SNF-REPORT NO. 27/2000

Do Numerical Simulation and Optimization Results Improve Management? Experimental Evidence

by

Kjell Arne Brekke

Erling Moxnes

SNF- project no. 3730
"Verktøyeksperiment"

This project is financed by the Research Council of Norway

Centre for Fisheries Economics
Report No. 64

FOUNDATION FOR RESEARCH IN ECONOMICS AND BUSINESS ADMINISTRATION
BERGEN, MAY 2000

ISBN 82-491-0055-7
ISSN 0803-4036

CONTENTS

FOREWORD

This report is produced in cooperation with Statistics Norway. An earlier version of the report is published as Discussion Paper No.228, September, 1988, Statistics Norway. At an initial stage of this project, the project group consisted of Asbjørn Aaheim, Magnus Hatlebakk and the authors. The authors are grateful for the discussions with the other project participants at this stage. We also had very useful discussions with Sigurd Tjelmeland at the Institute of Marine Research about the design of the virtual reality. Thanks also to Solfrid Malo for assistance during some of the experiments. Ådne Cappelen, Øystein Olsen and Karine Nyborg gave helpful comments on an earlier draft of the report. The usual disqualifiers apply. The project was funded by the Research Council of Norway.

JEL Classification: A11, C9, Q22

1. INTRODUCTION

Models of different forms are the core tool of economic analysis. The advises economists give decision makers may represent general insights derived from stylized models. Alternatively, economists may construct complex large scale models to simulate the consequences of different policy option, often without any definite conclusion as to what option is best. The advice may also be based on some intermediate approach. Decision makers rarely follow these recommendations without alteration. Typically, they will blend the results from numerous analyses, focusing on different facets of the real world. To evaluate the usefulness of model based analyses, we would have to take into account how the information provided by the model is "filtered through" the decision makers, and ask the question: Does the model improve the final decision made by the decision maker? This is the question we address in this paper. Note that this question differs markedly from the more usual question: What is the quality or accuracy of the analysis produced by a decision tool?

To answer this question in general is not possible since underlying problems differ. Decisions pertaining to simple, well defined problems, are likely to be greatly improved by analyses that provide precise answers to the questions that decision makers ask. For complex problems that are not easily formulated, and where available analyses built on partial models, it is harder to predict outcomes. However, by studying cases, strengths and weaknesses of decision tools can be revealed. With sufficient numbers of cases, certain patterns could also start to emerge. The derived insights should be beneficial for both modelers and decision makers.

The case in question is the management of fish stocks in the Barents Sea. The management problem is complex and the usefulness of different models is hard to evaluate. Actual management is a consequence of the choice of many different decision makers, with possibly conflicting objectives. Each individual decision maker has access to several sources of statistics, different models, and model based studies. It is hard to identify the contribution from one particular model. Perhaps one could identify dominating schools of thought in different fishing regions of the world, but then, differences among fishing grounds would complicate comparisons. To overcome these problems, we use an experimental approach.

We construct an experiment where students are asked to manage the stocks of cod and capelin in a computer-model of the two major fish stocks in the Barents Sea. The experimental treatments are numerical advises from two different models: a simulation model and a stochastic optimization model. Comparing the results for different groups of students, we identify the contribution from each model. We acknowledge at the outset that we focus the subjects' attention on one or two particular sources of information, without the competition from other advises. A possible effect is that the experiment overestimates the effects of the two decision tools at hand for situations in which they face competition. A first indication of this possibility is actually tested by the experiment itself. The interaction term between the two decision tools measures the effect of competition.

Also note that the experiment focuses on numerical advises from the two types of models for the purpose of quota setting. The purpose is not to test for underlying attitudes or misperceptions that for instance could prevent a quota system from being implemented in the first place. The quota system is taken for granted, and the remaining task is to set appropriate quotas. In this regard the experiment differs from previous experimental studies of renewable resources, e.g. Moxnes (1998b) and Moxnes (1998c).

We have not encountered similar experimental studies of the practical usefulness of models for social planning. A literature seems to be emerging in the management area. Oz et al. (1993) point to the need for experimental studies to assess the benefits of experts systems. Cavaleri and Sterman (1997) and Verstegen et al. (1995) make similar claims for systems modeling and information systems. All three find positive effects of decision support. Webby and O'Connor (1994) find that the usefulness increases with task complexity. They also find no difference between a deterministic and a probabilistic decision tool.

First we present the models and the experimental design, and then the econometric model. Next the results of the experiment and a post questionnaire are presented. Both decision tools are found to have significant positive effects, however for different reasons. Subject strategies are estimated and show interesting deviations from the strategy proposed by the optimization model. Finally we conclude and discuss findings that are likely to be found also in future investigations of decision tools.

2. THE EXPERIMENT

First we describe a two-species simulation model for cod and capelin. We refer to this model as the virtual reality, which the subjects are asked to manage. Then we describe the two decision tools. We have chosen to focus on two different kinds of model concepts to aid decisions: a simplistic two-species stochastic optimization model, and one complex deterministic simulation model consisting of two one-species models. While the optimization tool is a close replication of what has been presented in the literature, the simulation tool is a rough attempt to mimic how this tool is being used for real management. Finally we describe the experimental design.

2.1. The virtual reality

A model of cod and capelin in the Barents sea is taken as the virtual reality. The virtual reality is represented by a two-species, predator-prey model. The model is documented in Moxnes (1992), with minor changes documented in Moxnes and Nyhus (1994). The model has cohorts for both species, both weight and population numbers are represented. Predation is modeled with saturation, i.e. each predator (cod) has a limited capacity to eat the prey (capelin). Recruitments are nonlinear functions of spawning stocks, recruitment of cod is also negatively affected by the amount of juvenile cod, and both recruitments are influenced by randomness. Capelin is assumed to die after spawning. The biological part of the model is to a large extent based on an existing biological model, Tjelmeland (1990). The fishing gear for cod is more efficient for higher age classes than for lower age classes and the selectivity is fixed. The two species are caught independently, and costs depend on fish density and fleet capacity utilization. Capacity utilization also determines unemployment in the fishing sector. The criterion reflects present values plus stock values in the final year minus social costs of unemployment in the two fisheries. No activity (and maximum unemployment) in the cod and capelin fisheries were valued to respectively NOK 75 and 25 million per year. Payoffs to the participants were calculated as NOK 150 times the ratio of the obtained criterion value and the benchmark criterion value obtained by using the strategy suggested by the optimization tool.

2.2. The simulation tool

The simulation tool is a deterministic version of the biological part of the virtual reality, except that the linkage between the two stocks were broken, i.e. we used two one-species models. In all equations for capelin where information about the cod stocks was needed, an historical mean of the cod stock was used, rather than the model's own predicted cod stock, and vice versa. Otherwise, the model and its parameters were identical to the virtual reality with two exceptions. First, the economic part was left out, as it often is in models used for fishery management. Second, the simulation model was initialized each virtual year by that year's uncertain stock estimate. To simplify the programming, the distribution of fish on different age classes was identical to the one in the virtual reality. An excuse for this simplification is that in reality catch data produce rather good estimates of the relative strengths of age classes for which harvest is taking place.

Each virtual year, the simulation model was used to make two 4-year forecasts for each of the two stocks. The forecasts for cod were based on strategies with either 15 or 30 percent catch of the total stock, while for capelin, the two forecasts were based on 40 and 80 percent catch. While the simulation tool made no suggestions about optimal policy, these forecasts could produce certain framing effects in that they could be interpreted as indications of reasonable ranges for yearly catches.

2.3. The optimization tool

The optimization tool was a two species stochastic optimization model with capelin and cod. In continuous time, the growth equations are of Lotka-Volterra type[1]. The criterion to be maximized is the net present value of future catches. The model parameters were estimated from data generated by the virtual reality. The optimization model gives the optimal policy in the form of target escapements, i.e. the optimal stocks after the fishery seasons are over. For a further description of this model and the solution algorithm, see Brekke (1994). The exact

---

1 The growth of cod biomass T_t and of capelin biomass L_t is given as

$$\dot{T}_t = a'(\frac{\tilde{x}_t L_t^a}{L_t^a + \overline{L}})T_t - \tilde{m}_t T_t - E_{Tt} T_t^{b_T}$$

and

$$\dot{L}_t = \tilde{r}_t L_t - m(L_t)L_t - a(\frac{\tilde{x}_t L_t^a}{L_t^a + \overline{L}})T_t - E_{Lt} L_t^{b_L}$$

where r is recruitement, m is mortality, and E is effort. $\tilde{x}$ is a stochastic variable determining predation. $a, a', \overline{L}$, and b_i are parameters. Stochastic variables are marked with a tilda.

The objective is to maximize the net present value of future catches until time t, plus the value of remaining biomass

$$\max \int_0^t \{p_{Tt} E_{Tt} T^{b_T} + p_{Lt} E_{Lt} L^{b_T} - c_L E_{Lt} - c_T E_{Tt}\} e^{-dt} dt + S(T_t, L_t)$$

$S(T_t, L_t)$ is the value of remaining biomass at the end of the optimization period.

choice of optimization model was made to comply with existing literature on this topic, see Mendelssohn (1980).

The optimization model disregard much of the detailed information included in the virtual reality. Especially important is the exclusion of non-linearities and information about the year classes. The optimization incorrectly assumes that the two differential equations keep precise track of the biomass. However, this assumption is false in that capelin that has spawned dies, and that juvenile cod and capelin are not included in the respective biomass measures. Note also that unemployment, which is part of the criteria in the virtual reality, is not taken into account in the optimization model.

The optimal target escapement for capelin was found to be 7.0 million metric tons, For cod the optimal target escapement depends on the stock of capelin. The target is 0.8 million tons at very low capelin stocks. It increases linearly with the capelin stock until the capelin stock reaches 5.0 million metric tons, at which point the target is 1.35 million tons. For higher capelin stocks, the target escapement is constant at 1.35 million tons. The students that had access to this model were informed about the optimal target escapements and about the dependence of the cod target on the capelin stock.

Just to illustrate that the optimization model is not very close to the global solution for the virtual reality, we simulated the model with an adjusted strategy. The adjustment represents one first, somewhat random attempt of taking account of non-linearities and measurement errors that have been found to matter for the optimal solutions in less complex models than the virtual reality, Moxnes (1998a). We simply assume that the cod quota equals 20 percent of the cod stock and that the capelin quota is zero for capelin stocks at or below 3.0 million tons, and that the capelin quota increases with 50 percent of the increase in the capelin stock above this level. Simulations were carried out over the 16 different realizations of the random variables used in the experiment. On average the adjusted strategy beats the optimization strategy (the benchmark) by 30 percent.

Thus, the optimization model does not provide the best strategy we are able to develop, but it is the optimal strategy given the stated simplifying assumptions used in existing literature. As in reality, decision makers will only be informed about strategies that are optimal subject to some simplifying assumptions. In real world decision problems, an exact copy of reality is not available for simulations to test alternative strategies. Hence, even though better models are likely to lead to better strategies, one will not know with certainty that for instance a 30 percent improvement can be obtained. In some cases the simpler models could be the better ones.

2.4. Experimental design

A three by three factorial design was used. The two types of decision support represent the first two factors. The third factor was initial conditions, high or low stocks of both cod and capelin. For the economically most important species, cod, the low initial stock was close to the target level from the optimization model, while the high initial stock was from 60 to 80 percent above this level. The realization of the random variable varied among the subjects. However, the same 16 realizations were used for all four combinations of the two types of decision support. The realizations of the random variable will be viewed as a covariate.

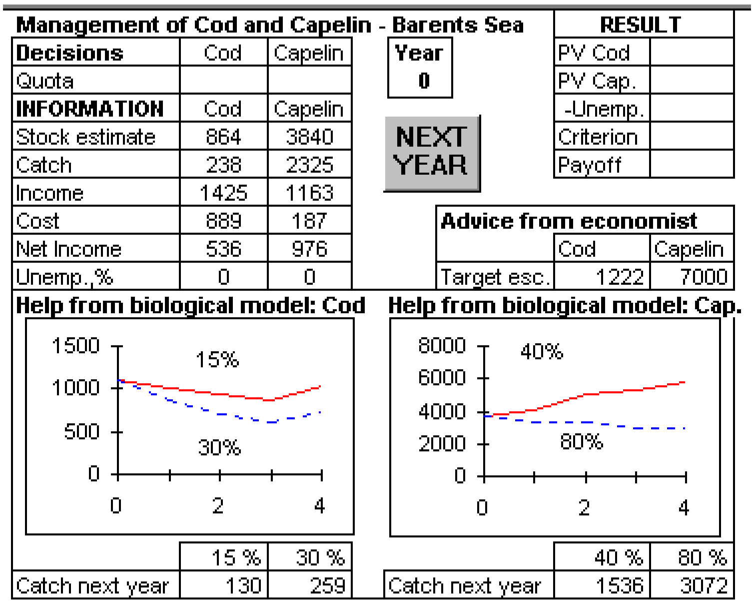

Figure 1: The computer screen.

The screen, as it appeared to subjects who got both instruments, is shown in Figure 1. The upper left quadrant shows data from the virtual reality. These data were updated each virtual year. Information about estimated stock sizes reflected the true stock sizes in the virtual reality plus a random error term. Information was also given about last year's catch, costs, net income, and unemployment. The subjects had to fill in the fields for quotas, and press the next year button to advance to the next year. The upper right hand corner revealed criterion values and payoffs after 25 years of management. Payoffs varied from NOK 38 to 380, i.e. from about 0.5 to 5 times a normal hourly wage for students.

The results from the two decision tools were either presented as shown in Figure 1 or they were blanked out for those who were not availed with one or both to the tools. The forecasts from the two one-species biological models were presented under the heading "Help from biological model". The target escapements from the optimization model were presented under the heading "Advice from economist".

In total 64 students participated in the experiment. Half of the students were from Bergen and the other half from Oslo. Around 50 percent were in the first or second year of their economics or business and administration studies. The remaining 50 percent were at more advance levels of the economics study, some with backgrounds in mathematics, technology and agriculture. Students were chosen for practical reasons. They were randomly selected for the different treatments.

The students were novices with respect to the actual management problem. Hence they do differ from experienced managers who are familiar with details of the analyses, and who know the positions of relevant interest groups. We can only speculate how students might behave differently from real managers. Novices with little knowledge of the system should benefit more than experts from the tools. Novices with a positive attitude towards analytical tools (as our subjects) should be expected to be less skeptical to the tools than managers. Real decision makers are presented with other goals, constraints, and information than the subjects in the experiment. They might even be presented with other, competing decision tools. Hence, for these reasons real decision makers are likely to put less weight on the two selected types of decision support than inexperienced students. Thus, the benefits of the tools could be overestimated in the experiment. On the other hand, lacking experience with the tools could also imply that they are not used to their full potential. While there are reasons to expect differences between students and actual decision makers, previous experiments indicate that they could be small and insignificant, at least when participating in a given experiment, see e.g. Bakken (1993) and Moxnes (1998b). Since the students were not acquainted with the problem at hand, and since there is no media focused controversy over multi-species management in the Barents Sea, the experiment should not suffer much from role playing (subjects making use of preferences, information, and strategies outside of experimental control). We found it unnatural and difficult to disguise the rather complex task as a neutral management problem.

The subjects received a written information (in Norwegian), and were encouraged to ask questions. Few questions were asked. The following pieces of information were given: They were told to see themselves as social planners with full control over the fisheries, and that historical harvests had varied quite a lot from year to year. The items on the computer screen and the technicalities of the experiment were explained. The criterion and its relation to the personal payoffs was explained. Facts were given about the virtual reality: the biology (predation, maturation delays, lifetimes, mortality, recruitment), harvesting (two fleets, gear selectivity, unemployment), economics (prices, fixed and variable costs), and randomness (recruitment, mortality, predation, variable costs, resource measurements).

Depending on treatment, the subjects also received explanations of the decision tools. In both treatments, subjects were told that the tools were based on simplified representations of the virtual reality. Regarding the simulation tool, they were told that the tool did not represent predation, randomness, correct initial conditions, and economics. Otherwise the model was a perfect representation. The scenarios and numbers shown on the screen, as well as the underlying strategies were explained. Regarding the optimization tool, subjects were told that the tool did not capture saturation in predation, had no age classes, disregarded measurement error, did not have continuous harvesting over the year, and put no weight on unemployment in its criterion. The target escapements shown on the screen were explained. Finally, the subjects were advised to decide for themselves to what extent they should follow the advice.

A pre-questionnaire was used to check that the subjects did understand what was to be maximized, and to check that they understood how the decision supports deviated from the virtual reality. Answers showed that the subjects understood what to maximize, although two subjects wrote that they were supposed to maximize quotas. Most subjects were also able to point out major differences between the virtual reality and the tools. Somewhat different simplifying assumptions of the tools were reported. A few subjects brought in their own general ideas about differences between tools and realities. We also asked for their age and their experience with and belief in economic models for the purpose of public management. The answers showed no significant differences between groups of subjects selected for the different treatments. Concerning their belief in models, the average rating was 3.5 on a scale from one to five (63 percent).

A post-questionnaire asked the participants about their willingness to pay for having access to each of the two decision tools in case they were to participate in a similar experiment. We also asked them to say if they tried to smooth quotas from year to year, if they tried to stabilize the stocks at the level in the initial year, and to what extent they saw the experiment as a rewarding learning experience. All three questions on a scale from 1 to 5.

3. ECONOMETRIC MODEL

Let Z_i denote the criterion value that person i achieved, and let Y_i denote the benchmark criterion value he would have achieved had he used the proposal from the optimization model without adjustment. (This criterion value can be computed irrespective of whether the person had access to the optimization model or not.) We assume that the criterion value depends on whether the student had access to the simulation model, represented by the dummy S_i, the optimization model O_i or whether the initial stocks were high or low H_i. The criterion value further depends on two stochastic variables. One representing the stochastic variables in the bioeconomic model, represented by the residual u_i and finally the management skill of individual i, represented by v_i. We thus assume

$$Z_i = \tilde{f}(O_i, S_i, H_i) + (1+e)u_i + v_i \tag{1}$$

where e is some parameter to be explained below.

A similar model will apply to the criterion value that i would have received if he had used the results of the optimization model without any adjustments, but then skill and access to the different models would not matter. Thus we define

$$Y_i = k'' + c''H_i + eu_i \tag{2}$$

To allow the bioeconomic uncertainties, represented by u_i, to have different impact on Y_i and Z_i we apply different parameters, $(1+e)$ and e respectively, but for simplicity normalized such that the difference is 1.

Let X_i denote the criterion value in excess of the benchmark, i.e. $X_i = Z_i - Y_i$. Then

$$X_i = f(O_i, S_i, H_i) + u_i + v_i \tag{3}$$

Note that according to this model

$$Z_i = aY_i + f(O_i, S_i, H_i) + u_i + v_i \tag{4}$$

with a=1. Testing the hypothesis a=1 is thus a test of the model above.

The design of the experiment requires some special considerations on how to handle the residual u_i. To reduce the noise in the comparison of models, we picked the same realization of the stochastic variable in the virtual reality for all different combinations of models. With 64 students in the experiment and with four different combinations of models, only 64/4=16 different (and independent) realizations of the stochastic variables were used in the virtual reality. Hence there are only 16 different realizations of u_i while there are 64 realizations of v_i. Thus the total residuals $u_i+ v_i$ are not independent. Still, estimating $\hat{k}''$ and $\hat{c}''$, we can approximate the residuals as

$$e\hat{u}_i = Y_i - \hat{k}'' - \hat{c}'' H_i \tag{5}$$

We then included this constructed variable as an explanatory variable in a regression version of (4). This turned out to have negligible effects on the results, and therefore we present only the results for the simplest equation where $u_i+ v_i$ is treated as an independent residual.

4. EMPIRICAL RESULTS

First we analyze criterion values, next we estimate decision strategies, then we discuss the results and finally the post-questionnaire.

4.1. Analysis of criterion values

We first estimate equation (4) to test the hypothesis that a=1.We find $\hat{a}$=0.98, not significantly different from 1.0, and with R^2=0.93. The hypothesis is clearly not rejected. The other estimates were very close to the ones found below.Thus the data are consistent with our model. We next estimated (5), to compute an estimate of eu_i. Including the estimated eu_i as a explanatory variable in (3) we found that $e \approx 10$. Hence more than 90 percent of the variation induced by the stochastic terms of the virtual reality is included in Y_i. For the error term in the X_i-equation, we find that the variance of v_i is almost 20 times that of u_i, and this explains why the correlation in error term does not influence the estimate. This finding also implies that the variation in X_i, is mainly due to skill, and not luck, whereas the variation in total score Z_i is more due to luck than to skill, since $(1+e)u_i$ has more than five times the variance of v_i.

Table 1: Means of X_i for different treatment combinations.

	H_i=0		H_i=1	
	O_i=0	O_i=1	O_i=0	O_i=1
S_i=0	1619	1877	-2251	2433
S_i=1	4340	4737	126	2898

Table 1 summarizes the results in terms of means for X_i for different treatment combinations, i.e. criterion values minus benchmarks. Mostly positive values of X_i indicate that subjects do better than the benchmark. We note the following tendencies. In the case with low initial stocks, there is a considerable effect of access to the simulation tool, while the optimization tool has only a minor effect. In the case with high initial stocks, the pattern is nearly reversed: there is a considerable effect of having access to the optimization tool, while the effect of simulation depends on the access to optimization. Without optimization the effect is considerable, while it is only minor in the case optimization is available. We also note for the cases without access to any of the tools, subjects do better than the benchmark when initial stocks are low, and they do worse in case of high initial stocks.

Table 2: ANOVA results for dependent variable X_i.

Variable	Estimate	t-ratio
Intercept	1973*	4.87
Optimization	1014*	2.51
Simulation	1053*	2.61
High stock	-1171*	-2.90
Opt.*Sim.	-222	-0.55
Opt.*High	850*	2.11
Sim.*High	-342	-0.85
All	-256	-0.64

All *-marked estimates are significant at a 5 percent level.

To estimate the different effects we conducted an ANOVA analysis. This corresponds to a regression of X_i, with dummies S_i, O_i, H_i, S_iO_i, S_iH_i, H_iO_i and finally $S_iH_iO_i$. The results are reported differently from the regression case, as deviation from the appropriate sample mean. The grand mean is positive, and there is a significant effect of access to either one of the two models. Moreover, students do worse when initial stocks are high. Finally, the benefit of the optimization model is higher when the initial stock is high. All the other estimates are insignificant.

For the ensuing discussion, we find the following linear regression result for the significant parameters (since the other effects were clearly insignificant, they are of little importance for the parameter values)

$$X_i = 1927 + 327 \cdot O_i + 2106 \cdot S_i - 4042 \cdot H_i + 3401 \cdot H_i O_i + u_i + v_i \qquad (6)$$

To see if the relative contributions of present values and unemployment to the criterion value change with the treatments, we perform another ANOVA. Fractions are formed where the present value of cod, the present value of capelin, or the costs of unemployment is divided by the total criterion value. Table 3 shows the significant results. The optimization tool leads to a considerable increase (decrease) in the value coming from the cod (capelin) fishery. This seems likely since the optimization tool reflects knowledge about the normally higher value of capelin as food for cod than for commercial harvesting. The simulation tool leads to a reduction in the costs of unemployment.

Table 3: Summary of ANOVA analyses of contributions to total criterion values.

Fraction	Significant factors (p-value)	Parameters (estimates)
Present value of cod/criterion	O(0.004), H(0.000)	O(36.0), H(32.9)
Present value of capelin/criterion	O(0.004), H(0.000)	O(-35.8), H(-34.2)
Cost of unemployment/criterion	S(0.037), H(0.034)	S(-1.40), H(-1.52)

S=simulation, O=optimization, H=high initial stocks. Significance level is 5 prosent.

4.2. Decision strategies

To investigate decision strategies, we turn to a regression analysis of the time-series data. We propose a simple model for each of the resources, explaining quotas as a function of last year's quota and the two stock levels. Thus in addition to the two determinants of yearly quotas in the optimization tool (cod and capelin stocks) we allow for a time lag or filtering. According to Moxnes (1998d), filtering can be used to improve decisions when there is measurement error. To separate the effects of the time lag and of the "indicated quota" from current stock estimates, we report the results according to the following equation

$$K_{t+1} = fK_t + (1-f)(b_{own}R_{own,t} + b_{other}R_{other,t} + b_0) \quad (7)$$

where K_t represents the quota in year t, and R_t is the stock estimate in year t. f is the weight on last year's quota and the right hand parenthesis represents the indicated quota. Table 4 shows average parameter values for all 64 subjects with corresponding t-ratios for the averages.

Table 4: Estimated strategy parameters for all subjects.

Species	Measure	f	b_{own}	b_{other}	b_0
Capelin	Average	0.23	0.31	-0.06	-42
	t-ratio average	8.1	13.8	-1.0	-0.5
Cod	Average	0.30	0.33	-0.010	-127
	t-ratio average	10.0	14.0	-3.6	-5.5

The average weight on last year's quota f is highly significant for both species (even the t-ratios for the individual regressions are high, averages are respectively 2.2 and 3.1 for capelin and cod). Hence we find evidence of a certain smoothing of the quotas, however with the largest weight on the indicated quota from current stock estimates. Also the average weights on the stock level for the own species b_{own} are highly significant (average t-ratios for the individual regressions are respectively 7.9 and 5.4 for capelin and cod, i.e. for the original coefficient for $(1-f)b_{own}$). For capelin there is no significant effect of cod, nor is the constant significant. For cod, there is a significant negative effect of capelin when looking at the average over subjects (the average individual t-ratio is 1.4). Similarly there is a significant negative average constant (the average individual t-ratio is 1.9).

We note that both quotas follow rules which have much lower slopes (0.31 and 0.33) than the slope of 1.0 implied by the target escapement rule predicted by the optimization model for stock sizes above target escapement[2]. Interestingly, the observed behavior deviates from the target escapement rule in the same direction as predicted by more elaborate optimization models valuing stability and incorporating increasing marginal costs and measurement errors, Moxnes (1998a). We also note that the cod slope is close to a slope estimated from data for the real Barents Sea cod fishery, 0.28, Moxnes (1999). Finally, the effect of capelin is in the direction that is suggested by the optimization model.

To see if there are significant differences in parameters between treatments we perform a full factorial ANOVA for each of the four parameters in Equation 7. Only significant factors are reported in Table 5 together with the parameter estimates for the significant factors.

Table 5: Summary of ANOVA analysis for estimated strategy parameters.

Species	Parameter	Significant factors (p-value)	Parameters (estimates)
Capelin	f	S(0.025), OS(0.005)	S(-0.23), OS(0.38)
	b_{own}	S(0.005), OH(0.01)	S(0.025), OH(0.27)
	b_{other}	-	-
	b_0	-	-
Cod	f	O(0.03), OH(0.02)	O(-0.028), OH(-0.34)
	b_{own}	O(0.001)	O(0.19)
	b_{other}	-	-
	b_0	O(0.02)	O(-130)

S=simulation, O=optimization, H=high initial stocks. Significance level is 5 prosent.

The important effects can be summarized in the following way. For capelin, the weight on the earlier quota is reduced by 100 percent in the treatment combination simulation and no optimization tool. The same weight increases by 65 percent in the case with both tools available. The effect of the capelin stock increases by 87 percent in the treatment combination optimization and high initial stocks. For cod, the weight on the earlier quota is reduced by 113 percent in the treatment combination optimization and high initial stocks. The effect of the cod stock increases by 58 percent with optimization. Finally, the constant decreases by 102 percent with optimization (absolute value increases).

2 If the subjects had followed a strict target escapement rule while the stocks fluctuated around the target, low slopes should be expected because we estimated a linear rather than a non-linear model. However, inspection of the individual data reveal that virtually no subject sets quotas equal to zero when the stocks are below the targets. The predominant pattern is a straight line.

4.3. Discussion

To get an idea of the size of the regression coefficients, see Equation 6, we compare with the average benchmark value$_i$, which is NOK 17.2 billion. The effect of the simulation tool is thus 12 percent. The average effect of the optimization tool, over high and low initial conditions, is 11 percent. When no tool is available, students beat the benchmark by 12 percent when initial stocks are low. When initial stocks are high, students are beaten by 12 percent.

Measured by the effect on X_i, the value of each tool is about NOK 2 billion. Thus the benefits are far greater than the development costs of both decision tools. In this connection, note that the virtual reality was calibrated to represent the actual Barents Sea fisheries. However, as we have argued before, the limited number of decision tools or advises in the experiment, are likely to cause an overestimation of the real benefits of the two tools. By itself, the present experiment does not support this reservation, since we did not find any significant negative interaction term for the two tools. However, we still expect the reservation to hold because students are likely to be more receptive to analytical advice than real decision makers, we expect that certain other advises carry more weight than the tested tools (e.g. lobbying), and because our decision tools are likely to be more correct than what tools are in general. (Recall that we knew the virtual reality perfectly before building models for decision support, and we could use time-series data from the virtual reality that were not corrupted in any way.)

What is it that makes the tools useful? Clearly the optimization tool is most useful when initial stocks are high. In this case the economically most important stock, the one for cod starts out 60 to 80 percent above its target level. The optimization tool gives a clear advice: reduce the cod stock. In case of low initial cod stocks, which are close to the target level for cod, there is not the same need for the advice from the optimization tool. This is demonstrated by the situation with none of the tools available. Then subjects do better in the case with low initial stocks than in the case with high initial stocks. Thus, there seems to be an element of luck involved, and this element of luck cannot be confined to only the initial year of the 25 virtual years. With no access to the optimization tool, we suspect that the initial stock level serves as an anchor for the assumed target level, i.e. the students follow a rule of thumb strategy to keep the stocks more or less constant.

Is there any evidence that the students actually used such a rule of thumb. To test this, we analyzed whether the average stocks after 10-15 years, or alternatively 20-25 years, were influenced by the initial stock. Following an optimal strategy, initial transients should not be observed at all after 10 years time. In accordance with this, we find that there is no effect of initial stocks on later stocks of capelin. However for the most important and less volatile species, cod, the students who got high initial stocks, kept a significantly higher cod stock both at 10-15 years and at 20-25 years of management than those with low initial stocks. This supports the hypothesis that they included initial stocks as an element in their rules of thumb. We also found that those with access to the optimization model kept a significantly lower stock after both 10-15 year and 20-25 year. This supports our explanation of why the optimization model had an impact. Access to the simulation model had no significant impact on the stock level over time.

Also note from Table 3 that access to the optimization tool leads to a considerable increase in the total value coming from the cod fishery. The effect is independent of whether initial stocks are high or low. The observation makes sense in light of the high target escapement for capelin, i.e. the optimization tool's consideration of the value of capelin as a food source for cod.

Looking at the analysis of quota strategies, the most obvious tendency is that access to the optimization tool leads to steeper quota strategies as functions of own resource estimates (the only exception is for capelin in the case of low initial stocks). Hence, subject strategies were influenced by access to the optimization tool. Whether this effect on strategies lead to better or worse decisions, we cannot say for sure. A more advanced non-linear, aggregate optimization model with measurement errors, Moxnes (1998a), predicts strategies with slopes considerably below 1.0. Access to the optimization tool might very well have biased decisions away from such a superior strategy.

Also the simulation model gives a significant contribution to the total performance. Why is the simulation model beneficial? The above results indicate that the simulation tool does not help to find a proper target for the cod stock. There seems to be a certain tendency for the simulation tool to reduce the costs of unemployment. Table 3 shows that the simulation tool decreases the costs of unemployment from 3.2 to 2.8 percent of the total criterion value. The reduction of 1.4 percentage points corresponds to NOK 240 million, and is only 11 percent of the estimated benefit of the simulation tool. The simulation tool had no effect on the fraction of the criterion value coming from the cod fishery. Neither has the simulation tool an important effect on the harvesting strategies. In the case of capelin, access to the simulation tool primarily implied that the lag coefficient became sensitive to the availability of the optimization tool.

One might suspect that the inclusion of cohorts in the simulation tool is what makes it most valuable, see Spulber (1985) and Mendelssohn (1978). The aggregation over cohorts in the optimization model implies that once the aggregate stock level is given the quota is given. Thus, aggregate models do not distinguish between a situation with mostly old fish and a situation with mostly young fish. However, in a model with a fixed gear selectivity, as in the virtual reality, there is a limited potential to benefit from the extra information derived from cohorts, Moxnes (1999). For example, if one wants to harvest strongly from a population with high average age, the higher fishing pressures will also affect juveniles, and consequently future harvesting possibilities.

It is perhaps not surprising that it is harder to explain the effect of the simulation tool than of the optimization tool. The optimization tool presents a clear-cut strategy, while the simulation tool presents forecasts based on two somewhat arbitrary strategies. Thus, while the given optimization strategy simply can be given weights by the subjects, the forecasts could be used in different ways depending on how the subjects perceive and formulate the problem.

4.4. The post-questionnaire

The subjects were asked about their willingness to pay (WTP) for having the tools available in case they were to repeat the experiment for another fishing area. On average the WTP for the optimization tool was NOK 53 and NOK 58 for the simulation tool. The difference is not significant. The WTP measures for the two tools were positively correlated. On average, all those who had one or two tools available in the experiment, had a total WTP for the two tools of 202 percent of the actual value of the two tools as measured by the experiment (significantly higher than 100 percent). Similarly, those who had no tool available had a relative WTP of 312 percent. According to these WTP measures, there is a tendency to overestimate the value of both tools.

The subjects were also asked how useful the experiment would be as a supplement to ordinary education. The average rating was 4.0 on a scale from 1 to 5 (75 percent). When commenting, the subjects pointed out the value of getting practical experience with the tools, of experiencing uncertainty, complexity, and dynamics which are often assumed away in education, and the value of experiencing the need for strategy.

When asked to what extent they tried to smooth fisheries from year to year, the average rating was 3.2 (57 percent), with no significant difference between the tools. When asked to what extent they tried to stabilize the resources at the level of the initial year, the average rating was 2.5 (38 percent). With the optimization tool available, the average was 2.4 compared to 2.7 when it was not. The difference is not significant. Both the tendency to smooth quotas and the weaker tendency to look to initial conditions corroborate our earlier findings.

5. CONCLUSIONS

We have performed a laboratory experiment to investigate the practical usefulness of two decision tools to aid quota setting for cod and capelin. An optimization tool was chosen to reflect economic literature on two-species management under uncertainty, while a simulation tool was used to represent biological single species models used to make forecasts. In total 64 students were asked to manage a virtual fishery with or without access to the tools.

The tools turned out to have approximately the same positive effect on management, but the models were useful for different reasons. The optimization tool helped the subjects identify appropriate target stocks. When the optimization tool was lacking, subjects tended to equate the target with historical stocks. The simulation tool had a slight stabilizing effect and it might have had a positive impact because of its richer dynamic structure than the optimization tool. The effect of each tool was an increase of 11 to 12 percent in net present values, while the effect of the tools combined was 23 percent. There was no significant interaction.

For the particular laboratory setting, we conclude that the two tools are not substitutes as a narrow methodological focus might imply. Rather the tools appear to be complements. Moreover, the tools have moderate rather than crucial impacts. This might come as a surprise, at least for the participants in the experiment who overestimated strongly the value of the tools.

Can we generalize from the laboratory results? Or perhaps better, are there findings or tendencies in the experiment that could be expected also in the real management of the Barents Sea fisheries and in other areas of social planning?

First, as found in previous studies, the benefits of tools are likely to depend on the complexity of tasks and the quality of tools. Hence the experiment is of little value with respect to predicting the value of tools in general. Nor can the experiment be used to make general conclusions about simulation versus optimization and economics versus biology.

Second, it might even be problematic to generalize from the experiment to the actual management of cod and capelin in the Barents Sea. If real managers have a better intuitive grasp of the management problem than students, the potential for the tools is reduced. If real managers are pushed by interest groups, while being uncertain about their own intuitive strategies, the tools could have a greater potential in reality than in the laboratory.

Third, it seems likely that for practical purposes, tools tend to be complements rather than substitutes. This should be expected if the tools attack different sub-problems. Complementarity could also follow from differences among decision makers, for whom it may matter how a story is told and who tells it. For instance there could be differences between decision makers with varying educational backgrounds, e.g. see the diffusion literature on "structural equivalence", see e.g. Harkola and Greve (1995).

Fourth, decision makers are not likely to follow advises closely, even when there is only one advice. For instance, most of those who received the optimization tool only, were far from using an exact target escapement policy. In the experiment, adjustments tended to improve the results. This might not always be the case. Hence one should be careful in inferring practical usefulness of a model from its theoretical properties.

Fifth, decision makers are not likely to compensate fully for weaknesses of decision tools. None of the tools, when used alone, produced a larger average improvement then 12 percent above the benchmark. A guessed at strategy, based on the results of a more advanced analysis, produced an average improvement of 30 percent. Thus the global solution is likely to be at least 30 percent above the benchmark. The insufficient adjustment for weaknesses of the tools is not surprising, considering the complexity of the task, complexities that are assumed away in many of the decision tools currently in use. By itself, this finding indicates that one should try to find optimal or near-optimal solutions to more complete models than our simulation and optimization tools. However, this conclusion is only justified to the extent decision makers are willing to accept results from more complex and possibly less tractable analyses. In this regard, there might be a greater acceptance for complex models in a rather well defined problem like ours, than in problems where new understanding is needed to make fundamental changes in policy making, for instance introducing a quota system in the first place.

Sixth, with tools that do not explicitly identify desired targets, there might be a tendency to set targets based on historical values. Concerning the historical management of Barents Sea cod, we note that the implicit fishing strategy after a quota regime came in place in the early 1980s, did not differ much from the implicit fishing strategy based on data for the period before quotas were implemented. The biologists' yield-per-recruit analyses and economic analyses have suggested a significantly lower fishing pressure for cod than what has been realized. There may of course be many reasons for this discrepancy. However, a general tendency to equate targets with historical observations, could explain why the discrepancy has not lead to stronger or quicker adjustments.

Seventh, it is not easy to evaluate the benefits to be reaped by having access to one or more decison tools. There was a clear tendency in a post-questionnaire to produce biased estimates among the students. Those who were availed with at least one of the tools, overestimated the tool's value by on average 200 percent. Those who did not have access to any of the tools, overestimated their value by more than 300 percent. While experience seems to pull in the right direction, the error is still considerable. While one should be careful in generalizing the tendency towards overestimation across tools and decision makers, accuracy should not be expected for complex problems.

Further research is needed to see if these tendencies carry over to other problems of social planning or management. Since the students gave the experiment a high score as a learning experience, the net costs of carrying out such experiments could be reduced by developing and using experiments for both research and education.

REFERENCES

Bakken, B.E. (1993). "Learning and Transfer of Understanding in Dynamic Decision Environments." Ph.D. Dissertation. Cambridge: MIT Sloan School of Management.

Brekke, K.A. (1994). "Optimal Quotas in a Stochastic Two-Species Fishery Model of the Barents Sea." Unpublished working paper. Statistics Norway.

Cavaleri, S. and Sterman, J. (1997). "Towards evaluation of systems thinking interventions: a case study." *System Dynamics Review* **13**:171-186.

Harkola, J. and Greve, A. (1995). "Diffusion of technology: Cohesion or structural equivalence." *Academy of Management Best Papers Proceedings* :422-426.

Mendelssohn, R. (1978). "Optimal harvesting Strategies for Stochastic Single-Species, Multiage Class Models." *Mathematical Biosciences* **41**:159-174.

Mendelssohn, R. (1980). "Managing Stochastic Multispecies Models." *Mathematical Biosciences* **49**:249-261.

Moxnes, E. (1992). "Multispecies Management under Uncertainty." SNF-report 95/92. Bergen: SNF.

Moxnes, E. (1998a). "Catch policy for a predator-prey system: measurement error and value of accuracy." SNF Report 56/98. Bergen: SNF.

Moxnes, E. (1998b). "Not only the tragedy of the commons, misperceptions of bio-economics." *Management Science* **44**(9):1234-1248.

Moxnes, E. (1998c). "Overexploitation of renewable resources: The role of misperceptions." *Journal of Economic Behavior and Organization* **37**(1):107-127.

Moxnes, E. (1998d). "Uncertain measurements of renewable resources: Harvest policy and value of accuracy." SNF Report 51/98. Bergen: SNF.

Moxnes, E. (1999). "Near-to-optimal harvesting strategies for a stochastic multicohort fishery." Report No. 57/99. Bergen: SNF.

Moxnes, E. and Nyhus, H. (1994). "Multispecies management under uncertainty: Optimization and/or simulation." SNF-Report 97/94. Bergen: SNF.

Oz, E., Fedorowicz, J., and Stapleton, T. (1993). "Improving quality, speed and confidence in decision-making - measuring expert systems' benefits." *Information and Management* **24**(2):71-82.

Spulber, D.F. (1985). "The Multicohort Fishery Under Uncertainty." *Marine Resource Economics* **1**(3):265-282.

Tjelmeland, S. (1990). "MULTSIMP: A simplified multispecies model for the Barents Sea." . Bergen: Institute of Marine Research.

Verstegen, J.A.A.M., Huirne, R.B.M., Dijkhuizen, A.A., and Kleijnen, J.P.C. (1995). "Economic value of management information systems in agriculture: A review of evaluation approaches." *Computers and Electronics in Agriculture* **13**(4):273-288.

Webby, R. and O'Connor, M. (1994). "The effectiveness of decision-support systems - the implications of task complexity and DSS sophistication." *Journal of Information Technology* **9**(1):19-28.