 'knowledge communication' between the agents involved in the learning process (Bredeweg and Forbus, 2003; Bredeweg and Winkels, 1998; Falkenhainer and Forbus, 1991).

In this article, we present implemented qualitative models and simulations of population and community dynamics based on qualitative representations of both theoretical and commonsense knowledge about populations and communities. Some of these models have been described elsewhere (Salles and Bredeweg, 1997; Salles et al., 2003) and follow the principles of constructing *articulate* representations as mentioned above. However, here we focus on the qualitative ecological knowledge captured by these models and not on their educational aspects (see e.g. Bouwer and Bredeweg, 2001; Salles and Bredeweg, 2001; Salles et al., 1997).

The organisation of this article is as follows. Section 2 summarises important characteristics of qualitative models and simulations, particularly focussing on the features relevant to our work. Section 3 discusses the ecological domain knowledge that is captured by the qualitative models presented in this article. Those models are then discussed in Section 4, which is divided into three parts, discussion single population behaviour, interacting populations and community behaviour, respectively. Section 5 highlights related work, focussing on alternative model-building efforts in ecology using qualitative representations. Section 6 concludes the article and stresses the potential of qualitative models for ecological modelling.

2. Qualitative modelling and reasoning

The models presented in this paper are fully implemented in GARP[1] (Bredeweg, 1992), a domain-independent qualitative reasoning engine based on traditional approaches to qualitative reasoning such as the component-based approach (de Kleer and Brown, 1984) and the process-based approach (Forbus, 1984). In this section we discuss some typical characteristics of qualitative models and simulations and how they are realised in GARP. For a detailed discussion on this branch of AI research see the original publications (Weld and de Kleer, 1990).

A qualitative reasoning engine takes a scenario as input and produces a state-graph capturing the qualitatively distinct states a system may manifest (Fig. 1). A scenario usually includes a structural description of the system. Such a description models the entities (e.g. physical objects and components) that the system consists of, together with statements concerning the structural organisation of these objects (e.g. a container *containing* a liquid). Often a scenario also includes statements about behavioural aspects such as relevant quantities, their initial values, and inequality statements between some of those quantities.

A state-graph consists of a set of states and state-transitions. A state refers to a qualitatively unique behaviour that the system may display (a 'possible state of behaviour'). Similar to a scenario, a state consists of a set of declarative statements that describe the structure of the system and the behaviour it manifests at that moment. The latter is characterised by a set of qualitative values about magnitude and direction of change of relevant quantities. A state-transition specifies how one behavioural state may change into another. A sequence of states, connected by state-transitions, is called a behaviour path, but is also referred to as 'a behaviour trajectory' or 'a possible behaviour' of the system. A state-graph usually captures a set of possible behaviours, because multiple state-transitions are possible from certain states.

To construct a state-graph, the reasoning engine uses a library of predefined partial models. These model fragments represent chunks of domain knowledge and, depending on the

[1] The software and models can be downloaded from: http://hcs.science.uva.nl/QRM/.

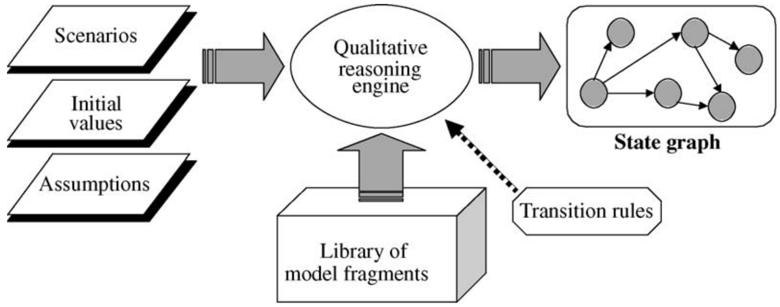


Fig. 1 – Basic architecture of the qualitative reasoning engine.

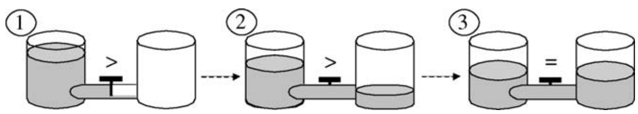


Fig. 2 – Two-tank system: an illustration of the behaviour observed in the real-world.

scenario details, subsets of these fragments are assembled by the engine. Below, the principles of qualitative reasoning are further explained using a two-tank system as an example.

Fig. 2 illustrates the behaviour of a two-tank system as it is observed in the real-world. One tank is filled with oil and the other is empty (1). The tanks are connected via a pipe (a fluid path) and after opening the valve the oil flows from the former into the latter (2) until the liquids in both tanks reach an equal height (3).

Fig. 3[2] shows part of the structural model as it appears in the scenario and in the states of the state-graph. Important representational primitives of this description are the *entities* that the system consists of (e.g. 'tank_left'), the *structural relations* between entities (e.g. 'contains'), and *structural attributes* of entities (e.g. 'openness'). The figure also shows the subtype hierarchy of the entities used by the qualitative simulator to determine the applicability of model fragments to a particular scenario (this is further discussed below).

The state-graph (with a single behaviour trajectory) as produced by the simulator on the basis of a qualitative model of the two-tank system is shown in Fig. 4. It depicts the scenario (grey, referred to as 'input') and the three qualitatively distinct states of behaviour (black-numbered circles). Each of these icons refers to a set of declarative statements representing structural and behavioural aspects of the two-tank system. Notice that, time is captured by the sequence of states that appear during the simulation.

Important aspects of the behaviour description are the *quantity values* in each state (Fig. 4) and the *causal* model (discussed below, see Fig. 5). A quantity value is represented as the pair ⟨Magnitude, Derivative⟩. Magnitude represents the amount of a quantity and the Derivative represents the direction of change over time. The values a Magnitude can take on are represented in a quantity space (QS). For instance, the amount of substance in a container can be represented as having three possible magnitudes: QS = {zero, plus, max}, respectively, meaning there is no substance, there is some substance, and the amount of substance in the container has its highest possible value: maximum. Values for the Derivative are also represented by a quantity space, namely QS = {min, zero, plus}, meaning the Magnitude is decreasing, steady, and increasing. Thus, if amount has the value *amount* = ⟨*plus, plus*⟩ this can be read as: there is an amount and in the current state it is increasing. Fig. 4 shows, among others, that the quantities *amount*, *level*, and (bottom) *pressure* for the oil on the right-hand side have the value ⟨*zero, plus*⟩ in state 1. Next, in state 2 these quantities have a positive value and are increasing ⟨*plus, plus*⟩. In state 3, they are steady ⟨*plus, zero*⟩. A sequence of quantity values is sometimes referred to as a value-history.

Determining the relevant quantity space for each quantity is an important aspect of constructing a qualitative model because it is one of the features that determines the variety of possible behaviours that will be found by the simulator when the model is simulated. Inequality statements (e.g. *level* oil_left > *level* oil_right) are also important in this respect. In fact, each qualitatively distinct state of behaviour is defined by a unique set of values and inequality statements. Transitions between behavioural states are the result of changes in these values and inequality statements. State-transitions are shown in the state-graph as arrows connecting the circles (Fig. 4). For example, while going from state 1 to 2, the magnitudes of *level*, *amount* and *pressure* (for oil_right) change from zero to plus. Going from state 2 to 3 the oil *levels* in the two tanks become equal (not shown in Fig. 4, but partly shown in Fig. 5) and the *flow* becomes zero. In addition, the *level*, *amount* and *pressure* for both tanks stop changing ($\partial = 0$). Notice that the concept of a qualitative state is a bit subtle in that it also represents 'changing' behaviour (and not only equilibriums). For instance, a contained liquid increasing in level is a qualitative state of behaviour (thus: *level* = ⟨*plus, plus*⟩).

An important aspect of a qualitative model is the notion of causality. Forbus (1984) points out that we use causality to impose order upon the world. When we think that '*A* causes *B*', we believe that if we want *B* to happen we should bring about *A*, and if *B* happens, then *A* might be the reason for it. Causality can also result from indirect influences: '*A* causes *C* indirectly' if '*A* causes *B*' and '*B* causes *C*.' In his qualitative process theory (QPT) Forbus introduces the *sole mechanism assumption* which defines that all changes in physical systems

[2] The diagrams are generated by VISIGARP, a graphical user interface on top of GARP that can be used to run and inspect models and simulations. For details see Bouwer and Bredeweg (2001).

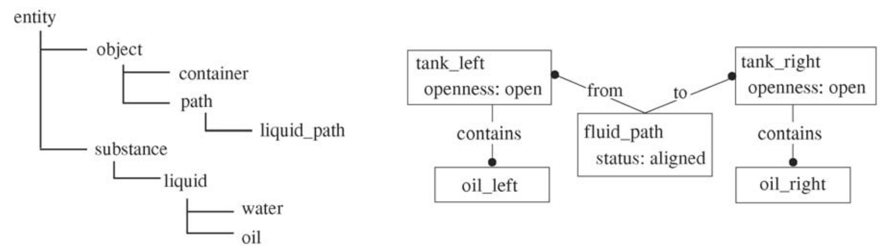


Fig. 3 – Two-tank system model: entity subtype hierarchy (left) and structural model (right).

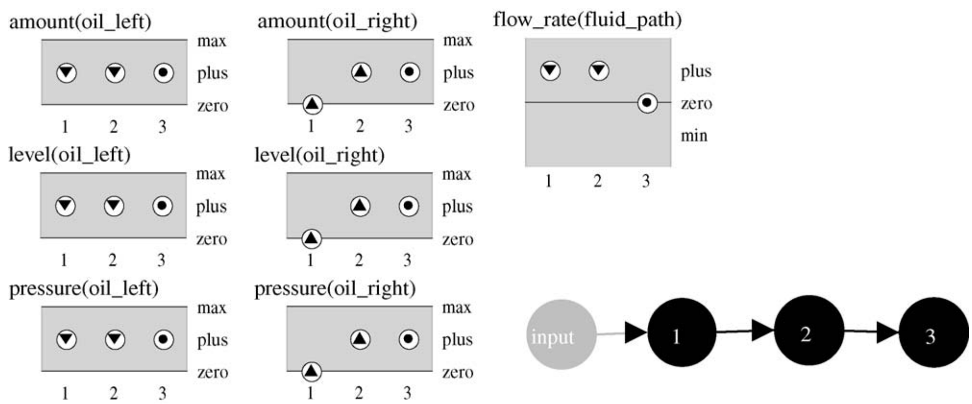


Fig. 4 – Two-tank system model: value-history and state-graph (right-low).

are directly caused by processes (modelled as {I+, $I-$}) or indirectly, by the propagation of their effects (modelled as {P+, $P-$}, and referred to as qualitative proportionalities).

The *causal directness hypothesis* poses certain limits on how direct and indirect influences have to be applied. First, all changes initiate with a process (and thus with a direct influence) and propagate through the whole system via proportionalities (indirect influences). Second, both direct and indirect influences are directed. Their effects propagate only in one direction. Thus, if $P+(B, A)$ holds, then A causes changes in

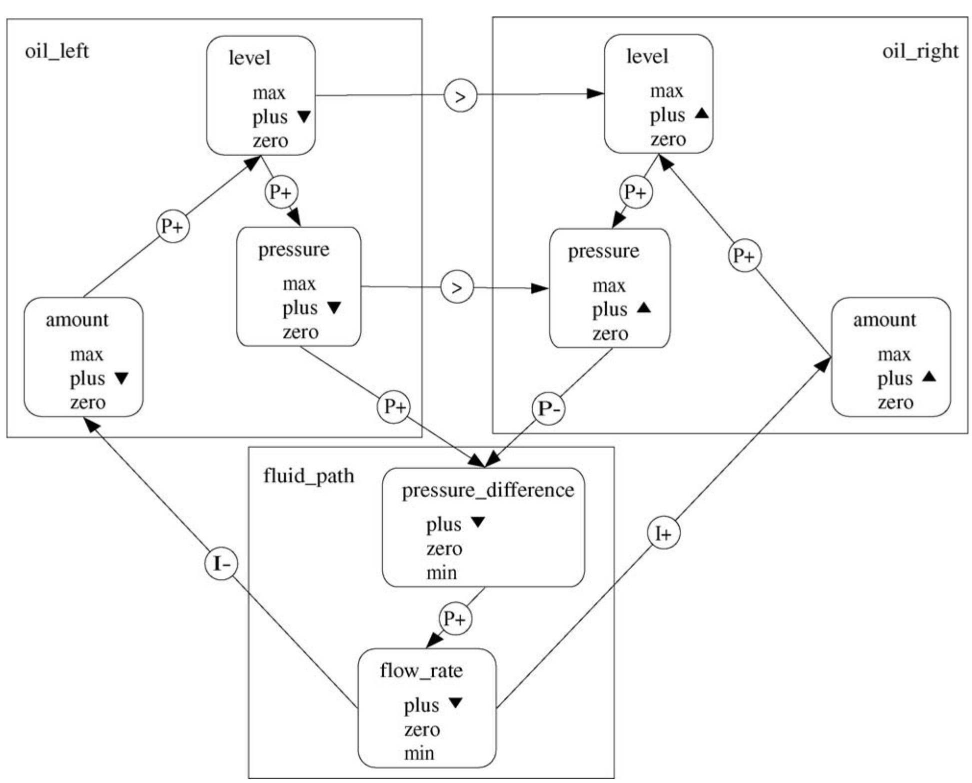


Fig. 5 – Two-tank system model: causal model (as active in state 2).

B and not the other way around. Third, no quantity may be influenced directly and indirectly at the same time, because it would violate the fundamental chain of causality.

Both direct influences and qualitative proportionalities are modelling primitives that express causal relationships between quantities, and have mathematical meaning. Direct influences determine the value of the derivative of the influenced quantity. For example, the relation $I+(Y, X)$ means that $dY/dt = (\ldots + X \ldots)$. Qualitative proportionalities carry much less information than direct influences. For example, the relation $P+(Y, X)$ means that there is some function (f) that determines Y, and is increasing monotonic in its dependence on X, such that $Y = f(\ldots X \ldots)$ and $dY/dX > 0$. A quantity that is not influenced by any process is considered to be constant (Forbus, 1984).

Notice that a single direct or indirect influence statement does not determine, by itself, how the quantity it constrains will change. Its effect must be combined with all the active influences on that quantity. Ambiguities may arise when positive and negative influences are combined. In these cases, the reasoning engine either considers all the possible combinations or any explicitly represented assumption that may constrain the system's behaviour. The mechanism of combining all the direct or indirect influences acting on the same quantity is called *influence resolution* (Forbus, 1984) and is further discussed in Section 4.

Summarising, a qualitative reasoning engine constructs a state-graph using representations of entities, structural relationships and attributes, quantities with value statements (including magnitude and derivative) and dependencies between quantities. The latter includes inequalities (used as constraints) and causal relationships (influences, qualitative proportionalities and correspondences). The set of dependencies is of particular importance to determine the specific behaviour in each state. Fig. 5 shows a subset of the dependencies that hold in state 2 of the simulation of the two-tank system model. Such a set of dependencies is often referred to as the *causal model*. The diagram shows that the two oils have unequal levels and pressures. The flow between the two oils depends on the difference between those pressures (is qualitative proportional to it) and has a negative influence on the oil with the higher pressure and a positive influence on the other, decreasing and increasing the two amounts of oil, respectively. Changes in the amounts propagate to changes in levels, which in turn change the pressures. Notice that this diagram also shows the quantity space for each quantity, the current value, and the direction of change. The latter is visualised by triangles pointing up (increasing), or down (decreasing), and by small black circles (steady) (as no quantity is steady in state 2, circles are not shown in Fig. 5). The direction of change icon is placed adjacent to the current value of the quantity, highlighting the latter in the context of its quantity space. For instance, the figure shows that, for the oil_left, *level* = ⟨*plus, min*⟩).

Due to space limitation, Fig. 5 does not show the active correspondences. Correspondences specify co-occurring values. For instance, if the magnitude of the amount of oil is zero, then the magnitude of the level must also be zero. Correspondences can be defined between specific values of two quantities (value correspondence) or between *all* the values of two quantities (quantity space correspondence), and may be directed or undirected. Quantity space correspondence requires that the two quantities have the same type of quantity space.

Some qualitative simulators take a fixed model structure (a set of constraints between quantities) to run a simulation (e.g. Kuipers, 1986). However, GARP implements a compositional-approach and automatically assembles the dependencies by consulting a *library* of model fragments (MFs). This implies, among other things, that the set of dependencies may change and can be different for alternative states. MFs can be seen as re-usable conditional statements that capture knowledge about the phenomena existing in a certain domain. They can be constructed using the representational primitives discussed above. In general a MF requires certain structural details to be true (e.g. a *container*, a *liquid* and a *contain* relation between these two entities). If the required structure exists the MF is instantiated for that structure and introduces the behaviour details that apply to it (e.g. the quantities *amount, height, pressure* and the *dependencies* that hold between them). Given a scenario a specific MF can be instantiated multiple times, namely for each occurrence of the structure to which it applies. A properly designed MF is context independent, obeying the *no-function-in-structure* principle (de Kleer and Brown, 1984). This means that a specific MF can be re-used in different models, because the laws defined by a MF do not presume the functioning of the system as a whole.

The causal model shown in Fig. 5, for state 2 of the two-tank system, was assembled by the simulator from two MFs: 'contained liquid' and 'liquid-flow' (see Table 1). The former was applied two times (namely for each container containing oil) and the latter was applied only once (namely for 'two contained liquids,' connected by a flow-path, and with a pressure difference). Notice that the conditional part of a MF may require other MFs to exist. It may also require behavioural facts to be true. For instance, in the case of the 'liquid-flow' the statement '*pressure* oil_left > *pressure* oil_right' must be true before it may be applied. For a detailed description of the two-tank system see Forbus (1984).

Given a sufficiently well-developed library for a certain domain the qualitative reasoning engine can predict the behaviour of all kinds of systems belonging to that domain. GARP performs this task using two basic inference steps. Given a scenario, it will examine the library of MFs. Each unique set (and interpretation) of MFs that is consistent with the scenario leads to a distinct state of behaviour. States may represent 'changing' system behaviour and thus each state is searched for possible terminations. For instance, increasing temperature of a liquid (current state) may reach its boiling point (a possible next state). State-transitions are found using domain-independent rules (for instance: an increasing quantity may reach the next higher magnitude in its quantity space). Possible successor states are again examined determining which MFs still apply and which new ones much be added given the transitions that have taken place. Constructing a state-graph is also referred to as creating an 'envisionment' (de Kleer and Brown, 1984).

A fundamental aspect of building a qualitative model is the construction of a library of MFs for a certain domain that can be used to reason about the behaviour of a set of systems belonging to that domain. In the past, considerable effort

Table 1 – A schematic representation of the model fragment 'liquid-flow'

```
liquid_flow(Path, ContainerA, ContainerB) isa process
  conditions:
    entities and relations:
      instance(Path, fluid_path)
      has_attribute(Path, status, aligned)
      has_attribute(Path, connected, ContainerA)
      has_attribute(Path, connected, ContainerB)
    quantities:
      bottom_pressure(LiquidA, PressureA, zpm)
      bottom_pressure(LiquidB, PressureB, zpm)
    quantity values:
    dependencies:
      PressureA > PressureB
    model fragments
      open_contained_liquid(ContainerA, LiquidA)
      open_contained_liquid(ContainerB, LiquidB)
  consequences:
    entities and relations:
    quantities:
      amount(LiquidA, AmountA, zpm)
      amount(LiquidB, AmountB, zpm)
      flow_rate(Path, Flow_rate, mzp)
      pressure_difference(Path, Press_difference, mzp)
    quantity values:
    dependencies:
      Press_difference = PressureA − PressureB
      Flow_rate = Press_difference
      prop_pos(Flow_rate, Press_difference)
      q_correspondence(Flow_rate, Press_difference)
      inf_neg_by(AmountA, Flow_rate)
      inf_pos_by(AmountB, Flow_rate)
      prop_pos(Press_difference, AmountA)
      prop_neg(Press_difference, AmountB)
```

has been put in building qualitative models for the domain of physics, but libraries for other domains still need to be developed. In this article we present qualitative models and simulations for population and community dynamics.

3. Aspects of the Brazilian Cerrado vegetation

The Cerrado biome is characterised as a tropical savannah because of the presence of an almost continuous and well-developed grass layer, and a discontinuous layer of trees and shrubs. It covers almost two million square kilometres in the central region of Brazil, where the climate is tropical with a well-marked dry season between May and September, and a wet season between October and April. The average annual rainfall ranges between 1100 and 1600 mm, 90% of which falls during the wet season. In most areas, soil is strongly or moderately acid, poor in nutrients and with high aluminium concentrations.

Cerrado holds great biological diversity and consists of many well-defined groups of species that occur together (physiognomies). Cerrado communities vary from open grasslands to closed forests, and have been studied by several researchers, for example: Coutinho (1990); Eiten (1972); Goodland and Ferri (1979); Moreira (1992); Ribeiro and Walter, (1998). In our work we consider five types of Cerrado communities, called in Portuguese *campo limpo*, *campo sujo*, *campo cerrado*, *cerrado sensu stricto*, and *cerradão*. They can be organised in a gradient according to the quantities of trees, shrubs, and grass in each community. The gradient goes from *campo limpo* (open grassland) to *cerradão* (dense forest). For the intermediate community types, trees and shrubs are increasing and grass and herbaceous are decreasing (see Goodland and Ferri, 1979 for a detailed study of this gradient).

Five factors typically determine the Cerrado vegetation: soil nutrient availability, soil moisture availability, herbivory, fire, and human actions (Moreira, 1992). Nutrients and water in the soil are the primary determinants. While herbivory do not have a great impact on the dynamics of the Cerrado, fire is an important influence on the vegetation (see Miranda et al., 1996, a collection of papers on this issue). Lightning is the main natural cause of fire in the Cerrado, but human actions are the main cause of fire. In the last 40 years, the impact of human actions on the Cerrado increased with the occupation of the central region of Brazil. As a consequence, deforestation and burning are becoming more frequent.

It is accepted that fire has a positive influence on grass and a negative one on trees. These differences are important for understanding population and community dynamics in the Cerrado. For example, it has been shown that, after burning, grass species are able to resprout quickly and therefore occupy the bare ground, out-competing woody species whose seedlings are sensitive to fire (Morosini and Klink, 1997).

Fire *frequency* is more relevant for changes in Cerrado structure than the *intensity* of isolated fire events. It is accepted that, on the one hand, if fire frequency is high, the vegetation becomes open, with more grass and fewer trees, and changes towards the *campo limpo* side of the gradient. This situation involves a positive feedback loop. The main components of the fuel in Cerrado are grass leaves and small pieces of wood, and communities such as *campo limpo* and *campo sujo* are rich in these components. Thus, increased fire frequency creates conditions for fire to become more frequent. On the other hand, vegetation in protected areas in which fire frequency is reduced tends to become woody and denser. These conditions favour the decrease of fire frequency.

Considering that the communities included in the gradient mentioned above can be seen as different successional stages of the vegetation, with *cerradão* being the local climax, it is possible to predict that a protected area of *campo* may evolve to become a *cerradão*. On the other hand, increasing fire frequency may change a *cerradão* or a *cerrado sensu stricto* into open grassland such as *campo sujo* or *campo limpo*.

Long-term studies are required to confirm that fire frequency indeed changes the physiognomy of the vegetation according to these successional stages, but available knowledge shows that this may be true in some Cerrado areas and under certain circumstances. Meanwhile, this hypothesis is part of commonsense theories about the dynamics of the vegetation, shared also by scientists, and provides a theoretical framework for teaching, training and management (Salles, 1997). In Section 4.3 we present an implemented qualitative model of this 'Cerrado succession hypothesis' (CSH). Simulations with this model explore the structure and behaviour of Cerrado communities under the influence of fire and other environmental factors.

4. The models and simulations

This section describes knowledge that supports reasoning about population dynamics, the *first principles* upon which we create representations of the physical world. The core of this model is a set of MFs about the dynamics of a 'single population' (Section 4.1). It includes knowledge about the key processes, such as natality, mortality, immigration, emigration, growth, colonisation and views about the types of biological entities, population size, and the required assumptions. In Section 4.2 this knowledge about the behaviour of a 'single population' is used as the basis for more complex models, particularly to generate and explain the behaviour of interacting populations. In Section 4.3 this is again taken a step further when it is used to generate and explain the Cerrado succession behaviour.

4.1. *Single population models*

The domain theory about population dynamics implemented in the models presented here is a qualitative reading of the basic equation:

$$\mathrm{Nof}_{(t+1)} = \mathrm{Nof}_{(t)} + (B + Im) - (D + E)$$

in which Nof stands for the number of individuals of a population in the beginning and in the end of the time interval, and *B*, *D*, *Im*, *E* are, respectively, the amount of individuals being 'born,' 'dead,' 'immigrated' and 'emigrated' during that period of time. These four quantities are the rates of the natality, mortality, immigration and emigration processes. Following the ontology provided by QPT, the representation for the four basic population processes and their effects on Nof would be:

$$I+(\mathrm{Nof}, B); \qquad I-(\mathrm{Nof}, D); \qquad I+(\mathrm{Nof}, Im); \qquad I-(\mathrm{Nof}, E)$$

During influence resolution, *B* and *Im* are added to the derivative of Nof, whereas *D* and *E* are subtracted from it. The final result depends on the relative amounts of *B*, *D*, *Im* and *E*. If there is knowledge available about the magnitude of these rates or about their relative size, GARP defines the resultant. If not, it tries all the possible combinations.

The domain theory includes feedback loops that represent the effect that Nof has on *B*, *D*, and *E*.[3] This is obtained by means of qualitative proportionalities:

$$P+(B, \mathrm{Nof}); \qquad P+(D, \mathrm{Nof}); \qquad P+(E, \mathrm{Nof})$$

This way, the combination $\{I+(\mathrm{Nof}, B), P+(B, \mathrm{Nof})\}$ reads as 'the amount of individuals being born should be added to the derivative of Nof' and 'when Nof changes (increases or decreases) the amount of individuals being born also changes in the same direction.'

How can we capture the insights described above in MFs? The basic structure of single population models consists of two entities: 'biological entity' and 'population,' the latter being defined as a set of the former. The entity 'biological entity' can be further detailed, creating representations for plants, trees, etc. The entity 'population' is characterised by the quantity Nof, which can take on zero or positive values. Most of the examples presented here use the QS = {zero, normal, max} for the magnitude of Nof (meaning: there is no population (zero), there is a population (normal range of values), and the population has its biggest size (maximum), respectively). Population size (its magnitude) may be increasing, stable or decreasing, modelled as the derivative of Nof. For the direction of change, the quantity space of the derivative of Nof is QS = {min, zero, plus}, meaning decreasing, steady and increasing as in the two-tank model.

The most fundamental distinction to be made in the models refers to existing and non-existing populations. If Nof > zero, there is a population, which is described in the MF 'existing population'; if Nof = zero, the situation is described in the MF 'non-existing population.' If the 'existing population' MF is active, the magnitude of Nof may assume any positive value, whereas the derivative may assume any value from its QS. If the MF 'non-existing population' is active, then Nof = ⟨zero, ?⟩. The question mark represents any of the possible values (zero or plus) for the derivative. Unless a process occurs, it is assumed to be stable, so that MF 'non-existing and stable' sets the value of Nof = ⟨zero, zero⟩.

The process 'natality' introduces the quantity 'born' (*B*), also a continuous variable with QS = {zero, plus} associated to the object 'population.' The MF 'natality' requires as a condition to become active that the MF 'existing population' is already active. That is, the relation Nof > 0 must hold for *B* to be introduced in the simulation, which is the simplest assumption one can make about natality in a population. The MF 'natality' introduces a direct influence, $I+(\mathrm{Nof}, B)$, and the indirect influence $P+(B, \mathrm{Nof})$. Additional knowledge introduced in this MF establishes a *correspondence* saying that if magnitude Nof = zero, then *B* = zero.

As explained above, when the natality process is active, *B* is added to the derivative of Nof. If this is the only direct influence, Nof will increase by an amount that equals the value of *B*. A simulation with this simple model (starting with an initial scenario in which the entities 'biological entity' and 'population' are introduced with Nof = ⟨?, ?⟩ and *B* = ⟨?, ?⟩) produces three initial states: state 1 = {Nof = ⟨zero, 0⟩}, state 2 = {Nof = ⟨normal, +⟩}, and state 3 = {Nof = ⟨max, +⟩}. When we further run the simulation, the only possible state-transition is [2 → 3], because Nof increases from 'normal' to value 'max.' The MF 'non-existing population' does not activate the MF 'natality' in state 1. Therefore state 1 remains as it is (Nof = ⟨zero, zero⟩; *B* = ⟨zero, zero⟩). Similar reasoning can be done with the other three basic processes.

An extension of this qualitative theory of population dynamics is obtained by aggregating quantities in order to get a different perspective on population growth. We define the growth process as a combination of the four basic processes, taking into account their effects on Nof. The quantity *Inflow* is calculated from the addition (*B* + *Im*) and *Outflow* is calculated from the addition (*D* + *E*). Both quantities are then used to cal-

[3] Immigration (Im) is not included in this feedback loop, because it is assumed that Im is exogenous to the system. That is, the amount of inflow resulting from Im does not depend on Nof. Instead, Im is seen as an external factor that is determined outside the scope of the system (it happens to the system). For a discussion on the notion of exogenous variables in qualitative reasoning see e.g. Iwasaki and Simon (1986).

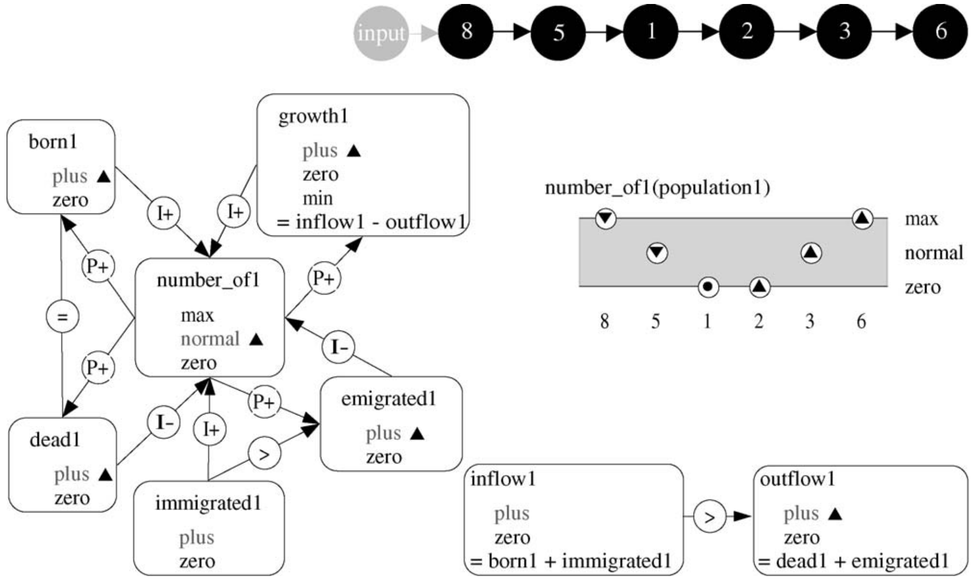


Fig. 6 – Causal model for a single population.

culate the value of the *Growth* rate: *Growth* = *Inflow* − *Outflow*. Along with these qualitative algebraic operations, the MF 'population growth' introduces the pair of causal relations {*I* + (Nof, Growth), *P* + (Growth, Nof)}. Differently from the four basic processes, *Growth* requires QS = {min, zero, plus} to take care of situations in which *Inflow* is smaller, equal or greater than *Outflow*.

Colonisation is modelled as a special case of Immigration. Descriptions of these two processes are very similar, in the sense that they require as conditions for becoming active the presence of the quantities Nof and *Im*. However, colonisation starts a new population where such population does not exist (Nof = zero). Therefore, a condition for the MF 'colonisation' to become active is that the MF 'non-existing population' is already active, whereas the condition for the MF 'immigration' to become active is that the MF 'existing population' is already active. Both MFs represent a process in which the immigration rate (*Im*) directly influences Nof.

Fig. 6 shows part of a simulation that uses the ideas discussed so far. It concerns a situation in which all the six processes are active. The initial scenario mentions two entities: 'biological entity' and 'population,' and introduces Nof and seven other quantities (but no values are assigned to them yet). The initial scenario also mentions the assumptions 'open population' and 'born and dead correspond.' The former is an operating assumption[4] that applies to 'population' and must hold in order to activate MFs 'immigration,' 'emigration' and 'colonisation.' The latter is a simplifying assumption and also applies to 'population.' It introduces a value correspondence between *B* and *D* that holds for all the values in their QS. This way, the simulation is simplified and changes in Nof depend primarily on the relative values of *Im* and *E*.

The simulation produces eight states. All the possible magnitude and derivative values of Nof are combined but one: Nof = ⟨zero, min⟩. Fig. 6 also shows the causal model with direct and indirect influences on Nof in state 3, and a 'value-history diagram' of the quantity Nof. Here the four basic processes are active and the rates *B*, *D*, *Im* and *E* can be seen in each box, with their QS, their current value and direction of change. The figure also shows the direct (*I*+, *I*−) and indirect influences (*P*+, *P*−). Notice that *Growth* is a direct influence whereas *Inflow* and *Outflow* are not influences, but intermediate quantities used for computing the *Growth* magnitude and derivative. Notice that the condition *Inflow* > *Outflow* is also valid in state 3.

The sequence [8 → 5 → 1 → 2 → 3 → 6] of qualitative states shows the longest behaviour path obtained in this simulation. It starts with a maximum sized population (state 8) that becomes extinct (state 1). Next colonisation starts a new population (state 2) that will grow up to its maximum size (state 6).

Colonisation is simulated as a two-step mechanism. First, a steady state with an extinct population (state 1), then a state in which a 'non-existing' population (Nof = zero) receives immigrants so that its derivative becomes positive {Nof = ⟨zero, plus⟩} and colonisation takes place (state 2). In the following state (3) population size no longer has the value zero, and all the basic processes are active (the latter is shown in Fig. 6).

4.2. *Two-population models*

Interactions between populations of different species can be classified either by the *mechanism* or by the *effects* of the interaction. Describing the mechanisms of interaction includes the particularities of each species' life cycle. If these details are left out and just the effects are considered, the interactions can be represented as combinations of the symbols {−, 0, +} so that: (a) '−' means that one population is negatively affected by the

[4] Operating assumptions are concerned with the scope of the simulation, that is, which behaviours to include in the first place. Simplifying assumptions cause the simulator the reason about certain behaviour in less detail. This notion of assumption types is introduced by Falkenhainer and Forbus (1991).

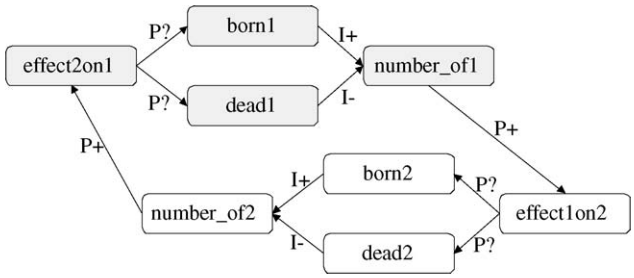


Fig. 7 – Base model for two populations.

other; (b) '0' means that one species suffers no effects from the other population; and (c) '+' means that one population is benefited by the influence of the other species.

Effects of interactions between two populations, as described by Odum (1983), was the basis for the implementation of a set of simulation models of interactions between two populations. We implemented models of neutralism (0, 0), amensalism (0, −), commensalism (0, +), predator–prey (+, −), symbiosis (+, +) and competition (−, −) (see Salles et al., 2003). These signs indicate that a positive or a negative term has to be added to the growth equations. It means introducing a new quantity that will affect the other population's *B* and/or *D* positively or negatively. A base model for capturing interactions between two populations is shown in Fig. 7. Population1 produces some effect (*Effect1on2*), which in turn affects natality (*B2*) and mortality (*D2*) of population2. In the same way, population2 produces an effect on population1 (*Effect2on1*) which influences natality (*B1*) and mortality (*D1*) of population1. These influences are modelled by qualitative proportionalities. Notice that *P?* must be instantiated with one of the signs $\{-, 0, +\}$ to implement a specific interaction pattern. 'closed population' is assumed for both populations and *Im* and *E* are thus included in the model with value zero for their magnitudes and derivatives. As a consequence, the growth equation is redefined, being $Inflow = B$ and $Outflow = D$.

Suppose we want to run a simulation involving two populations that do not interact (neutralism). Given that there are no constraints, Nof can assume all possible combinations of magnitudes and derivatives values, and thus all possible behaviours each population can exhibit (alone) are expected to appear in the simulation. If we run a simulation of neutralism in which the initial scenario introduces instances of 'biological entity' and 'population' and the quantities Nof1 = ⟨normal, ?⟩ and Nof2 = ⟨normal, ?⟩, this results in a full simulation consisting of 25 states representing all the possible behaviour paths, as expected.

However, if any influence creates a dependency between the two populations we expect that some behaviours will be restricted in one or in both of them. That is, only part of the set of possible behaviours can be expressed by one population in the presence of the other. Modelling population interactions is therefore a matter of putting constraints on the attributes of the two populations such that their behaviours express the six effect-based types of interactions. In order to build the structure of each particular interaction model, we created specifications for: (a) defining parameters related to the basic processes in both populations and new parameters for representing the effect of one population on the other; (b) establishing causal links between parameters Nof, *B*, *D* and *Effect*[5]; (c) establishing correspondences between Nof, *B*, *D* and *Effect*; (d) controlling values assumed by Nof, *B*, *D* and *Effect*; (e) representing conditions for existing and non-existing populations.

To illustrate the model-building process and the resulting simulations, let us consider the predator–prey (+, −) model. A MF 'predator–prey' introduces the entity 'predator–prey interaction' and defines population1 as the predator and population2 as the prey. This MF also introduces the quantities *Consumption* (related to the predator population) and *Supply* (associated with the prey population). It is assumed that there is a full correspondence between the QS of quantities *Consumption* and Nof1, and between quantities *Supply* and Nof2 (see also Fig. 8).

A second MF, 'predator–prey interaction,' establishes that increasing *Consumption* increases the mortality of prey population $\{P+(D2, \text{Consumption})\}$, and that increasing *Supply* increases natality and decreases mortality of the predator population $\{P+(B1, \text{Supply}); P-(D1, \text{Supply})\}$. This MF also introduces the following assumptions: (a) the maximum value of *Consumption* is equal to the maximum value of *Supply*; (b) *Supply* should be equal or greater then the *Consumption*; and (c) *Consumption* cannot increase faster than *Supply*.

The non-existence of populations is described by two MFs: 'predator–prey but no predator' and 'predator–prey but no prey.' The former states that there are no constraints on the prey population growth if the predator population is zero. The latter says that if the prey population is zero, then the predator population should also equal zero. As a consequence, it never happens in the simulations that the predator survives without the prey population (although the contrary is possible). This is a biologically acceptable *simplifying* assumption.

Finally, the MF 'predator–prey assumptions' puts constraints on the quantities Nof, *B* and *D* of both populations, in order to simplify the simulations. It is assumed that (a) for the predator population, the derivatives of *B1* and *D1* are equal. Remember that these quantities are both influenced by a positive feedback loop with Nof1 and by the other population, via quantity *Supply*. The current assumption resolves possible ambiguity between these influences by giving the two rates the same direction as the change of *Supply* (thus the three quantities will simultaneously increase, decrease or stay steady), that is, assuming that the external influence is bigger than the predator population's self-control. (b) With respect to the prey population, it is assumed that the derivative of *B2* is smaller or equal to the derivative of *D2*. This way, the derivative of *D2* may increase because of the effect of *Consumption* and therefore becomes bigger than the derivative of *B2*. As a consequence, *D2* itself may become bigger than *B2*. (c) In both prey and predator populations, the quantities *B* and *D* go to zero at the same time. (d) Finally, in both populations the size of *D* can never exceed the Nof. As mentioned above, the former has to decrease if Nof decreases.

A simulation with this model is shown in Fig. 9. Starting with an initial scenario in which both populations have Nof = ⟨normal, ?⟩, four initial states are found. The full simula-

[5] *Effect* is used here to refer to both *Effect1on2* and *Effect2on1*.

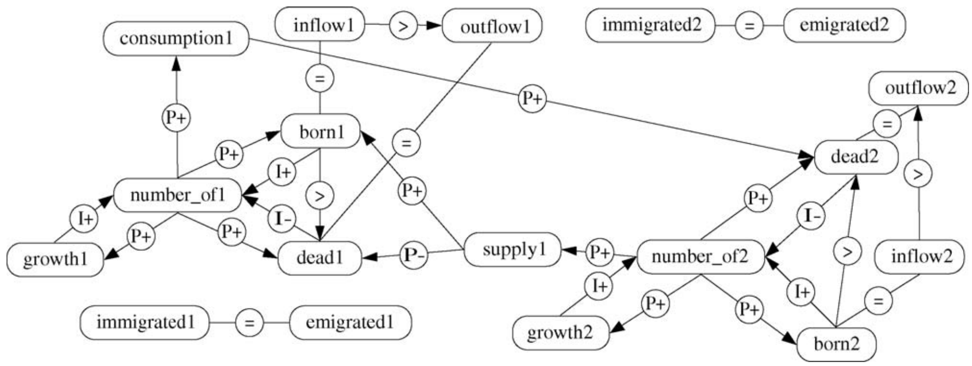


Fig. 8 – Causal model for predator–prey interaction.

tion produces 11 states showing all the possibilities under the set of constraints: coexistence with normal sized populations (state 2), both populations at their maximum size (state10), predator extinct and prey at maximum (state 8) and both populations extinct (state 6). Notice, there is no state of behaviour in which the predator population becomes bigger than the prey population.

4.3. *Models for simulating the 'Cerrado succession hypothesis'*

This section describes the implementation of the 'Cerrado succession hypothesis' (CSH). According to this hypothesis, if fire frequency increases Cerrado communities tend to become less dense and the graminoid layer becomes more important. On the other hand, if fire frequency decreases, the Cerrado vegetation tends to become woody and denser.

The model consists of the library of MFs developed for the one-population models extended with a representation of the Cerrado vegetation. An entity 'cerrado' is defined as consisting of three other entities: 'tree,' 'shrub,' and 'grass.' Each is represented as a 'population,' and there are MFs defining 'tree population,' 'shrub population' and 'grass population.' A set of 13 MFs, hierarchically organised, represents knowledge about the Cerrado and provides the context. The most general MF is 'Cerrado sensu lato,' a composition view that represents the general vegetation. Subtypes of that MF define 12 cerrado community types, each by means of a separate MF. Three main types ('campo,' 'campo cerrado,' 'woody cerrado') are subdivided as follows: {'campo' includes 'campo limpo' (subdivided in 'campo limpo with grass' and 'campo limpo with less grass') and 'campo sujo' (subdivided in 'campo sujo with no tree' and 'campo sujo with tree')}; {'campo cerrado' is not subdivided}; and {'woody cerrado' includes 'cerrado sensu stricto' and 'cerradão' (the latter is subdivided in 'open cerradão' and 'climax cerradão')}.

The CSH model includes three populations, without constraints between them (neutralism). Each represents a functional group of plants (tree, shrub, and grass). The QS adopted for Nof in the three populations is QS = {zero, low, medium, high, max} (or {zlmhm}) to represent existing populations with sizes either in the intervals 'low' and 'high' or equal to the points 'medium' and 'max.' This larger QS increases the representational power, so that each community type can be defined according to the proportion of trees (T), shrubs (S) and grass (G). The hierarchical representation of MFs permits inheritance of knowledge. For example, 'woody cerrado' ($T >$ medium; $G <$ medium) has subtype 'cerradão' ($G =$ zero; $S =$ high; $T >$ medium), which in turn has subtype 'climax cerradão' ($T =$ max). Therefore 'climax cerradão' is a community in which ($G =$ zero; $S =$ high, and $T =$ max).

The MF 'Cerrado sensu lato' introduces relevant environmental factors associated with the entity 'cerrado,' represented by the following quantities: *Litter*, *Moisture*, *Light*, *Soil_temperature*, *Nutrient* and *Fire_frequency*, all associated with QS = {plus}, meaning that they are always present. Another quantity, *Cover*, represents the amount of shade on the ground caused mainly by the canopy of trees. This quantity is related to Nof trees, and uses the same QS = {zlmhm}. In the same MF, these quantities are related to each other by means of qualitative proportionalities ($P+$, $P-$), in order to build the following

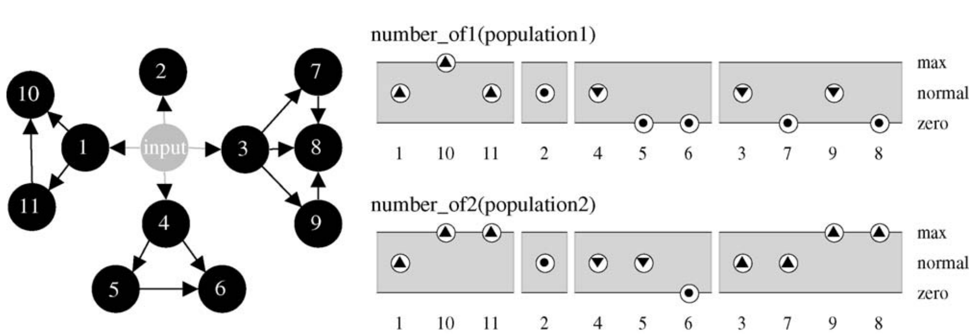


Fig. 9 – State-graph showing predator–prey behaviour.

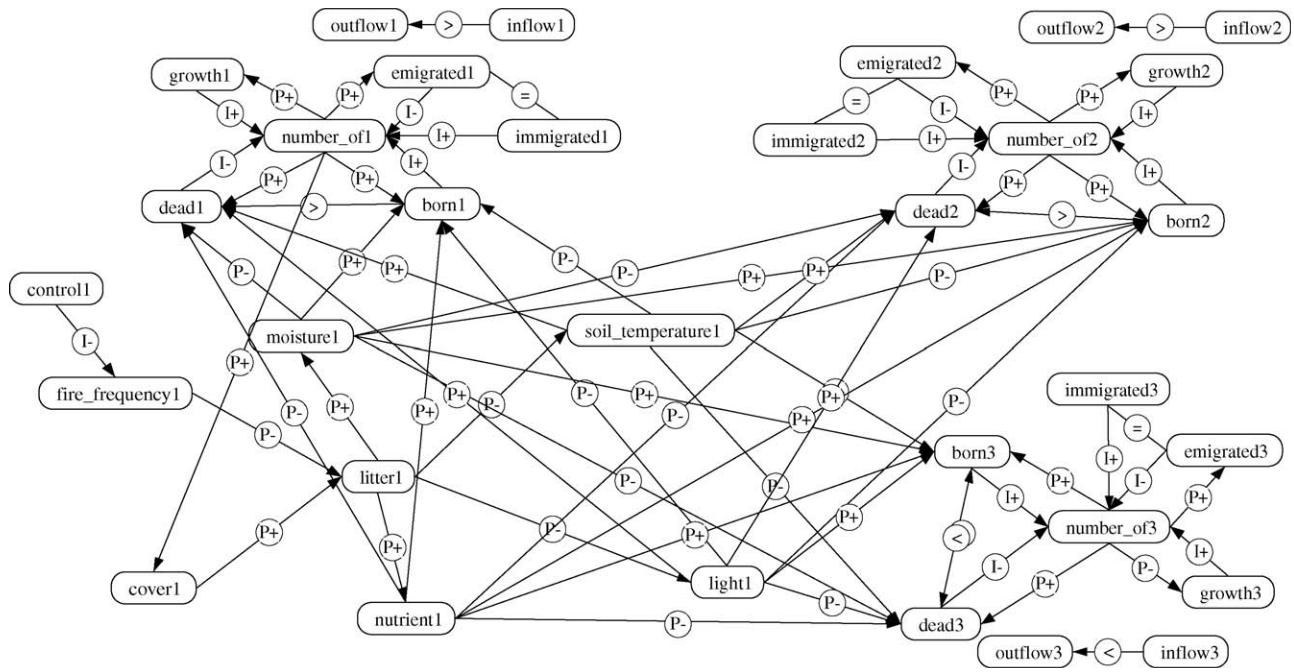


Fig. 10 – Causal dependencies for the Cerrado succession hypothesis.

causal chain:

$$\begin{aligned} \mathit{Fire\_frequency} &\rightarrow \mathit{Litter} \\ &\rightarrow \{\mathit{Moisture}, \mathit{Light}, \mathit{Nutrient}, \mathit{Soil\_temperature}\} \\ &\rightarrow \{B, D, Im, E\}. \end{aligned}$$

A different set of MFs encodes knowledge about how these environmental factors affect the natality and mortality processes in *T*, *S*, and *G* populations. In this case, qualitative proportionalities affect the quantities *B* and *D*, creating the typical behaviour of each population. A part of the causal model is shown in Fig. 10.

Effects of human actions on environmental factors are also shown in Fig. 10. Human actions, in the context of fire control, involve several types of activities and processes. This complex situation is modelled in GARP as a 'compound' process introduced by an 'agent model.' In this case, the agent model puts a direct influence on the quantity *Fire_frequency*.

The agent model includes the entities 'cerrado,' 'manager' and 'fire control' and the quantity *Fire_frequency* (associated to 'cerrado'). The rate of this process is called *Control*, with QS = {min, zero, plus} associated to 'manager.' The *Fire_frequency* is constrained by a direct influence $\{I-(\text{Fire\_frequency, Control})\}$. There are two versions of this agent model, implemented in two MFs: 'decrease fire frequency' (in which *Control* = ⟨*plus, zero*⟩) and 'increase fire frequency' (in which *Control* = ⟨*min, zero*⟩). They read as 'when *Control* has value plus, *Fire_frequency* decreases,' and 'when *Control* has value *min*, *Fire_frequency* increases.' The effect is assumed to be constant during the simulation.

Inspecting the causal model, it is possible to evaluate the consequences of human actions in controlling fire frequency. The influence on the community is indirect — it propagates through the network of environmental factors and finally influences the basic processes of *T*, *S*, and *G* populations. Note that {*T*, *S*} and {*G*} are differently influenced by *Light* and *Temperature*, what is assumed to be ultimately the cause of their different behaviour. Note also that there is a positive feedback loop involving Nof tree and *Cover*.

Simulations with this model show the transitions between the community types caused by control measures on fire frequency as predicted by the hypothesis about the succession in the Cerrado. For example, Fig. 11 shows the state-graph and the values of the state variables in a simulation in which fire frequency decreases and the succession moves towards forest-like communities. Notice that the order of states (1–19) in the value diagram does *not* refer to a behaviour trajectory. Instead, it is just an enumeration of all states. Behaviour trajectories, on the other hand, can be found in the state-graph. They are sequences of states connected by state-transitions (the latter shown as arrows between the black circles).

The initial scenario of the simulation shown in Fig. 11 introduces instances of the 20 objects and entities, and a set of assumptions involving the quantities *B*, *D*, *Im*, *E*, the flows and the community types. There are 32 quantities associated with the objects, and only the magnitudes of Nof trees, shrubs and grass are specified (*T* = zero; *S* = zero; *G* = max). Four states are initially generated for this scenario, states 1, 2, 3, and 4. Given that a 'manager' performs the task 'fire control' which decreases *Fire_frequency*, the derivative of G is negative in all the initial states. These initial states present the conditions for the colonisation process in *T* and/or *S*. The full simulation produces 19 states, ending in a 'climax cerradão' community. A possible behaviour path including the states [1 → 4 → 5 → 12 → 14 → 15 → 18] describes the typical succession. With *Fire_frequency* decreasing, a community of 'campo

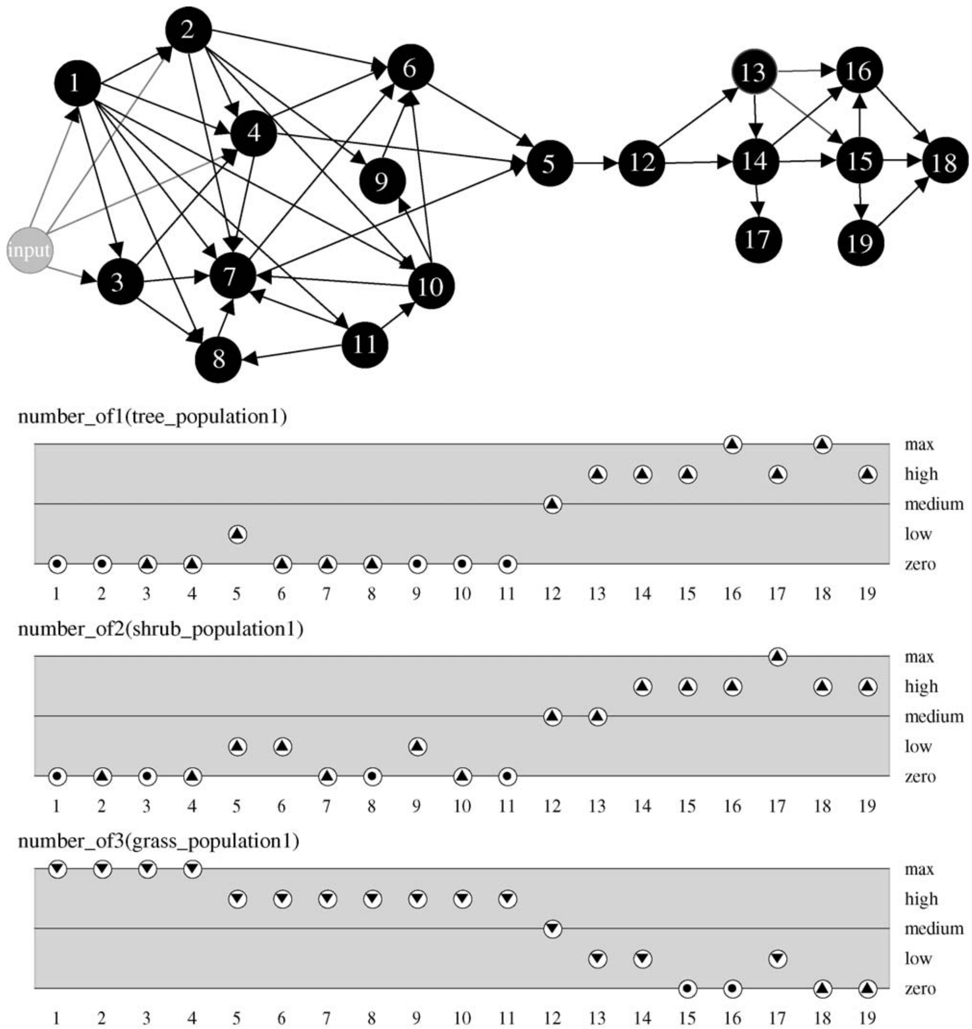


Fig. 11 – State-graph capturing the Cerrado succession hypothesis and values of three state variables.

limpo' (state 1) may change to 'campo sujo' (states 4 or 5) because *G* will decrease, and *S* and *T* will increase. Next, 'campo sujo' can also change and become denser like a 'simple campo cerrado' (state 12), and keep changing (states 14 and 15) until a 'climax cerradão' (state 18).

The assumption ('simple campo cerrado') introduced in the initial scenario plays an important role in simplifying the simulation. It is expressed in the MF 'simple campo cerrado,' a subtype of MF 'campo cerrado,' and defines a community in which ($T = S = G =$ medium). The effect of this assumption can be seen in Fig. 11. There is a bottleneck between a 'campo sujo with tree' (state 5) and a 'simple campo cerrado' (state 12), the transition from the 'campo' to the 'woody cerrado' types of communities. Without this simplifying assumption more states are generated, resulting in many more behaviour paths showing the same succession phenomenon, albeit with small variations in the intermediate states of behaviour.

The whole library has currently some 120 MFs. To give an idea of CSH model size, consider state 12, the 'simple campo cerrado' community. This state is described by means of 61 MFs (remember that the same MF can be applied simultaneously to different objects, e.g. MF 'Natality' in *G*, *S* and *T*). They introduce 20 entities, associated with 32 quantities and influenced by 19 process (natality, mortality, immigration, emigration, population growth, colonisation for *T*, *S* and *G*, and the agent model for human actions). These quantities are constrained by 127 different relations. There are 39 inequality relations involving the magnitudes and 13 involving the derivatives of the quantities. Also 16 correspondences (between specific values or QS of quantities) are included in the system description. The processes introduce 16 direct influences on quantities (remember that colonisation process is described as a kind of Immigration process and therefore both introduce the same influence). The effects of these processes are propagated via 43 indirect influences and qualitative proportionalities.

The current implementation, of all the models described in Section 4, has nearly 100 initial scenarios. This number can be easily augmented, by changing the initial values of quantities (magnitudes and derivatives), without introducing any further additional knowledge to the library.

5. Related work

Capturing qualitative knowledge in simulation models is an outstanding problem for ecological modelling. Guerrin and Dumas (2001a,b) describe models about the functioning of spawning areas of salmon (salmon reeds) and its impact on mortality rates at early stages. These models combine processes occurring at different time scales (fast and slow). Faster changes in the salmon development depend upon the daily water temperature. Slower changes depend on the accumulated water temperature and the dissolved oxygen concentration over time. The resulting model shows, for example, that when rain increases, the flow of water in the river also increases, increasing suspended solids and sediments and reducing the dissolved oxygen. These factors increase fish mortality, as expected from experts and the literature.

The approach taken by Guerrin and Dumas is implemented in QSIM (Kuipers, 1986) and therefore significantly differs from our approach. One distinction is that QSIM models do not explicitly represent causality, which hampers the use of these models for automated explanation generation, particularly in the context of educational applications. Another aspect is that QSIM does not use model fragments to compose models. By using a compositional-approach our qualitative domain theory about population dynamics can be used to implement variations of the same model and for scaling up to more complex problems (as we did by applying it to the CSH). This approach gives the modeller more flexibility and the possibility of explicitly representing and manipulating assumptions. On the other hand, models like those described by Guerrin and Dumas have a more detailed representation of time than our models. However, this is not a crucial aspect for the applications we foresee for our models.

Different models about population dynamics, like the S-shaped model (based on the logistic equation), have been implemented both in QSIM (Kuipers, 1994), and in GARP (Kamps and Péli, 1995). Although very popular among ecologists, these models are based on strong assumptions, like the density-dependency. We believe that simpler models about the exponential growth are still very rich in representing population behaviours of interest.

Modelling communities with qualitative reasoning techniques as described here is not found in the literature. May (1973) did a qualitative analysis of the results produced by differential equation models about interactions between populations to study the relationship between complexity and stability of biological communities. Most of the modelling approaches to these problems make assumptions about the magnitudes of the interactions between species in the community. May's research question was to investigate what can be said if only the topological structure of the trophic web is known, i.e. knowing only the signs (−, 0, +) of the interaction between the species and reasoning with changes over time (with the derivatives of the quantities). May showed that a less complex community met the conditions for stability, while the more complex one was not stable. Therefore, the 'commonsense wisdom' that more complexity means increased stability may not be true. He pointed out that such qualitative analysis can be a useful approach for modelling quite complex food webs and to capture the general tendencies of the system, bypassing long and complicated steps required by numerical models.

Noble and Slatyer (1980) proposed an approach for building qualitative models about the dynamics of communities subject to recurrent disturbance (such as fire). This approach is based on a small number of attributes of the plant's life history (vital attributes) that can be used to characterise the potentially dominant species in a particular community, under different types and frequencies of disturbance. Simulations typically produce a replacement sequence that depicts the major shifts in composition and dominance of species that occur following a disturbance. At each state, the community is defined by the presence of a sufficient number of individuals in particular stage of their life cycle (for example propagules, juvenile or mature). While further developing this modelling approach, Moore and Noble (1990, 1993) describe a simulation model that is also based on the vital attributes, combined with knowledge about the abundance of the populations and their survival according to the availability of environmental resources. Population sizes and certain vital attributes (e.g. germination rate) are stored internally in the model as real numbers. However, their values are mapped into discrete scales and presented in qualitative terms such as {low, medium, high} in the final output. This makes the qualitative simulation model useful for predictions and facilitates its use in supporting decision-making.

Recently McIntosh (2003) describes a modelling language for dealing with partial and imprecise ecological knowledge. Borrowing some concepts from QR, such as the representation of quantities (including the distinction between amount and derivative, both having two value components, magnitude and sign, and a set of possible qualitative values), the author implements his ideas using a rule-based approach and presents an example about vegetation dynamics. Succession is based on state-transition modelling and describes changes in communities due to fire and grazing.

Our work differs from these three approaches to community modelling in a number of points. Beside those already discussed (vocabulary, causality, compositional modelling), our models do not use any numerical values and the qualitative simulations are based only on symbolic representations of values and calculations. Moreover, our models capture knowledge that explains *why* changes happen, given that they use 'deep knowledge' (basic population dynamics) to represent the underlying mechanisms that cause changes in the system and to support reasoning about community phenomena.

6. Discussion

There is a great need for computer-based approaches for reasoning and making predictions about the behaviour of ecological systems with incomplete knowledge, expressed in qualitative terms. This paper shows that qualitative reasoning (QR) methods can be useful for using such knowledge to build models in which predictions and explanations are grounded on explicit representation of objects, typical situations, representative quantities, relations between quantities and processes in the domain of population dynamics.

The basis for the models presented here is a library of MFs representing knowledge about basic processes such as natality, mortality, immigration/colonisation, emigration, and population growth. This core knowledge is considered to represent the *first principles*, used to support simulation models about single populations, interactions between two populations and succession in the Cerrado vegetation. These models have been included in a model inspection tool VISIGARP that has proven to be useful in educational settings (Bouwer and Bredeweg, 2001).

The selection of the quantities and quantity spaces is one of the most important activities for a qualitative modeller. However, ecological systems do in general not have landmark values with strong meaning such as freezing and boiling points (as e.g. known from physics), although the carrying capacity (*K*) can be seen as an example. Therefore, ecology puts new requirements for the QR modeller to decide about. In the models presented here we adopted some quantity spaces based on concepts such as normality (e.g. QS = {zero, normal, max}), interval-based representations of size (e.g. QS = {zlmhm}) and a maximum value for the variable. These are very intuitive representations of quantities that fit well in the ecological systems we are dealing with.

Single population models provide different perspectives for the description of processes involved in population growth, either using only the basic processes (natality, mortality, emigration, and immigration/colonisation) or defining an aggregated of processes (the population growth process). These options expand the vocabulary encoded in the library and provide means for the modeller to test different hypotheses about population dynamics. In order to illustrate this point, consider the work done by Cromsigt et al. (2002). After comparing the results of five different mathematical population models in the context of management of black rhinos, these authors clearly demonstrate that relatively small differences in model structures can result in large differences in model predictions. They conclude that “it does not seem wise to base management decisions solely on the outcome of one model” and “strongly encourage managers to compare the outcome from different model structures, based on different biological assumptions”. Changes in the structure of a qualitative model like those presented here can be used to explore different assumptions and to compare predictions.

Assumptions are explicitly represented and the examples presented here illustrate many important modelling decisions. For example, alternatives for the notion of ‘closed’ population can be implemented by considering only {*B* and *D*} or considering the four basic processes (as in an open population), but assuming that *Im* and *E* correspond, $Im = E$ or $Im = E = 0$. It is important to notice that a modeller can change the assumptions, resulting in different behaviour of the system.

Community behaviour can be seen as the result of a complex web of relationships and interactions between pairs of populations. Understanding such interactions constitutes an important part of ecological theory and practice. We have presented a set of qualitative models and simulations that capture knowledge about the interactions between two populations. With these models it is possible to derive complex community behaviour from *first principles* in ecology.

The CSH simulation models represent commonsense knowledge about the effects of fire on the Cerrado vegetation. Managers and students have difficulties understanding the causal chain of reasoning involving environmental factors and changes in populations and communities. The added value of simulation models such as those presented in the CSH is to make the relationships between them explicit. The hypothesis implemented in the current version of the library may be changed or expanded, exploring alternative causal chains and adding more details to the models.

The combination of direct influences and qualitative proportionalities provides a language that is sufficiently powerful to express, albeit with less precision, any system of ordinary differential equation free of simultaneities whose independent parameter is time (Forbus and de Kleer, 1993). These authors argue that this restriction is analogous to the state space widely used in engineering modelling, where systems are decomposed into a set of state variables and dependent variables. In our qualitative models, directly influenced parameters play the role of state variables and indirectly influenced parameters play the role of dependent variables. Direct and indirect influences can also be compared to the ‘flows’ and ‘influences’ concepts adopted in System Dynamics. However, the qualitative representations of such mathematical functions include the notion of causality and the restriction of monotonic functions for proportionalities.

Feedback systems, which are general mechanisms for controlling biological and ecological systems, always imply a time delay, that is, it takes some time for an effect to happen. These relationships always contain a derivative relationship, which is modelled by a direct influence, instead of a qualitative proportionality (Forbus, 1984). Therefore, it is possible to represent a feedback loop following the pattern {*P?*(*A*, *B*); *I?*(*B*, *A*)}. For examples, see the interactions between the quantities Nof and *B*, *D*, *E* and *Im* presented in Section 4.1.

There is some added value and a price to pay for those who adopt a QR approach, as done in this article. QR is not a substitute for mathematical models. QR techniques are particularly useful in situations where the available knowledge is incomplete or expressed in qualitative terms. However, OR can also be used along with quantitative data. We believe the ecological modelling community will profit from exploring QR methods for designing systems and experiments, testing hypothesis, solving problems, explaining system behaviour, and doing diagnosis.

Acknowledgements

We would like to thank Bob Muetzelfeldt and the anonymous reviewers for their valuable comments on earlier versions of this article.

REFERENCES

Bouwer, A., Bredeweg, B., 2001. VisiGarp: graphical representation of qualitative simulation models. In: Moore, J.D., Luckhardt Redfield, G., Johnson, J.L. (Eds.), Artificial Intelligence in Education: AI-ED in the Wired and Wireless Future. IOS-Press, Amsterdam, pp. 294–305.

Bredeweg, B., 1992. Expertise in Qualitative Prediction of Behaviour. PhD thesis, University of Amsterdam, Amsterdam.
Bredeweg, B., Winkels, R., 1998. Qualitative models in interactive learning environments: an introduction. Interact. Learn. Environ. 5, 1–18.
Bredeweg, B., Forbus, K., 2003. Qualitative modeling in education. AI Mag. 24, 35–46.
Bredeweg, B., Struss, P. (Eds.), 2003. Current topics in qualitative reasoning. AI Mag. (special issue) 24, 13–130.
Coutinho, L.M., 1990. Fire in the ecology of Brazilian Cerrado. In: Goldammer, J.G. (Ed.), Fire in the Tropical Biota, Ecological Studies, vol. 84. Springer-Verlag, Berlin, pp. 82–105.
Cromsigt, J.P.G.M., Hearne, J., Heitkönig, I.M.A., Prins, H.H.T., 2002. Using models in the management of Black rhino populations. Ecol. Model. 149, 203–211.
Eiten, G., 1972. The Cerrado Vegetation of Brazil. Bot. Rev. 38, 201–341.
Falkenhainer, B., Forbus, K., 1991. Compositional modeling: finding the right model for the job. Artif. Intell. 51, 95–143.
Forbus, K., 1984. Qualitative process theory. Artif. Intell. 24, 85–168.
Forbus, K., de Kleer, J., 1993. Building Problem Solvers. MIT Press, Cambridge, Mass.
Gillman, M., Hails, R., 1997. An Introduction to Ecological Modelling: Putting Practice into Theory. Blackwell Science, Oxford.
Goodland, R., Ferri, M.G., 1979. Ecologia do cerrado. Editora Itatiaia-Editora USP, São Paulo.
Guerrin, F., Dumas, J., 2001a. Knowledge representation and qualitative simulation of salmon reed functioning. Part I. Qualitative modelling and simulation. BioSystems 59, 75–84.
Guerrin, F., Dumas, J., 2001b. Knowledge representation and qualitative simulation of salmon reed functioning. Part II. Qualitative model of redds. BioSystems 59, 85–108.
Haefner, J.W., 1996. Modeling Biological Systems: Principles and Applications. Chapman & Hall, New York, NY.
Iwasaki, Y., Simon, H.A., 1986. Causality in device behavior. Artif. Intell. 29 (1), 3–32.
Jørgensen, S.E., Bendoricchio, G., 2001. Fundamentals of Ecological Modelling, 3rd ed. Elsevier, Oxford.
Kamps, J., Péli, G., 1995. Qualitative reasoning beyond the physics domain: the density dependence theory of organisational, ecology, in: Bredeweg, B. (Ed.), Proceedings of the 9th International Qualitative Reasoning Workshop, 16–19 May, University of Amsterdam, Amsterdam.
de Kleer, J., Brown, J.S., 1984. A qualitative physics based on confluences. Artif. Intell. 24, 7–83.
Kuipers, B., 1986. Qualitative simulation. Artif. Intell. 29, 289–338.
Kuipers, B., 1994. Qualitative Reasoning: Modeling and Simulation with Incomplete Knowledge. MIT Press, Cambridge, Mass.
May, R.M., 1973. Qualitative stability in model ecosystems. Ecology 54, 638–641.
McIntosh, B.S., 2003. Qualitative modelling with imprecise ecological knowledge: a framework for simulation. Environ. Model. Softw. 18, 295–307.
Miranda, H.S., Saito, C.H., Dias, B.F.S. (Eds.), 1996. Anais do Simpósio "Impactos de Queimadas em Áreas de Cerrado e Restinga", in: Proceedings of the 3rd Congress of Ecology of Brazil, 6–11 October, Universidade de Brasília, Brasília.
Moore, A.D., Noble, I.R., 1990. An Individualistic model of vegetation stand dynamics. J. Environ. Manage. 31, 61–81.
Moore, A.D., Noble, I.R., 1993. Automatic model simplification: the generation of replacement sequences and their use in vegetation modelling. Ecol. Model. 70, 137–157.
Moreira, A., 1992. Fire Protection and Vegetation Dynamics in the Brazilian Cerrado. PhD thesis, Harvard University, Cambridge, Mass.
Morosini, I.B.A., Klink, C.A., 1997. Interferência do capim-grodura (Melinis minutiflora Beauv) no desenvolvimento de plântulas de embaúba (Cecropia pachystachya Trécul). In: Leite, L.L., Saito, C.H. (Eds.), Contribuição ao conhecimento ecológico do cerrado. Universidade de Brasília, Departamento de Ecologia, Brasília, pp. 82–86.
Noble, I.R., Slatyer, R.O., 1980. The use of vital attributes to predict successional changes in plant communities subject to recurrent disturbances. Vegetation 43, 5–21.
Odum, E.P., 1983. Basic Ecology. CBS, New York.
Ribeiro, J.F., Walter, B.M.T., 1998. Fitofisionomias do Bioma Cerrado, in: Sano, S.M., Almeida, S.P. (Eds.), Cerrado: Ambiente e Flora. Planaltina, DF, EMBRAPA-CPAC, pp. 89–166.
Rykiel, E.J., 1989. Artificial Intelligence and expert systems in ecology and natural resource management. Ecol. Model. 46, 3–8.
Salles, P., 1997. Qualitative models in ecology and their use in learning environments. PhD thesis, University of Edinburgh, Edinburgh.
Salles, P., Bredeweg, B., 1997. Building qualitative models in ecology, in: Ironi, L. (Ed.), Proceedings of the 11th International Workshop on Qualitative Reasoning, 3–6 June. Instituto di Analisi Numerica C.N.R., Pavia (Pubblicazioni no. 1036).
Salles, P., Bredeweg, B., 2001. Constructing progressive learning routes through qualitative simulation models in ecology, in: Biswas, G. (Ed.), Proceedings of the 15th International Workshop on Qualitative Reasoning, 17–18 May, Saint Mary's University, San Antonio.
Salles, P., Bredeweg, B., Araujo, S., Neto, W., 2003. Qualitative models of interactions between two populations. AI Commun. 16, 291–308.
Salles, P., Bredeweg, B., Winkels, R., 1997. Deriving explanations from qualitative models. In: Boulay, B., Mizoguchi, R. (Eds.), Artificial Intelligence in Education: Knowledge and Media in Learning Systems. IOS-Press, Amsterdam, pp. 474–481.
Weld, D., de Kleer, J. (Eds.), 1990. Readings in Qualitative Reasoning about Physical Systems. Morgan Kaufmann, San Mateo, CA.