A Bayesian network for investigating the decline in fish catch in Switzerland

Mark E. Borsuk, Patricia Burkhardt-Holm, and Peter Reichert
Swiss Federal Institute for Environmental Science and Technology (EAWAG)
8600 Dübendorf, Switzerland
(mark.borsuk@eawag.ch)

Abstract: Catches of brown trout have decreased about 50% in many rivers and streams in Switzerland in the past 15 years. Additionally, the health status of numerous brown trout populations has been assessed to be impaired. In order to evaluate the causes for these phenomena, a nationwide interdisciplinary project named “Fischnetz” was launched in 1999. Twelve hypotheses for the fish population declines were proposed and laboratory and field research projects were initiated to investigate these suggested causes. To apply the results of these investigations to the task of discerning the relative causal importance of each of the hypotheses, a Bayesian probability network is being developed. The development of a “Bayes net” begins with eliciting mental models about the cause-and-effect relationships among system variables from subject-matter experts. Represented as a graphical network, these models imply a set of assumptions about the conditional dependencies among the variables, which simplifies the problem of working with imprecise knowledge. Hard-to-derive joint probability distributions are replaced by a set of conditional distributions, which can be characterized using either: (1) experimental investigation, (2) collected field data, (3) process-based models, or (4) elicited expert opinion. Such information, available as a result of the “Fischnetz” research program and from the scientific literature, will be integrated into the network, thus quantitatively summarizing all relevant information. The quantified network will then be used to assess the historical causal importance of anthropogenic changes, as well as predict the effect of proposed management actions. Analyses will be carried out for individual streams using site-specific information as evidence to update less specific prior beliefs. The results can be used form the basis for preliminary management and to prioritize future research projects based on their ability to reduce uncertainty in model-based assessments. In this paper, a first prototype of the network is presented and the methodology for its construction and application is discussed.

Keywords: causal attribution; probabilistic modelling; expert elicitation; decision support

1. INTRODUCTION

Several indications over recent years have suggested that fish populations in many Swiss rivers and streams have experienced serious declines. Annual catch records of anglers indicate a decrease of up to 50 % since the 1980s [Friedl 1999]. These declines are especially apparent in the more anthropogenically impacted midlands and northern regions of Switzerland. Of the most commonly caught fish species, brown trout *(Salmo trutta fario)*, grayling *(Thymallus thymallus)*, and nase (*Chondrostoma nasus*) are those with the greatest declines [Frick et al. 1998]. In parallel with the indications of decreasing fish catch, fish health monitoring since the 1980s has produced evidence of an impaired health status of native species. Brown trout with both macroscopic visible lesions and histopathological tissue alterations have been observed in a number of Swiss streams [Wahli et al. 1998, Bernet et al. 2000]. The causes of the widespread health problems and decreased abundance of fish in Swiss rivers are not readily apparent.

In January 1998, representatives of the cantonal fisheries administrations, the Swiss federal environmental administration, and several research institutions, met to discuss the observed problems in Swiss fish populations. As a result of this meeting, a nationwide research network named “*Fischnetz” ("Netzwerk Fischrückgang Schweiz”)* was initiated [Burkhardt-Holm et al. 2002].

Among the goals of the Fischnetz project are the following:

- to collect and evaluate available, but scattered, data on the status of Swiss rivers and on fish catches, fish health, and fish populations,
- to improve communication and coordination of relevant individual research activities in various Swiss universities, research institutions, and cantonal and private organizations,
- to initiate new research activities wherever significant gaps in information are identified, and
- to use the resulting research findings to identify the most important causes of the present situation and consider opportunities for improvement.

To achieve these integrative aims, a method is required to combine quantitative data and qualitative knowledge into a coherent analytical framework. We have found causal Bayesian networks [Pearl 1988] to be one of the most promising methods for performing such types of ecological assessment and prediction [Borsuk et al. 2002]. These models focus attention on cause-and-effect relationships of direct scientific or policy relevance and then represent the effects of remaining influences with probabilistic expressions. In this paper, we describe Bayesian networks and their relevance for the Fischnetz project. A preliminary causal model is then presented based on twelve working hypotheses currently being used to organize the Fischnetz research. This network is then expanded using results from the published literature and causal mental models elicited from Fischnetz project coordinators. The assumptions underlying this network allow the complex chain linking anthropogenic causes to ecological effects to be factored into an articulated sequence of conditional relationships. Each of these relationships can then be quantified independently using an approach suitable for the type and scale of information available. Although quantification has not yet occurred, we describe anticipated methods, results, and suggested uses of the final integrated model.

2. CAUSAL BAYESIAN NETWORKS

Fundamental to developing and using Bayesian networks is viewing the model as a graph. In the graph, rounded nodes represent important system variables, and an arrow from one node to another indicates a dependent relationship between the corresponding variables. Such networks can be easily drawn using conventional scientific notions of cause-and-effect. The interesting point that is made explicit in the graph is the conditional independence implied by the *absence* of connecting arrows. These independencies allow each relationship indicated by the *presence* of an arrow to be quantified separately, perhaps based on disparate forms of information [Reckhow 1999]. Quantification of these relationships consists of parameterizing conditional probability distributions that reflect the aggregate response of each variable to changes in its "up-arrow" predecessor, together with the uncertainty in that response.

Conditional probability relationships may be based on either: (1) experimental investigation, (2) collected field data, (3) process-based models, or (4) elicited expert judgment. Observational field data that consist of precise measurements of the variable or relationship of interest is likely to be the most useful and least controversial, form of information. Unfortunately, appropriate and sufficient data may not always exist. Experimental evidence may fill this gap, but concerns may arise regarding the applicability of this information to the natural, uncontrolled system, and appropriate experimental data may also be limited. As a consequence, the elicited judgment of scientific experts may be required to quantify some of the probabilistic relationships. Established techniques exist for performing these elicitations [e.g. Morgan and Henrion 1990, Meyer and Booker 1991], and help to assure accurate and honest assessments.

Once all relationships in a network are quantified, probabilistic predictions of model endpoints can be generated conditional on certain values for "up-arrow" causal variables. These predicted endpoint probabilities, and the relative change in probabilities between alternative scenarios, convey the magnitude of expected system response to historical changes or proposed management while accounting for predictive uncertainties. In addition to prediction, probability networks can be used to perform probabilistic inference when observations of certain model variables are made. Inference is the process of probabilistically estimating the value of all other variables (or distributional parameters) in the network given the values for the observed variables. Inference is useful for updating probabilities based on new observation and for assessing the likely cause of an observed event (fish declines, for example) when data on causal variables is not available. These tasks are particularly valuable when additional monitoring is likely to occur concurrent with the management effort. Statistical inference involves the use of Bayes theorem, thus the term *Bayesian* network