2 CONTEXT AND MOTIVATION: FORMALIZING KNOWLEDGE ON INTERNATIONAL ENVIRONMENTAL REGIMES

The overall objective of this thesis is to make the knowledge included in environmental computational models explicit, in order to enable correct interpretation and reuse. In this chapter, we illustrate the relevance of this objective in a modeling study in the field of environmental governance.

In order to address global environmental problems, policy makers have developed systems of rights and obligations and related decision-making procedures, also known as *international environmental regimes*. In this chapter we present an innovative approach to formalize knowledge on the effectiveness of these regimes. We extract general findings from literature on international regimes and translate these into a set of knowledge rules. We formalize these into a conceptual framework for the systematic analysis of conditions that influence regime effectiveness, which we implement in a computer model using fuzzy logic methodology. As we make the knowledge on environmental regimes accessible and explicit, scientists from both the political science and the environmental modeling domain are able to understand it, discuss it and contribute to it.

This chapter is based on a paper coauthored by Peter Janssen, Marcl Kok, Sofia Frantzi, Eleni Dellas, Philipp Pattberg, Arthur Petersen, and Frank Bierman "Formalizing knowledge on international environmental regimes: A first step towards integrating political science in integrated assessments of global environmental change"(Vos et al., 2013), which has been published in the Environmental Modelling & Software journal.

2.1 INTRODUCTION

Global environmental problems have features that distinguish them from traditional scientific problems. They concern 'global public goods', are global in their scale and long-term in their impact and they are characterized by high uncertainty, complexity and multiple interests, requiring transdisciplinary approaches to deal with them (Funtowicz and Ravetz, 1993; Jakeman, Letcher, and Norton, 2006; Schmolke et al., 2010b; Sluijs, 2002). In order to address global environmental problems most effectively and efficiently, policy makers have developed series of systems of rights and obligations and related decision-making procedures in international environmental policy, also known as international environmental regimes (Carter, 2007). These regimes are a type of institution, where an institution is understood as a "cluster of rights, rules, and decision-making procedures that gives rise to a social practice, assigns roles to participants in the practice, and guides interactions among occupants of these roles" (Young, 2008). International environmental regimes are a distinct type of institution dealing with issue-specific environmental concerns at the international level (Hasenclever, Mayer, and Rittberger, 1997). International environmental regimes are considered key factors in dealing with global environmental problems (Biermann, 2007; Kates et al., 2001), but their development and implementation may be costly and difficult. It is therefore important to understand if and how regimes are effective in tackling these problems. This requires knowledge on the potential impact of environmental regimes as well as on their political feasibility.

The creation and performance of regimes to solve international environmental problems is studied by the field of international relations and more specifically by environmental regime theory. Scholars in this field have applied different qualitative and quantitative approaches to defining and measuring regime effectiveness (Haas, Keohane, and Levy, 1993; Miles et al., 2002; Young, 1999). When measuring the performance of international regimes, scholars by and large focus on the behavioral change of key actors (i.e. states)and not on environmental improvements (Easton, 1965; Underdal, 2001). Even though some scholars recognize the need to look at environmental improvement, measurement of the specific impacts of a regime is difficult because disentangling these impacts from influences that are independent from the regime is complicated and often practically almost infeasible. An additional complication is that a regime can only influence behavior, achieve its goals, and address an environmental problem once it has been formed and implemented (Underdal, 2001), and thus assessing the effectiveness of proposed regimes is inevitably speculative. When trying to draw reliable and policy-relevant lessons on a wide variety of existing and proposed regimes the current approaches to defining and measuring regime effectiveness suffer from problems of comparability and generalizability (Biermann, 2007). A further challenge is that these types of analyses are often not combined with environmental outcomes. One way to improve this is through a better cooperation between the fields of environmental regime theory and integrated assessment of global environmental change.

Integrated assessment is a methodology to analyze global environmental problems by combining knowledge from the social, environmental and economic domains relying strongly on quantification and computer simulation, but also by incorporating participatory methods to include stakeholders in integrated assessments (Siebenhüner, 2002). Integrated assessment models have become essential tools in supporting environmental decision making by exploring the consequences of alternative policies or scenarios (Jakeman, Letcher, and Norton, 2006; Schmolke et al., 2010b; Sluijs, 2002). Scientists in the field of integrated assessment acknowledge that it is important to include knowledge on environmental regimes in their analyses (Reid et al., 2010; Rotmans and De Vries, 1997; Turner et al., 2003), but they have not yet been successful in doing so (Ostrom, 2009) due to inherent difficulties to model human and social dimensions. Models of social institutions ultimately rest on assumptions about human behavior, which is 'substantially nontrivial' (Braumoeller and Sartori, 2004) and might be more complex than a model suggests or is able to capture. Another core problem is the difficulty in conceptualizing key social concepts like power and legitimacy that are essential in explaining regime effectiveness, but are difficult to operationalize in a form that can be used for quantitative modeling approaches. Therefore, knowledge on environmental regimes is often disregarded in integrated assessments of sustainable development (Biermann, 2007).

In this chapter the multi-disciplinary challenge of bringing together the worlds of integrated assessment and regime theory is taken up. We here offer an innovative approach to formalize knowledge on the effectiveness of environmental regimes. The aim of this chapter is to make this knowledge explicit so that scientists from both domains can understand it, discuss it and perform systematic analysis of the effectiveness of environmental regimes. We constructed a conceptual framework for the systematic analysis of conditions that influence regime effectiveness based on regime theory. We implemented the framework in a computer model using fuzzy logic methodology (Zadeh, 1965), a simple and straightforward way of linguistic reasoning. We expect added value of this framework and the computer model as a tool to perform assessments of international environmental regimes and of options to improve their effectiveness.

This chapter is structured as follows. Section 2.2.2 describes the construction of the conceptual framework, the selection of a suitable modeling method and its translation into a fuzzy logic model. Section 2.3 evaluates the model in an analysis of four existing environmental regimes and Section 2.4 discusses main findings on the usefulness and prospects of our conceptual framework and model. Throughout the chapter we will reflect on methodological challenges we encountered in our research.

2.2 METHODOLOGY

The following section explains our method of formalizing knowledge on environmental regimes (Figure 2.1). We formulated a definition of regime effectiveness, gathered knowledge on regime effectiveness in a literature review and formalized this knowledge in rules and a conceptual framework. In a review of possible modeling approaches we selected fuzzy logic as the most appropriate technique. The last part of this section describes the translation of the conceptual framework into a fuzzy model.

2.2.1 *Assessing regime effectiveness: building on international regime literature*

In the development of international environmental regimes, regime formation and regime implementation can be considered as two important distinct phases. While the former includes the negotiations among states, the latter includes the process of putting the regime's stipulations into practice. Although successful regime formation and implementation may not be sufficient to guarantee effective ways to deal with the environmental problems at hand, we consider them as necessary preconditions for the functioning of international environmental regimes. In this chapter we therefore use the terms 'likelihood of regime formation' and the 'likelihood of regime implementation' as proxies for their effectiveness. Within the research field of international regimes, the literature provides numerous hypotheses regarding the factors that influence

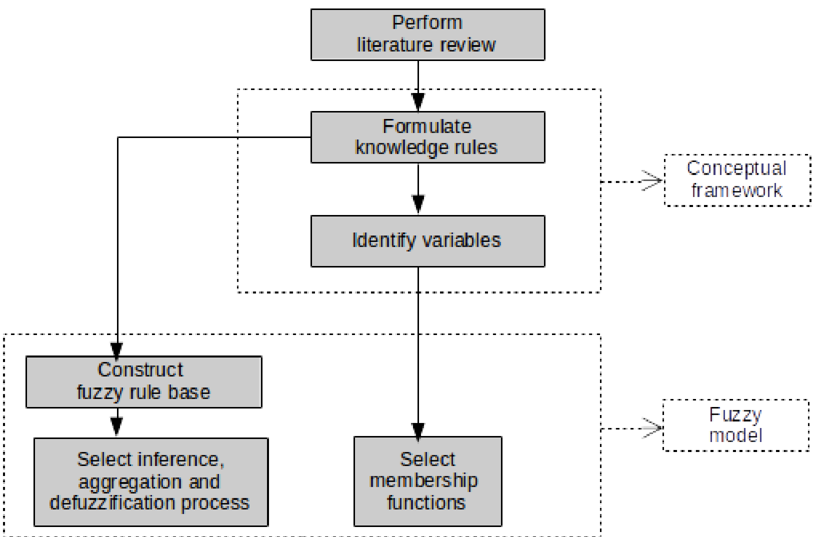

Figure 2.1: Description of the various steps in the process of developing a conceptual framework and a fuzzy model to analyze regime effectiveness.

regime formation and implementation, (for instance: (Breitmeier, Young, and Zurn, 2006; Miles et al., 2002; Mitchell, 2008; Underdal, 2001; Young, 2008). In a review of this literature (Dellas et al., 2011) we identified robust general findings and translated them by expert judgement into a set of 64 clear rules on the likelihood of regime formation and implementation (see Appendix A.1 for an overview), with a view towards formalizing this knowledge in a conceptual framework as a basis for further analysis of regimes and formal modeling.

2.2.2 *Constructing a conceptual framework*

From the rules we identified the determinants of likelihood of regime formation and implementation (Table 2.1 and 2.2). The next step was to relate the determinants, i.e., 'input variables', and the two phases in regime development together in a conceptual framework, using the likelihood of regime formation and of regime implementation as respective 'output variables' (Figure 2.2).

We have distinguished the input variables in 'context' and 'design' variables (Table 2.1 and 2.2). There are three categories of 'context' variables: 'problem structure', which refers to attributes of the environmental problem (Hovi, Sprinz, and Underdal, 2003; Miles et al., 2002; Young, 1999), '(state) actors', which refers to the (state)actors who take part in negotiations on the regime and its implementation (Hasenclever, Mayer, and Rittberger, 1997), and 'regime environment' which concerns the background against which regime formation or implementation takes place including the influence of other institutions and norms (Oberthür and Gehring, 2006). The 'design' variables refer to the choices that policy makers can make during regime formation or implementation, and they in fact can mitigate or enhance the impact that the 'context' variables have on regime effectiveness. As the 'context' variables vary between different environmental problems, the appropriate institutional structure or regime 'design' will also differ (Young, 2008). For example the context variable 'asymmetry of states' interests' describes the difference between states with respect to the responsibility for causing an environmental problem, the capacity to address it, and the vulnerability to its impacts (Underdal, 2001). High asymmetry may negatively affect the likelihood of regime formation or implementation (rules C3 and C4, see Appendix A.1). Designing policy measures such that 'differentiated responsibili-

Table 2.1: Input variables of regime formation listed per category

Context variables			Design variables
Problem structure	Actors	Regime environment	Negotiation process
Regulation costs[n]	Asymmetry of interest of powerful states[n]	Preceding agreement[b]	Negotiation costs[n]
Public concern[n]	Asymmetry of interest of important states in issue area[n]	Scientific advisory bodies[b]	Differentiated rules[b,m]
Systemic/cumulative problem[b]	Support of powerful states[n]		Side payments[b,m]
Uncertainty[b]	Support of important states in issue area[n]		Transaction costs[b,m]
Collaboration/coordination problem[b]	Number of economic sectors[n]		Framework treaty[b,m]
	Homogeneous states[b]		Informal agreement[b,m]
	Urgency[b]		Incentives[b,m]
	Cumulative cleavages[b]		Positive issue linkages[b,m]
	Powerful pushers[b]		
	Powerful laggards[b]		

[n] Numerical variable: expressed on a scale from 0(low) to 10(high); translated into linguistic categories using the mf from Fig middle panel
[b] Binary variable: expressed as 0(no) or 10(yes), translated into linguistic categories using the mf from Fig upper panel
[m] Mitigating variable: (design) variable; appears in combined rules; mitigates the negative impact of other input (context) variables

Table 2.2: Input variables of regime implementation listed per category

Context variables			Design variables
Problem structure	Actors	Regime environment	Regime design
Regulation costs[n]	Participation government[b]	institutional framework[b]	Knowledge mechanism[b]
Collaboration/coordination problem[b]	Participation powerful states[n]	Negative interplay[b]	Differentiated rules[b,m]
Systemic/cumulative problem[b]	Participation important states in issue area[n]	Positive interactions[b]	Compliance mechanism[b,m]
	Number of economic sectors[n]		Side payments[b,m]
	Outvoting of important states[b]		Compliance mechanism[b,m]
	Asymmetry of interest of powerful states[n]		Precise rules[n]
	Asymmetry of interest of important states in issue area[n]		Legally binding rules[b]
			Strong secretariat[b]
			Reporting mechanism[b]
			Consensus voting[b]
			Broad issue coverage[b]
			Public awareness mechanism[b]

[n] Numerical variable: expressed on a scale from 0(low) to 10(high); translated into linguistic categories using the mf from Fig middle panel
[b] Binary variable: expressed as 0(no) or 10(yes), translated into linguistic categories using the mf from Fig upper panel
[m] Mitigating variable: (design) variable; appears in combined rules; mitigates the negative impact of other input (context) variables

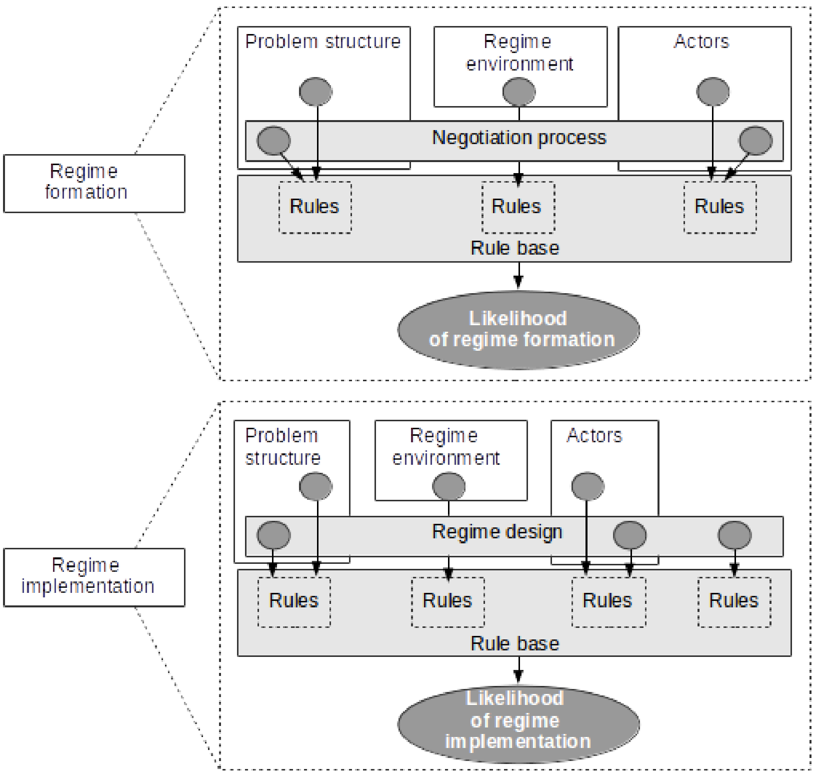

Figure 2.2: Conceptual framework for the analysis of the likelihood of regime formation and implementation

ties' are allowed for the different (state)actors, can help to mitigate the negative impacts of the asymmetry in their interests, thus enhancing likelihood of regime formation or implementation (Young, 2008). Some 'design' variables refer to choices that policy makers can make, especially in the implementation phase, which do not have a mitigating effect on context variables but can act independently. For example having 'precise rules' or a 'strong secretariat' will always have a positive effect on the likelihood of regime implementation, independent of the given context. The above conceptual framework serves as a simple representation of determinants of regime performance and the categorization of the identified context variables into 'problem structure', 'regime environment' and '(state) actors' provides some structure and overview. Although the rules that underlie the framework find their basis in regime theory, the applied categories are not commonly recognized as separate entities for which, e.g., associated aggregate indicators can be envisioned which express their importance for regime performance. Besides, also the aspect of time is not explicitly addressed in this framework, since we could not find clear and robust information in the literature on the temporal dynamics involved, although it is obvious that regime formation and implementation are processes in time. A similar remark holds for interactions between the input variables (determinants): as we did not find clear information in literature on the influence of interactions, we so far treated all input variables as independent entities. The literature review leading to the collection of knowledge rules did not provide equal amounts of information for the different categories of input variables on their influence on the likelihood of regime formation and implementation. As a consequence the four categories of input variables do not have an equal number of variables in the model and not an equal number of rules in the fuzzy rule base. Categories with many variables therefore contribute relatively more often to the overall score of likelihood of regime formation and implementation than categories with few variables. Since the literature does not provide clear information on the relative importance of one category of input variables over the other, we decided not to put additional weighting to these categories in the model. Additionally, not all relevant literature on cooperation between actors (such as economic or psychological theory) were reviewed, so that the concepts included here are derived from a limited and clearly defined field of research on international institutions, thus guaranteeing a minimum coherence in terms of underlying ontological and epistemological assumptions. We, however, acknowledge that our theoretical bias towards neo-liberal institutionalism (regime theory) excludes other plausible accounts of the likelihood and effectiveness of international cooperation. Lastly, the rules included are by no means meant to be conclusive, but rather a preliminary list that could be added to or refined.

2.2.3 *Review of modeling approaches*

Sprinz and Wolinsky-Nahmias (2004), and Young et al. (2006) indicate a broad range of methods and approaches that can be helpful in studying international regimes, institutions and governance issues. In our specific search for modeling methods enabling to code the current body of knowledge on regime effectiveness that we retrieved from environmental regime theory, we have considered a number of well-known formalized modeling methods, including system dynamic modeling, game theoretic modeling, agent-based modeling, qualitative reasoning models and qualitative simulation models. In judging which method to choose, we primarily focused on the question whether it could offer good possibility to deal directly with the extracted regime theoretic knowledge, taking also implicitly into account its potential to 1) allow systematic analysis and comparison of environmental regimes, 2) deal with often ambiguous social science concepts, 3) deal with quantitative and qualitative knowledge from different sources as well as with missing knowledge and uncertainty and 4) allow for incorporation of the behavior and interactions of actors and institutions.

System dynamic modeling describes and simulates the feedback processes in complex dynamic systems (Sterman, 2000). Stylized system dynamical models have been used to study institutional aspects in man-environment interactions (Anderies, 2000; Good and Reuveny, 2006) while at a larger scale the system dynamic model International Futures (Hughes, 2001) covers aspects of socio-political structures and processes. Depending on complexity and required detail system dynamic models often require a large amount of quantitative data for development and testing, especially if one wants to include information on the heterogeneity of actors and on the diversity of their relationships and behaviors in modeling social systems. Aside from these restrictions, we judged the system dynamics approach as inappropriate for our current modeling purpose since the very rationale for applying this approach was missing in our case: the knowledge rules on regime effectiveness that we extracted from environmental regime theory did not explicitly cover any dynamic feature in regime formation and implementation.

modeling methods like game theory and agent-based modeling focus especially on the behavior of actors and their mutual relationships. These methods have been used to study how institutions constrain the choices available to actors or influence the interaction with their environment (Gotts, Polhill, and Law, 2003; Janssen and Ostrom, 2006) and how bargaining and cooperation can lead to coalition formation in international environmental treaties (Finus, 2008; Kilgour and Wolinsky-Nahmias, 2004). Although these methods have at places been criticized for the sometimes rather restrictive and unrealistic description of the behavior of actors (Green and Shapiro, 1996), the difficulties to understand the logic of their results (Axelrod, 1997; Earnest and Rosenau, 2006) and their limited ability to describe hierarchical systems based on authority (Bousquet et al., 2001; Earnest and Rosenau, 2006), these objections can to a certain extent be resolved. The main reason for us not to use game theory or agent-based modeling in this stage was that the body of rules that we extracted from environmental regime theory was reflecting the behavior of actors and their interests and drive to cooperate only in a very general way. It presently didn't yet provide sufficient detailed information to enable a sensible use of game theory or agent-based modeling.

Qualitative reasoning and qualitative simulation (Forbus, 2004; Kuipers, 1994) provide means and tools for formally representing and reasoning with incomplete, uncertain knowledge that is difficult to quantify. Qualitative simulation uses 'common-sense' system-dynamical insight and expertise in combination with consistent reasoning, as a suitable tool for this purpose. It has, e.g., been applied in research on the 'syndromes of global change', where the complex processes of social and environmental changes are described as syndromes encompassing various better or lesser understood variables (Biermann, Petschel-Held, and Rohloff, 1999; Petschel-Held et al., 1999). Main reason for not considering qualitative simulation as our present choice for modeling was similar as for system dynamic modeling: the knowledge rules that we took as a basis for studying regime effectiveness did not explicitly cover dynamical aspects.

A different branch of qualitative reasoning methods, which focuses less on the dynamical characteristics of the system and is therefore more suitable for our purpose is formed by methods linked to Bayesian belief networks (BBNs) and fuzzy logic (FL). These methods also focus on a more global characterization of the system, in terms of relationships between parts/components/subsystems, combining quantitative as well as qualitative knowledge on these characteristics and using some form of inference rules, being probabilistic (as in BBNs) or logico-linguistic (as in FL). E.g., Qualitative Bayesian Belief Networks have been used to study the role and effects of institutional settings in complex water management issues (Saravanan, 2008), and enable also linkage with system dynamic modeling (Vankouwen, Schot, and Wassen, 2008).

Given that our conceptual framework is based on rather qualitative rules expressing to what extent certain determinants influence the likelihood of regime formation and implementation, the use of such a qualitative probabilistic modeling set up might seem rather appropriate to put our conceptual framework into a model. This would however require that experts on environmental regimes have a certain familiarity in expressing their knowledge on environmental regimes adequately albeit qualitatively in terms of chances/probabilities reflecting certain levels of likelihood, which was considered as a somewhat restrictive condition. Though our conceptual framework in speaking of likelihood of regime formation and implementation certainly contains some probabilistic elements, its focus is actually more on the logical and rather qualitative rule-base. Therefore we consider the fuzzy logic approach as an adequate candidate to formalize this framework into a model, especially since the set of rules serves as direct input basis to this form of modeling. In the next sections it is elaborated how we have operationalized this.

2.2.4 *Our approach: fuzzy logic*

Fuzzy logic reasoning (Zadeh, 1965), in fact meets all the abovementioned criteria. A basic property of fuzzy logic is that it uses non-numeric linguistic variables to express vague or imprecise knowledge which cannot be stated in exact numerical form, e.g., 'public concern is high', and enables further processing of this information on the basis of fuzzy rules which link these vague/fuzzy propositions, thus performing computation with words rather than with numbers. It therefore provides a systematic and transparent way of dealing with the propositions obtained from environmental regime theory on regime effectiveness and its determinants, for which a straightforward quantification was impossible. Furthermore, fuzzy logic provides formalization methods to combine and integrate quantitative and qualitative knowledge from different domains and sources and yield concrete answers which could eventually be related to quantitative approaches in integrated assessment analyses. Fuzzy logic methods have been used to combine social and ecological knowledge in sustainability assessments (Acosta-Michlik et al., 2008; Phillis and Andriantiatsaholiniaina, 2001).

In the following sections we explain the various steps that we took in developing a model to assess regime effectiveness in a 'fuzzy logic' manner (Figure 2.1). For the implementation of our model we have used the freely available fuzzy logic toolbox of Babuska 1993, which runs under Matlab.

2.2.5 *Constructing a fuzzy model*

The next step was to translate the conceptual framework into a formal model that would enable a systematic analysis and evaluation of the factors determining regime effectiveness, and facilitate future integration of this knowledge in integrated assessment analyses.

Quantification and fuzzification of variables

If variables in the conceptual framework can be quantified in one form or another, the first step in our modeling approach is to translate this quantitative information into linguistic categories (like, e.g., 'public concern is high', or 'public concern is medium') which form the core elements in the fuzzy logic approach. This process is called fuzzification. Not only for explicitly available quantitative variables this is done, but also for variables on which no quantitative information (e.g., in terms of 'measurable' indicators) is available: the latter variables are first given numerical values based on expert-judgement, e.g., on a scale running from 0 to 10 as a means to express their magnitude, before establishing the fuzzification. In the end, also the explicitly available quantitative variables are transformed to that same scale to provide for ease of use and reference a common basis for the subsequent translation into linguistic categories.

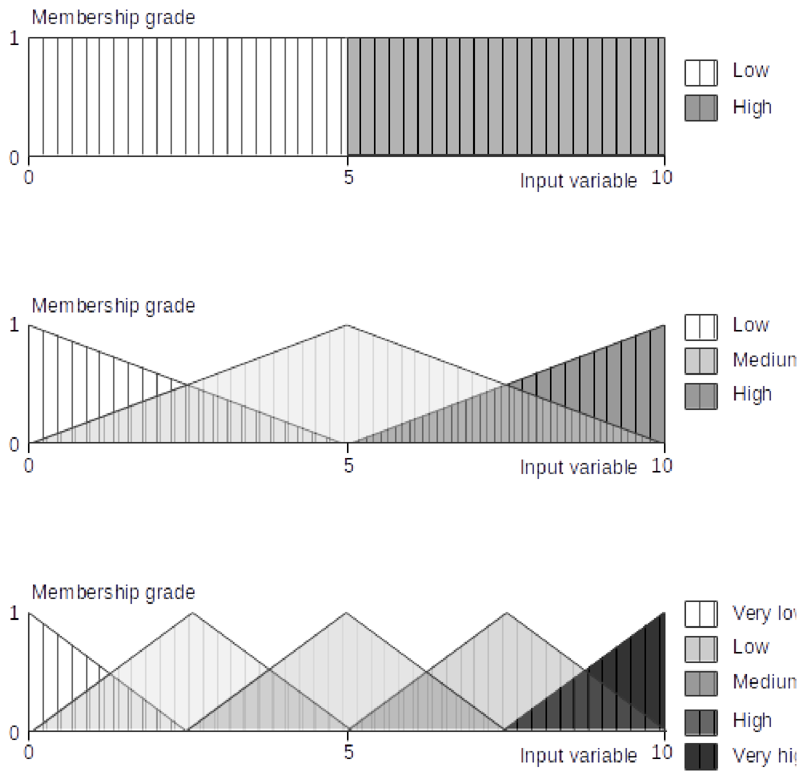

Figure 2.3: Membership functions for the binary input variables (upper panel), numerical input variables (middle panel) and output variables (lower panel) used in the fuzzy model.

For this translation into linguistic categories we have employed three categories to characterize real-valued input variables, viz. 'low', 'medium' and 'high'. Each of these categories is represented by a membership function (Figure 2.3), which assigns to each value of the input variable a membership grade between 0 and 1 which expresses to what extent this specific value belongs to the specific linguistic category (i.e. to what extent the value can be considered as 'low', 'medium' or 'high'). For example, if the variable 'asymmetry' has (transformed) numerical value 6, it will be categorized as mainly 'medium' (membership grade 0.8) and a little bit as 'high' (membership grade 0.2), while not as 'low' (membership grade 0). For output variables we employ 5 linguistic categories, viz. 'very low', 'low', 'medium', 'high' and 'very high', to obtain more differentiation in our statements on the likelihood of regime formation or implementation. This completes the fuzzification processes for numeric values. For variables that initially have binary values instead of numerical ones, the fuzzification process is straightforward: they are assigned to two linguistic categories 'yes' and 'no' (or equivalently, present/absent, true/false), and subsequently are given membership grades 1 or 0 for the complementary (non-overlapping) membership functions characterizing category 'high' or 'low' (Figure 2.3, upper panel). In fact binary variables are treated as crisp (and not as fuzzy).

Fuzzy rule base construction

The heart of the fuzzy logic method is the fuzzy rule base (available as Supplementary information [1]), a set of reasoning rules that reflects the knowledge on the system of interest, which in our case is based on the robust findings from international environmental regime literature as represented by the conceptual framework and its associated knowledge rules. Below we present a number of typical examples which illustrate how we have created the fuzzy rule base for our model from the knowledge rules associated to the conceptual framework. The fuzzy rules are described in terms of IF-THEN statements, and relate input variables with the output variable.

1. Some knowledge rules in the framework describe the effect of a single variable on the output variable 'regime forma-

1 The online version of this article (doi:10.1016/j.envsoft.2012.08.004) contains supplementary material, which is available to authorized users

tion' or 'regime implementation', for example rule C3 (Appendix A.1) "Great asymmetry of powerful states' interests decreases likelihood of regime formation". Translation of this rule into a fuzzy rule could, e.g., yield IF asymmetry is high THEN likelihood of regime formation is very low.

2. Some rules in the framework describe the combined effect of a two input variables, for instance rule B2 (Appendix A.1) "If a problem is marked with great asymmetry of powerful states' interests, differentiation of rules increases likelihood of regime formation". These combined rules are in fact further specifications of single rules to illustrate the mitigating effect of design variables ('differentiation of rules') on the negative impact of context variables ('asymmetry of powerful states'). In 'translating' this rule B2 to the fuzzy rule base the effect of single context variables on the output variable was valued as 'low' and not as 'very low' to account for the mitigating effect of design variables. Rules from the framework that contained the same input variable were combined in the fuzzy rule base to prevent unnecessary rule conflict. Translation of rules B2 and C3 would therefore yield the following set of fuzzy rules

   a) IF asymmetry is high AND differentiation is true THEN likelihood of regime formation is low

   b) IF asymmetry is high AND differentiation is false THEN likelihood of regime formation is very low

   Design variables in the regime formation phase almost all act as mitigating variables (Table 2.1) which give them less importance than context variables, since they only contribute when context variables have a negative impact but not when context conditions are neutral or favorable. In the regime implementation phase, on the other hand, many design variables act independently upon regime performance variables (Table 2.2), rendering them a bigger contribution to likelihood of regime implementation.

3. Since literature on regime theory did not provide information on regime effectiveness in all possible situations, the rules in the conceptual framework do not cover all possible linguistic values of the input variables. E.g., in situations where asymmetry is low or medium, there are no rules available, and thus it is not clear whether these specific conditions

affect the likelihood of regime formation or not. A possible approach would be to only include rules in the rule base for situations on which we have information, but this would lead to many situations were no inferences can be made. As a consequence the model results would then be based on only few variables which are assigned disproportionate importance. One can argue that this is a rightful consequence of our limited knowledge on the effects of some determining factors. However, since we wanted all variables to be included in every model analysis, we have chosen for an alternative approach, where we have artificially extended the fuzzy rule base by adding rules, e.g., assuming a neutral likelihood of regime formation in situations on which no information was available:

a) IF asymmetry is medium THEN likelihood of regime formation is medium

b) IF asymmetry is low THEN likelihood of regime formation is medium

About one third of the rule-base currently consists of rules assuming a neutral likelihood. A sensitivity study (available as Supplementary information) shows that results of the limited rule-base and the artificially extended rule-base are comparable. However, we consider the rule-base as an ongoing body of work and hope regime theorists feel encouraged to perform additional study and improve the rulebase by replacing artificially added rules by genuine knowledge rules.

Fuzzy inference

With the fuzzy rule base thus constructed we could, for a given set of input variables, evaluate the likelihood of regime formation and implementation. This activity is called fuzzy inference and we applied the commonly used Mamdani's minmax inference algorithm (Jang et al., 1997) which works with a simple 'min-max' operation structure: Each fuzzy IF-THEN-rule is activated by first determining the degree of fulfillment of the rule's antecedent, which is equal to the membership-grade of the condition in the IF-part. The inferred implication of the rule's antecedent is established subsequently by redefining the membership-function of the rule's conclusion, i.e., the THEN-part. This redefinition is performed by applying the min-operator, which clips the membership values of the rule's conclusion by the value of the antecedent's degree of fulfillment. Next, in an aggregation process, all activated IF-THEN rules are combined by applying the max-operator on all the clipped membership functions of the activated rule's conclusions. This results in an encompassing membership function which assigns a weighting to each output value (i.e. likelihood of regime formation or implementation); see (Jang et al., 1997) for further details on this inference process.

Defuzzification:calculation of regime effectiveness

The fuzzy inference of the previous step has resulted in an overall fuzzy conclusion on the likelihood of regime formation or implementation, represented in terms of an encompassing membership function obtained in the aggregation process. The final step in our fuzzy logic framework involves the back-translation of this fuzzy information into a crisp value for the likelihood of regime formation or implementation. For this defuzzification we use the 'centre of gravity' method, which determines the specific output value (centroid) which divides the area under the membership function into two equally sized subareas, see (Jang et al., 1997). This final output value gives an numerical indication of the likelihood that a regime will be formed or implemented in a given situation. A low likelihood does not necessarily mean that regime formation or implementation does not take place, but rather indicates that formation or implementation is very difficult in the given circumstances. In addition to this single overall output value, we have also calculated 'partial results' by considering the specific contribution of the subset of rules which are linked to each of the three categories of context variables (State Actors, Regime Environment, Problem Structure). In this way we gain some additional insight in the contribution of each separate category to the final output. An illustration of this is given in next section where we discuss the application of our model. The interpretation of the final quantitative output of the model may be difficult, as there is neither a universally accepted definition nor metric of the likelihood of regime formation and implementation that could be applied in the model. However, the quantitative model results enable systematic and meaningful comparisons between regimes.

2.3 EVALUATION OF THE MODEL

We evaluated the model by analysing three existing regimes, i.e, the Vienna Convention for the Protection of the Ozone Layer (plus the Montreal Protocol), the UN Framework Convention on Climate Change (plus the Kyoto Protocol), the Convention on Biological Diversity and one regime that was never established in the form of a signed treaty, the International Forest Regime. The focus of our model analysis was on the circumstances under which the regimes have been developed and whether the formation and implementation process could be considered easy or difficult for the participating states. As the formation and implementation of these regimes have been studied by different scholars over the past years (e.g., (Barret, 1999; Dimitrov, 2003)), the model results could also be compared with observations in reality. These observations, the input data used in the model analysis and the knowledge captured in the fuzzy rule base do, however, not originate from fully independent sources of information. This analysis should therefore not be seen as a validation of the model results, but rather as an exercise to assess whether the model is able to reflect and reason with knowledge from environmental regime theory.

Figure 2.4 provides an overview of the different steps in the process of applying the model and is explained in more detail in Section 2.5.3, 2.5.4 and the next two sections.

2.3.1 *Collection of model input*

Input data for the analysis of the four regimes were obtained by means of expert-elicitation. Per regime scientists with experience in the field of international environmental policy scored the values of all input variables and as a kind of self-assessment the corresponding levels of confidence that they had in these scores. Variables were directly scored on a scale from 0 (extremely low) to 10 (extremely high) for numerical variables, or 0 (no)/10 (yes) for binary variables (Table 2.1 and 2.2 a and b). All experts' scores are available as Supplementary information. Seven experts scored the variable values for the Climate and Forest regime; for the Ozone and Biodiversity regime respectively six and two experts completed the scoring form. As the forest regime was never established in the form of a signed treaty, only variables on regime formation were included in its analysis.

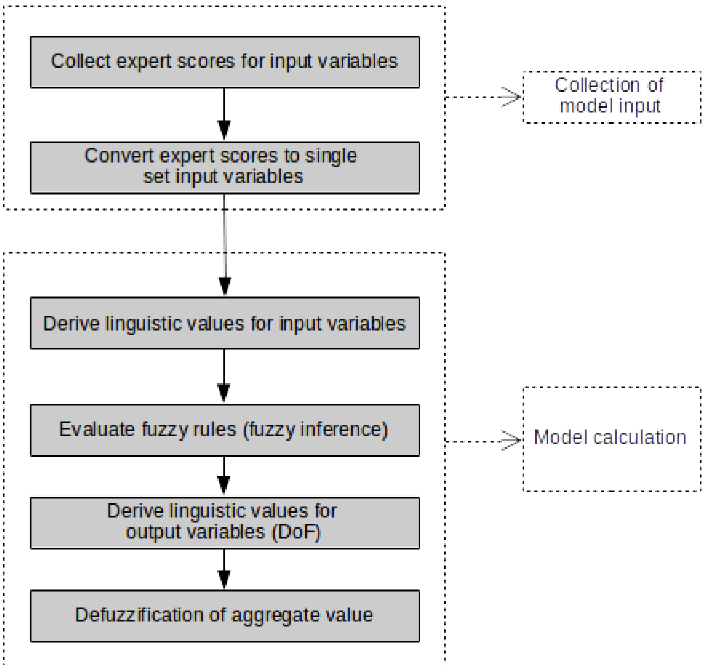

Figure 2.4: Description of the different steps in the process of applying the model to analyze four existing regimes.

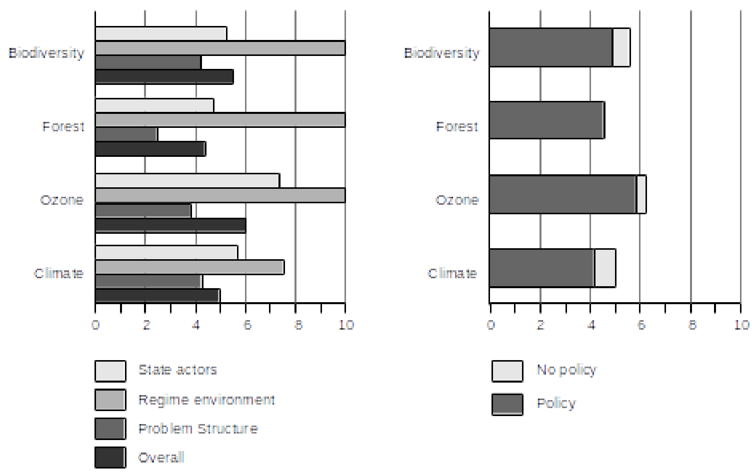

Figure 2.5: Scores of four regimes on likelihood of **regime formation** (based on weighted average of expert responses). Left panel: overall and partial scores on three categories of context variables. Right panel: contribution of design variables (overall scores are compared to a situation without policy measures).

Per regime scores of all experts were converted to a single set of input values for the model by calculating the weighted averages of the expert scores, where the weighting factors were based on the corresponding confidence levels specified by the experts. The values of the input variables showed much diversity in expert opinions. As averaging expert values may hide this diversity and give an incomplete picture of the input information, we also performed model calculations with the separate sets of expert responses (Supplementary information). Although we were not able to test the statistical significance of our results, these showed no systematic under or overscoring by the experts. We therefore considered expert scores both on the level of overall likelihood and in the different categories comparable.

Finally we included also an artificial, i.e., not based on expert values, situation called 'no policy' to assess the effect of policy measures taken in the formation and implementation phase of the regimes. Input values for the 'no policy' situation were derived by for all design variables choosing values that would have most negative impact on likelihood of regime formation or implementation, while for all context variables their regular values (i.e. the weighted averages) were used.

2.3.2 *Model calculation*

We calculated model results following the different steps that are explained in Figure 2.4 and in Sections 2.5.3 And 2.5.4. During this process we made some additional decisions regarding the methodology. If none of the experts scored a certain input variable, the corresponding rules in the fuzzy rule base were excluded from the model analysis. In the analysis of regime formation the rules on 'powerful pushers', 'powerful laggards' and 'positive issue linkages' were excluded for the biodiversity regime, and rules on 'differentiated rules' and 'informal agreement' were excluded for the forest regime.

For pragmatic reasons we have decided to choose the same form of membership functions for all input variables (Fig. 3) to offer the experts a common ground for scoring all variables. A more differentiated choice, which would allow for different choices in shapes and positioning of the memberships for the various input variables, would add a rather complicated and demanding task in the scoring sessions for the experts. After performing some initial

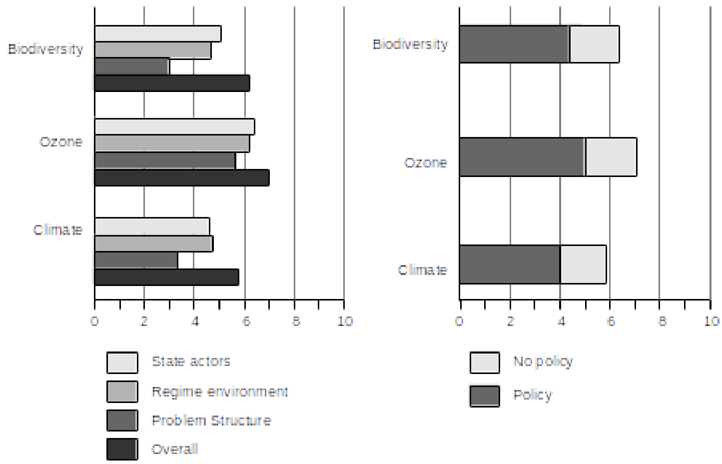

Figure 2.6: Scores of four regimes on likelihood of **regime implementation** (based on weighted average of expert responses). Left panel: overall and partial scores on three categories of context variables. Right panel: contribution of design variables (overall scores are compared to a situation without policy measures).

sensitivity analyses (available as Supplementary information) to study the effects of other feasible membership function choices, we finally decided to stick to our pragmatic choice of uniformity.

2.3.3 *Model results*

Representation of model results

The final overall output values of our model, i.e., likelihood of regime formation and implementation (Figures 2.5 and 2.6; left panel) give an indication of the circumstances under which the four regimes were formed and implemented. A value of 0 means that circumstances are highly unfavorable and consequently the formation or implementation process will be extremely difficult for the participating states, whereas a value of 10 indicates highly favorable circumstances and a very easy formation or implementation process. In addition to these single overall output values we

Table 2.3: Summary of the degree of fulfillment tables for regime formation in the four analyzed regimes

Context variables	Likelihood of regime formation			
	Biodiversity	Forest	Ozone	Climate
Problem Structure				
Regulation costs	±	−	±	−
Public concern	±	±	±	±
Systemic/cumulative problem	−	−	+	+
Uncertainty	±	±	±	±
Collaboration/coordination problem	±	−−	±	±
State actors				
Asym. powerful states	±	±	±	±
Asym. important states issue area	±	±	±	±
Support of powerful states	±	±	+	±
Number of economic sectors	+	+	+	±
Homogeneous states	±	±	++	±
Urgency	+	+	+	+
Cumulative cleavages	−	±	±	−
Powerful pushers		±	++	++
Powerful laggards		−	±	±
Regime environment				
Preceding agreement	++	++	++	±
Scientific advisory bodies	++	++	++	++

have also calculated 'partial results' (also represented in Figures 2.5 and 2.6; left panel) by considering only the contribution of the subset of rules which are linked to each of the three categories of context variables separately (State Actors, Regime Environment, Problem Structure), which give some additional insight in the contribution of each separate category to the final output. More detailed information on the determinants of effectiveness of the four regimes can be obtained from the fuzzy rule base by looking at the Degree Of Fulfilment (DOF) tables, which are summarized in Tables 2.3 and 2.4. Full DOF tables are available as Supplementary information [2].

The contribution of design variables to the process of regime formation and implementation is shown by comparing the scores

2 The online version of this article (doi:10.1016/j.envsoft.2012.08.004) contains supplementary material, which is available to authorized users

Table 2.4: Summary of the degree of fulfillment tables for regime implementation in the three analyzed regimes

Context variables	Likelihood of regime implementation		
	Biodiversity	Ozone	Climate
Problem Structure			
Regulation costs	±	±	−
Collaboration/coordination problem	−	±	−
Systemic/cumulative problem	±	+	+
State actors			
Participation government	+	±	++
Participation powerful states	±	+	±
Number of economic sectors	+	+	±
Outvoting of important states	−	±	−
Asym. powerful states	±	±	±
Asym. important states issue area	±	±	±
Regime environment			
Institutional framework	±		±
Negative interplay	−	±	−
Positive interactions	++	++	++

of the overall output values of the regimes to a situation without policy measures, i.e., where all design variables have the most unfavorable value (Figures 2.5 and 2.6; right panel).

Description of model results

Overall scores show that formation was easiest for the ozone regime and most difficult for the forest regime (Figure 2.5; left panel). Implementation of the Climate regime went less smoothly than implementation of the ozone and biodiversity regime (Figure 2.6; left panel). However, the overall scores of the regimes in the formation and implementation phase do not differ much. The partial scores in the three categories of context variables provide some insight in the type of factors contributing to the effectiveness of the four regimes. These partial scores are different for the formation and implementation phase, as the categories in both phases contain different variables (Table 2.1 and 2.2).

The problem characteristics of deforestation were least favorable for regime formation, while characteristics of ozone pollution were most favorable for both regime formation and implementation (Figures 2.5 and 2.6; left panel). Deforestation and biodiversity loss are cumulative problems, i.e., they are local in nature, but are globally replicated, which decrease the likelihood of regime formation and implementation (Table 2a and b). Ozone pollution and climate change, on the other hand, are systemic problems with both global causes and effects, which encourages regime formation and implementation. Furthermore, regulation costs of the ozone and biodiversity regime were lower than those of the climate and forest regime (Tables 2.3 and 2.4).

State actors involved in the ozone regime were more harmonious (Figures 2.5 and 2.6; left panel), represented by homogeneity and limited asymmetry (Tables 2.3 and 2.4), compared to those involved in the other regimes, which increased likelihood of regime formation and implementation. In contrast with the other regimes, the forest regime had only laggards and no pushers, i.e., that there were no state actors trying to achieve cooperation and regime formation and hence that regime was never turned into a signed treaty. Conditions concerning state actors were less favorable for the climate regime due to the higher number of economic actors involved in regime formation and implementation. The regime environment in the formation phase was least favorable for the climate regime (Figure 2.5; left panel) as therewas no preceding agreement or policy dealing with this issue (Table 2.3 and 2.4). Institutional interactions had both positive and negative effect on the implementation of the climate and biodiversity regime, but merely positive influence on the implementation of the ozone regime.

The influence of design variables, i.e., policy measures, in the regime formation phase contributed little to the overall scores on likelihood of regime formation (Figure 2.5; right panel). Measures taken in negotiation process of the ozone and climate regime, like side payments and the formation of an initial framework treaty, had most effect. Policy measures taken in the implementation phase had a bigger impact on the overall score. Likelihood of implementation of all three regimes was increased by measures like legally binding rules, a strong secretariat and mechanisms for increasing public awareness.

2.3.4 *Reflection on model results*

Finally we compared the established model results with observations in reality on the formation and implementation process of the four regimes. Model results indicate that several components of the problem structure of deforestation are currently unfavorable for regime formation, which reflects the current situation of international negotiations on forest issues. Indeed, deforestation is often considered a prime example of failed regime formation (Dimitrov, 2003). At the same time, our analysis also suggests ways to mitigate the barriers to regime formation: for example, high interest asymmetry between state actors on the problem of deforestation can at least partially be mitigated by rule differentiation. With respect to the climate and ozone regimes, there is a wide consensus that while the ozone regime has resulted in a close-to-complete phase-out of ozone depleting substances (and the corresponding recovery of the stratospheric ozone layer), climate change is still unabated despite the existence of the Kyoto Protocol. Using the scores from the expert judgement, our model identified that several context variables are more favorable in the case of ozone. This observation is in line with authoritative scholarly analysis, which identifies ozone as the more successful example of regime formation and implementation (Barret, 1999). Model results for the four different regimes did not differ much and covered a limited range in the output space, which is a logical consequence of our decision to extend the fuzzy rule base with rules which conclude on neutral likelihood in situations where we lack information (Section 2.5.2). Nevertheless, the model enabled a meaningful comparison between the four regimes and the factors influencing their effectiveness. The set of artificially added rules represent the knowledge that is missing in the literature on environmental regime theory in order to perform a systematic analysis of regime effectiveness. Additional research on these topics is needed to improve the fuzzy rule base and consequently the discriminative power of the model. The results of the model analysis imply that the category 'state actors' has a greater impact on regime effectiveness than 'regime environment', since more variables and rules are involved in this 'state actors' category (Table 2.1 and 2.2). This difference in impact may be a true representation of reality, but may also be caused by a bias in literature for example because factors concerning state actors are easier to study than factors concerning regime environment. Additional study is needed to assess whether the set of rules for input variables should be adapted or a weighting system of the categories should be implemented in the model. Design variables contributed little to the overall scores on likelihood of regime formation, but had more impact in the implementation phase. As such the model reflects our decision to consider them as mitigating factors in the formation process and as independently acting variables in the implementation process (Section 2.5.2). The model thus implies that policy measures taken in the implementation phase have bigger influence on regime effectiveness than measures taken in the formation phase. This fits with our understanding of reality. A serious limitation of our approach is the subjectivity in the scoring of input variables and in the definition of membership functions, which were based on expert knowledge instead of direct empirical evidence/data. Some authors have populated their fuzzy models with quantifiable indicators that were selected for their data availability (Acosta-Michlik et al., 2008; Phillis and Andriantiatsaholiniaina, 2001). As a consequence they left out data-limited variables that may have been important in explaining their understanding of the modelled system (Carr and Kettle, 2009). Most of the variables in our conceptual framework are data-limited, but nonetheless important to include in the fuzzy model as they capture the expert knowledge on the effectiveness of environmental regimes. However, the rather indicative and pragmatic approach that we have currently taken in this study could in the future be improved by invoking better methods for expertelicitation (Cornelissen et al., 2003).

2.4 GENERAL DISCUSSION AND CONCLUSIONS

Model results of the four analysed regimes were in line with observations from reality. This indicates that the model is able to reflect and reason with knowledge from environmental regime theory and that robust general knowledge rules on the effectiveness of environmental regimes can be used to analyze individual regimes. The model can be used as an aid in analysing the effectiveness of existing or future regimes, highlighting which determinants contribute to success or failure, and it enables systematic and meaningful comparisons between regimes and policy measures. Besides, the model can be used for experimenting with the factors that affect regime effectiveness, which may offer useful insights in environmental regime theory and encourage discussions between natural and political scientists.

Formalizing knowledge on environmental regimes in a conceptual framework and a model enhanced its transparency and deductive power as it forced us to be explicit about our choices and assumptions regarding determining factors and their potential effects on regime effectiveness (Krugman, 1997; Powell, 2002). As such, the conceptual framework and model are expected to offer researchers in integrated assessment useful insights in environmental regime theory. However, since formalizing entails the danger of discarding important aspects of an over-complex reality (Biermann, 2007), results of analyses performed with the conceptual framework and model should whenever possible be compared with empirical analysis, and be interpreted withdue regard to the limitations and simplifying choices and assumptions involved in the study.

Developing and using the framework and model also revealed the lacunae in knowledge in environmental regime theory. In order to make a systematic analysis of the effectiveness of environmental regimes there is more information needed on the situations that are currently covered by artificially added rules in the fuzzy rule base, the relative importance of different categories of input variables, the temporal dynamics of the input variables and their mutual relationships and influences. The set of fuzzy rules represents the knowledge base in our model and its form makes it easy to understand, discuss and add to the knowledge that is captured in the model. We consider this rule set as an ongoing body of work. The lacunae in knowledge identified in this chapter may inform regime researchers to further structure and increase their knowledge on the functioning of environmental institutions (Biermann et al., 2010) and encourage them to contribute to the fuzzy rule base.

By making knowledge on environmental regimes explicit and understandable we have taken an important step towards a better integration of social science in integrated assessment. We think that, in integrated assessments, it is necessary to consider both the potential impacts of environmental regimes and their chance of successful formation and implementation. This is in line with studies in sustainability science stating that in order to find effective ways of dealing with global environmental problems integrated assessments should be strengthened in their policy relevance (Reid et al., 2010) and should pay as much attention to social science as to natural science (Kates et al., 2001; Ostrom, 2009). Recent integrated approaches (Delden et al., 2011; Nabel et al., 2011; Schaldach et al., 2011) which combine information from various sources (data, models, experience) and domains (economic, environmental, social) and connect explicitly to policy context, interest groups and end-users, provide useful and promising means to assess the implications of proposed policy options in a broader context. Typically they assume that considered policy options will be(come) active, and they do not explicitly address the question what the feasibility of realizing these policy options is. In our approach by focusing on conditions and factors that determine a successful formation and implementation of environmental regimes we expect to contribute to this often neglected important issue.

We believe, however, that that at this stage it is not yet possible and may be even undesirable, to actually include knowledge on environmental regimes directly into the computer models used in integrated assessments. Attempts to do so (Anderies, 2000; Smajgl et al., 2009) focus on the impact of institutions on human behavior and the resulting impact on the environment, but do not consider explicitly their formation and implementation. Our model and framework, on the other hand, especially focus on the political feasibility of environmental regimes but their scope and variables differ presently too much from those used in integrated assessment models to bridge the gap in a direct way and enable successful and meaningful integration. The current contribution of our model lies primarily in enhancing the interpretation of results from integrated assessment models by examining the political context more explicitly. Furthermore, results from analyses with our model can be used to improve scenario storylines for use in integrated assessment modeling, which could account for the development and existence of promising future environmental regimes and provide the factors that need to be in place for their effectiveness. Currently most scenarios used in assessments lack policy relevance as they are devised by scientists and do not consider specific policy options (Perrings et al., 2011). Summarizing, we believe that the integration of the knowledge on environmental regimes in integrated assessments is still in its early days and requires further attention in the future.