**Advances in
Geosciences**

Agents, Bayes, and Climatic Risks – a modular modelling approach

A. Haas and C. Jaeger

Potsdam Institute for Climate Impact Research, P.O. Box 60 12 03, 14412 Potsdam, Germany

Abstract. When insurance firms, energy companies, governments, NGOs, and other agents strive to manage climatic risks, it is by no way clear what the aggregate outcome should and will be. As a framework for investigating this subject, we present the LAGOM model family. It is based on modules depicting learning social agents. For managing climate risks, our agents use second order probabilities and update them by means of a Bayesian mechanism while differing in priors and risk aversion. The interactions between these modules and the aggregate outcomes of their actions are implemented using further modules. The software system is implemented as a series of parallel processes using the $CIAM^n$ approach. It is possible to couple modules irrespective of the language they are written in, the operating system under which they are run, and the physical location of the machine.

1 Introduction

The experience of the Kyoto protocol shows how difficult it is to establish a plausible climate policy in the short-term (Michel, 2004). At the same time even if the Kyoto protocol had been implemented quickly and successfully, it would be far from solving the long-term challenge of climate change (Hasselmann et al., 2003). For the future evolution of climate policy, it will be vital to develop complementary approaches to trigger a sustained process of social learning in the face of this global challenge.

So far, scientific thinking about climate policy has been mostly framed in terms of a single goal function that is assumed to be somehow appropriate for humankind as a whole. We will call such a function a Leviathan function, following Hobbes' classical view of a super-agent taking care of co-ordination problems shared by society as a whole. In research on climate policy, two kinds of Leviathan functions

Correspondence to: A. Haas
(armin.haas@pik-potsdam.de)

are currently used. On the one hand, natural scientists often assume that some quantitative measure of a stable climate can be given as a guideline for policy-making. In this spirit, increases in global mean temperature of 1/10 degree Celsius per decade have been proposed as a limit for the stress ecosystems can cope with (WMO, 1988). A total increase of one to two degrees Celsius has been proposed as an additional limit (Rijsberman and Swart, 1990; Hare et al., 2004). The resulting goal function is pretty simple, yielding only one of two values – 1 and 0, or "acceptable" and "inacceptable" – for possible trajectories of the climate system. On the other hand, economists often assume that a slightly more complex goal function is appropriate. In this case, a discounted time-integral of global Gross Domestic Product is proposed as Leviathan function. Climate change then is to be avoided to the extent to which it lowers GDP. Non-market effects are taken into account by monetary correction terms, sometimes including highly controversial estimates of different economic values of human lives in different countries.

In the future, it will be increasingly important to model the climate problem in a multi-agent setting where different goal functions are associated with different agents. As Tesfatsion (2006) argues, agent based computational economics has reached a degree of maturity where this seems a feasible goal. However, two kinds of challenges have to be met if useful insights are to be obtained. The first kind arises from the need to identify and represent social agents in a suitable way (Sect. 2). The second is the need to find a reasonable representation of learning processes in the face of far-reaching uncertainties (Sect. 3). On this basis, it is possible to gain substantial new insights on pitfalls and opportunities in the management of climate risks (Sect. 4).

2 Agents

The definition of agents for the present purpose requires a balance between standard economic concepts of utility, probability, and constraints on one hand, and more Artificial

Intelligence oriented concepts of rule following on the other hand (Sect. 2.1). The design of suitable modelling tools in turn requires a series of choices about how to implement software agents in a modular and integrated way (Sect. 2.2).

2.1 Human Agents

There are two major approaches for modelling human agents, be they individuals or organisations: maximising expected utility and rule following.

In economics, the standard concept states that agents do have preferences about the possible consequences of alternative choices. In practice (although not necessarily in theory) it is assumed that these preferences can be represented by an ordinal utility function. When confronted with an uncertain outcome from choices, agents are depicted as combining their preferences with their expectations for the occurrence of different outcomes. This is represented by a cardinal utility function combined with a probability distribution. Human agents are supposed to be able to identify and choose a single action which maximises their expected utility under constraints given by the situation they find themselves in.

In this approach, the agent's attitude towards risk can be handled in two ways. First, the shape of the utility function can be used to represent risk averse, neutral, or risk seeking behavior. Second, when maximising utility, additional constraints can be added to depict an agent's behavior towards risk. We use, for example, a non-ruin constraint stating a maximally accepted probability for going into ruin.

In information sciences, agents are often conceived as rule following entities. When confronted with a decision situation, they choose an action using a set of rules. These rules typically are if-then statements which trigger an action once a condition is met. In a decision situation, the core challenge for this approach is to come up with a single action as typically several conditions are met and the resulting set of stated actions has more than just one element. A standard mechanism to tackle this problem is to attach weights to the rules and apply the one with the highest weight. The weights themselves are adjusted in a learning process like the bucket brigade mechanism suggested by Holland (1992).

These two ways of modelling the decision processes of agents are most often perceived as alternatives. We interpret them slightly different. Each constraint of a maximisation problem can be interpreted as a rule the agent has to obey. In the case of the non-ruin constraint applied in our work, the rule is: whatever you do, the probability of going into ruin shall not exceed a specified threshold. The expected utility then can be used to come up with a single action amongst those satisfying the rules relevant for the situation at hand.

It is important to realise that even if each agent comes up with a single action in a given situation, this by no means implies the existence of a single optimal equilibrium for the whole range of agents. This hints at the potential richness of such a perspective as it opens up the way to hybrid models achieving greater descriptive and normative content by linking standard approaches of economics and of information sciences.

2.2 Software Agents

Human agents can be modelled by suitable software agents. We have done so with the model family LAGOM (LAGOM is a Swedish word denoting a sense of balance and harmony, perhaps akin to the chinese "Tao"). One way of designing the required software agents is to look at them as functions linking specific data types. It is useful to think about the building blocks of the program in terms of a hierarchy of such data types. At the most general level, we can distinguish between perception types and action types. A human agent, then, is implemented as a function whose domain consists of one or several perception types while the range consists of one or several action types.

In order to represent the uncertainties the agent is faced with, a third data type is needed, namely expectations. Human agents form expectations by comparing a given state of affairs with their previous expectations. Expectations are a state variable of the agent in question. This is reminiscent of Turing machines, as these map a perception into an action on the basis of an inner state which is redefined at each step of their operation. Accordingly, we use functions mapping perceptions and expectations to actions and expectations:

Human Agent:
   [[Perceptions], [Expectations]]
       → [[Actions], [Expectations]]

The perceptions of standard economic agents are characterised by two data types: the prices found on the market and the resources owned by the agent. Their actions relate to two more types: supply quantities and demand quantities. Economic agents can then be further specified into firms and households of various kinds. In a typical version of LAGOM, the economic agents are producers, households, and insurance companies. LAGOM relies on a general purpose optimisation algorithm (produced by Dan Ontanu, Bukarest, and Cezar Ionescu, Potsdam) that can handle an extremely large class of optimisation problems. In the future, additional modules may implement "fast and frugal" searching algorithms needed to model choices of non-optimising agents (Gigerenzer and Selten, 2001).

Human actions have consequences, and in social interactions these consequences depend on the actions of several agents. This pattern can be modelled by a second kind of software agent representing various kinds of social and biophysical contexts. A context can be implemented as a function mapping states of affairs and actions into new states of affairs:

Context:
   [[State of Affairs], [Actions]]
       → [State of Affairs]

The standard economic context is a market mapping supply and demand under given prices and a given resource allocation into new prices, traded quantities and the resulting resource allocation.

Finally, a third kind of software agent is needed to get a fully specified model. These are the program components enabling the software agents representing human agents and contexts to run in an appropriate way. They are bundled in a driver performing the following steps. First, the driver launches the other modules as parallel processes. Second, it gives each one of them the addresses it is supposed to talk or listen to. All modules have variables for input and output ports. The driver assigns values to these (modifying them at runtime if needed), thereby making it easy to exchange modules of the same kind in a plug and play mode. Third, the driver listens to selected modules in order to relaunch a next iteration until the goal of the simulation has been reached. And finally, it shuts down the whole process.

Figure 1 shows a typical LAGOM implementation (producers have been left out to avoid overcrowding the diagram, the expectation manager will be discussed in the next section). There are two layers of communication between the different modules. On one layer, the driver performs his business as described above. On the other layer, the different modules interact in a structured manner.

In the present example, there is a climate module (dubbed "stormy weather" in the figure) that damages households and is observed by the expectation manager and is influenced by actions of households. The expectation manager helps insurers and households to up-date their respective – usually different – expectations on the basis of new observations. The Walras finder, representing the "invisible hand" of the market, finds prices matching demand and supply for weather insurance by running through a fast (hence the short dashes) sequence of interactions with households and insurers. LAGOM operates at multiple time scales. Market interchanges can be and usually are much faster than climate change. Industrial production introduces a third time scale, usually lying between those two. By running a set of such agents recursively – i.e., feeding their output of step n as input into step n+1 – one gets a dynamic system. Computer experiments with LAGOM have shown that this format can be used to represent economic agents according to standard economic theory. An economic equilibrium then is a fixed point of the recursive mapping defined above. In systems with several fixed points, non-equilibrium dynamics can obtain when the initial conditions are not identical to a fixed point. They are particularly interesting when combined with some stochastic process (Haas, 2001).

The design pattern of LAGOM is based on the $CIAM^n$ software platform ($CIAM^n$ is an acronym for "Community Integrated Assessment Modules to the power of n" and refers to a software platform that we developed together with other researchers at PIK and in the broader context of the European Climate Forum – cf. the link "$CIAM^n$" at www.european-climate-forum.net). The platform combines an understanding of software management as a social process
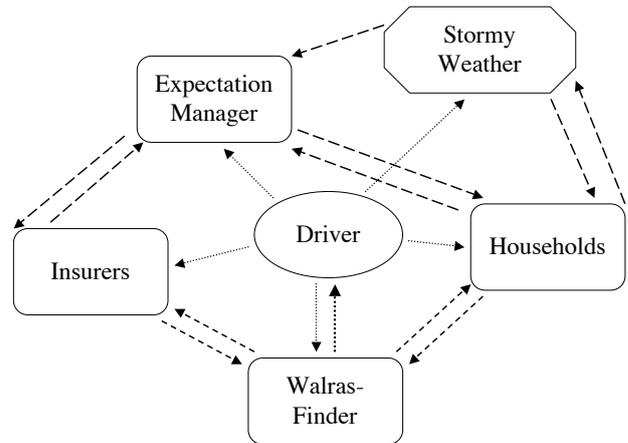

Fig. 1. Scheme of typical LAGOM implementation.

(Jaeger et al., 2002) with a logical analysis of the plurality of domains of discourse required for integrated assessment modelling (Jaeger, 2003). It provides algorithms for coupling modules programmed in a dynamical simulation mode with modules programmed in intertemporal optimisation mode – a key problem when combining natural science and economic models (Leimbach and Jaeger, 2004).

Using this platform requires the user to do three things: assign each piece of code to exactly one module, provide a driver module suitable for the problem at hand, and program input and output to each module according to a shared protocol for typed data transfer – TDT.

The TDT protocol (see the $CIAM^n$ link indicated above as well as www.pik-potsdam.de/software/tdt) provides a simple way of handling input and output to each module by means of two functions, one for writing and one for reading. These functions take addresses as parameters and work across different programming languages, different operating systems, and different machines in different physical locations. Computer experiments with LAGOM and other models have shown that drivers using this protocol can effectively co-ordinate highly heterogeneous multi-agent systems.

3 Bayes

As mentioned above, human agents often act on the basis of expectations that they try to improve in the course of action. To model such processes of expectation formation, some kind of Bayesian reasoning is needed.

A reasonable management of climatic risks is impossible without some way of making the relevant expectations of various agents – including scholars engaged in climate research – explicit. With LAGOM, we tackle this problem in two steps. First, we represent heterogeneous expectations of a variety of agents, accepting that they may have perfectly reasonable ways to entertain different expectations. Second, given these expectations, we model how learning agents can update their expectations on the basis of further experience.

Whether this updating will lead to converging expectations, and if so, at what speed, then becomes a question amenable to scientific inquiry.

For this purpose, it is necessary to consider not only frequentist probabilities, but also guesses about them. Every child learns to deal with relative frequencies: it learns to distinguish fast rythms from slow ones, it learns that certain teams win more often than other ones, it learns how to roll a dice, etc. Relative frequencies can be observed, they can be forecast, and they matter for many actions. The mathematics of probability measures can be used to deal with relative frequencies where a more in-depth analysis is required. Probabilities then appear as limits of relative frequencies defined over sequences of events.

Children also learn to make guesses about things they do not know, and to take action on the basis of such guesses – one may guess that a certain branch of a tree will not crack and climb on it, etc. Often, guesses have nothing to do with relative frequencies, as when one guesses what is the capital of some foreign country. But sometimes, we need to take action in the face of a situation where unknown relative frequencies matter. In these cases, guesses about frequentist probabilities occur.

The mathematics of probability measures turns out to be helpful for analysing guesses about frequentist probabilities, too. This is due to a result known as de Finetti's theorem (see Bernardo and Smith, 1994 for an exposition). The theorem shows how an unknown probability distribution can be approximated if additional samples become available step by step. For practical purposes, an important point is that the approximation can start from very different initial guesses as long as these guesses do not exclude the limiting distribution.

To apply this framework, one must distinguish between first order and second order probabilities and consider updating schemes for the latter ones. If two possibilities are given with unknown probabilities, then the relevant first order probabilities are given by the open interval (0,1). Complete ignorance then implies a distribution of second order probabilities corresponding to the probability density function $y=1$ over this interval.

An agent can improve upon this initial distribution under two conditions. First, she must be able to gather additional evidence. Second, she must have some opinion about the structure of the underlying process, which may be quite complex. If she is willing to consider key parameters of the process as characterized by some unknown frequentist probability, then the following updating scheme can be shown to be efficient.

$$p_{2,t+1}(p_i(s))$$

$$= p_{2,t}(p_i(s)) \cdot \frac{p_i(s)}{\int_{i \in I}(p_{2,t}(p_i(s)) \cdot p_i(s))dp_i(s)}.$$

p_2 stands for second order probabilities, t for time, p for first order probabilities, P for the set of all first order probabilities suitably indexed by an index set I, p_i for a specific member of the set of first order probabilities, and s for a specific situation obtained at time t out of a set of all possible events S.

In LAGOM, this updating mechanism has been implemented in a separate module, the expectation manager. It may be used as a skeleton for representing scientific communities along with mass-media amplifying their claims, as both decision-makers and the general public update their expectations concerning climate change by interacting with these institutions.

The updating process is similar to the formula known as Bayes' rule. But there is an important difference between Bayes' rule as applied to a static situation and this dynamic updating process. In a static situation, Bayes' rule follows immediately from the definition of conditional probabilities. The updating mechanism introduced above can be shown to be efficient for dynamic processes with a limiting frequentist probability, but this is far from being trivial.

Therefore, it is important to notice that the model design used for the updating of expectations in LAGOM can also be applied to non-Bayesian updating schemes, as when one agent imitates the behavior of another one.

4 Climatic Risks

The various IPCC publications (for an overview see Watson, 2001) are the authoritative source on climatic risks. They show beyond reasonable doubt that humankind is altering the global climate system in ways that can lead to serious damages to humans and to things they value – by sea-level rise, droughts, floods, storms, and the like. These publications derive their authority to a considerable extent from the fact that their production is controlled by a carefully crafted consensus-building mechanism. This, however, makes it impossible to reach an in-depth assessment of these risks. There are two reasons for this state of affairs: the difficulty to reach an agreement on the value to be attached to non-marketable goods like human lives or the beauty of coral reefs, and the difficulty to reach an agreement on the likelihood to be attached to events without statistical track record. A good example for these difficulties as well as for their importance is the Pentagon study on abrupt climate change (Schwartz and Randall, 2003). The study considers the possibility that anthropogenic climatic change will lead to a shut-down of the thermohaline ocean circulation in the North Atlantic. It does so by taking as an analog a similar shift in ocean currents that occurred about 8000 years ago, looking at its effects then, and asking what might be the worst conceivable implications of these events in today's world. This leads to scenarios of social turmoil and international tensions, including violent conflict and dramatic losses of welfares in Europe and North America. Similar scenarios, geared to events like shifting monsoon patterns, melting permafrost, intensified El Niño events, and the like, can be produced for other parts of the world. While the available evidence is sufficient to reach agreement on the fact that these scenarios describe possible courses of events, there is no way to attach

any frequentist probability measures to them; nor is there a meaningful Leviathan function providing the one and only yardstick for the evaluation of such risks.

With the model of learning agents presented in this paper, another approach is feasible: One can study the options of agents acting on the basis of heterogeneous expectations and evaluations in the face of climatic risks. In particular, one can study to what extent their expectations may converge on the basis of additional evidence and what sort of compromise they may find in order to reduce the risks about which they care most. While it is extremely unlikely that problems of this kind have a single optimal solution, it is almost certain that stepwise improvements can be achieved through a combination of negotiations and learning by doing.

Along these lines, the model can be used to study the management of climatic risks without the need to assume an unwarranted Leviathan function. This is not to say that no agent can play the role of Hobbes' Leviathan. Quite the opposite, the model can be used to study how a specific agent may acquire that role, which is then not assumed, but explained. The model can also be used to study multilateral constellations where climate risks are managed without any agent assuming the Leviathan role. Simulations show that such constellations are hardly ever characterised by a single fixed point of the relevant dynamics. Therefore, a multilateral solution of the climate problem will need to rely on some selection mechanism, most likely of the kind described by Schelling as a focal point (Schelling, 1960).

This state of affairs implies that processes of supply and demand – even if amended by instruments like carbon taxes and emissions permits – will be insufficient to bring about a solution. For this purpose, economic mechanisms and policy instruments must be complemented by suitable processes of expectation formation. While Bayesian learning is by no means the only mechanism of this kind, it provides a useful starting point for multi-agent modelling of climate risks.

5 Conclusion

Modular multi-agent models representing learning agents faced with climatic risks are both desirable and feasible. They are desirable because climate risks are not amenable to a purely frequentist analysis and because the use of a Leviathan function dodges the question of how to reach agreement over the valuation of non-marketable goods. They are feasible because software tools are available to couple highly heterogeneous software agents, because mathematical concepts like second order probabilities can be used to represent learning agents, and because the scientific evidence on climatic risks is sufficient to identify salient risks. In addition, modular modelling has the advantage of facilitating co-operation between researchers with different fields of expertise, thereby improving the reliability of the resulting computations.

References

Bernardo, J. M. and Smith, A. F. M.: Bayesian Theory, Wiley, New York, 1994.

Gigerenzer, G. and Selten, R.: Bounded Rationality: The Adaptive Toolbox, MIT Press, Cambridge, Massachusetts, 2001.

Haas, A.: Kartell und Marktprozeß, Duncker & Humblot, Berlin, 2001.

Hare, B., Schaeffer, M., and Meinshausen, M.: What is dangerous climate change? Initial results of a symposium on key vulnerable regions, climate change, and article 2 of the UNFCCC, Buenos Aires, 2004.

Hasselmann, K., Latif, M., Hooss, G., Azar, C., Edenhofer, O., Jaeger, C. C., Johannessen, O. M., Kemfert, C., Welp, M., and Wokaun, A.: The challenge of long-term climate change, Science, 302, 1923–1925, 12 December 2003.

Holland, J. H.: Adaptation in Natural and Artificial Systems, MIT Press, Cambridge, Massachusetts, 2nd edition, 1992.

Jaeger, C. C.: A note on domains of discourse. Logical know-how for integrated environmental modelling, www.pik-potsdam.de/~cjaeger/working_papers/pr86.pdf, 2003.

Jaeger, C. C., Leimbach, M., Carraro, C., Hasselmann, K., Hourcade, J.-C., Keeler, A., and Klein, R.: Integrated assessment modeling: Modules for cooperation, FEEM Nota di lavoro 53-2002, www.european-climate-forum.net/pdf/modular_iam.pdf, 2002.

Leimbach, M. and Jaeger, C. C.: A modular approach to Integrated Assessment modeling, Enviromental Modeling and Assessment, 9, 207–220, 2004.

Michel, D. (Ed.): Climate Policy for the 21st Century, Meeting the long-term Challenge of Climate Warming, Johns Hopkins University, Washington, D.C., 2004.

Rijsberman, F. and Swart, R.: Targets and indicators of climate change, Report of Working Group II of the Advisory Group on Greenhouse Gases, WMO/UNEP, Stockholm Environment Institute, Stockholm, 1990.

Schelling, T.: The Strategy of Conflict, Harvard University Press, Cambridge, Massachusetts, 1960.

Schwartz, P. and Randall, D.: An abrupt climate change scenario and its implications for united states national security, www.icomm.ca/survival/pentagon_climate_change.pdf, 2003.

Tesfatsion, L: Agent-Based Computational Economics: A Constructive Approach to Economic Theory, in: Handbook of Computational Economics, Volume 2: Agent-Based Computational Economics, edited by: Judd, K. L. and Tesfatsion, L., North-Holland, accepted, 2006.

Watson, R. T. (Ed.): Climate change 2001: synthesis report, Cambridge University Press, Cambridge, 2001.

WMO: Developing policies for responding to climatic change, WMO/TD no. 225, World Meteorological Organization, Geneva, 1988.