Available online at www.sciencedirect.com
SCIENCE DIRECT®

Environmental Modelling & Software 19 (2004) 423–433

Environmental Modelling & Software

www.elsevier.com/locate/envsoft

A knowledge-based model of watershed assessment for sediment

J.J. Dai [a,*], S. Lorenzato [b], D.M. Rocke [a]

[a] Center for Image Processing and Integrated Computing (CIPIC), University of California, Davis, CA 95616-8553, USA
[b] California Department of Water Resources, DPLA, P.O. Box 948236, Sacramento, CA 94236-0001, USA

Received 1 May 2002; received in revised form 17 July 2002; accepted 19 June 2003

Abstract

A watershed is a complex ecosystem. Assessment of watershed condition entails consideration of numerous issues and factors. The problem is complex, the issues are not well defined, and data are often lacking. These characteristics suggest that a knowledge-based approximate reasoning approach is especially useful for watershed assessment. This paper describes a knowledge base for watershed assessment for sediment (WAS). The knowledge base is designed for protection of fish habitat and control of excessive sediment, and is evaluated in the Ecosystem Management Decision Support (EMDS) system. The WAS model allows experts from diverse fields to contribute to an integrated assessment of watershed condition. As a decision support tool, the model provides a means to assemble key pieces of information and reasoning that support land use or regulatory decisions, and to communicate among diverse audiences the basis for those decisions. The paper also presents an application of the model to assess the condition of a coastal watershed in northern California.

Keywords: Water pollution; Environmental modeling; Decision support; GIS; California

Software availability
Name of software: WAS-KB
Developer: J.J. Dai, University of California at Davis
E-mail: jjdai@ucdavis.edu
Year first available: December 2001
Software required: EMDS and ArcView GIS
Program language: NetWeaver
Program size: ~80 kb
Availability: by contacting the developer

1. Introduction

Water pollution is one of the most frequent and widespread environmental problems. In the United States the Clean Water Act and its related regulations require states to identify impaired water bodies and to develop Total Maximum Daily Loads (TMDL) for the pollutants. A TMDL is a quantitative assessment of water quality problems, contributing sources, and control actions needed to attain water quality protection goals. In recent years focus on TMDLs has shifted from point sources (end-of-pipe discharges) to nonpoint sources (diffuse sources such as run-off) (Ruffolo, 1999). Examples of nonpoint source pollution (NPS) include excessive sediments, nutrients, and chemicals from the watershed. NPS pollution is often strongly influenced by human activities in the watershed, as well as the physical characteristics of the watershed (Lisle, 1989; Waters, 1995). For example, accumulative changes in land uses and land cover in the upland and riparian area could have a profound impact on the quality of waters in the watershed. Because water quality depends on the health of the watershed, it is important to assess and monitor watershed conditions in order to control water quality problems.

Assessment of watershed condition is a complex task (NCRWQCB, 1994; NCWAP, 2001). It deals with issues at different spatial scales (e.g. stream reach, subwatershed, and watershed). It requires consideration of numerous watershed functions, anthropogenic influences, and management concerns. It involves characterizing aquatic, riparian, terrestrial features and manage-

* Corresponding author. Tel.: +1-530-752-6501; fax: +1-530-752-8894.
E-mail address: jjdai@ucdavis.edu (J.J. Dai).

1364-8152/$ - see front matter
doi:10.1016/S1364-8152(03)00158-0

ment issues within a watershed and developing linkages between watershed processes and environmental concerns. Watershed assessment must also be relevant to management decisions related to pollution control, land use practices, and natural resource stewardship. In the management arena, decisions must be well supported and easily communicated to various constituencies. Thus, watershed assessment needs an approach that can handle complex problems but is easy to implement, that is flexible but consistent, that can be applied at different spatial scales, and that can be readily translated into easily communicated descriptions related to management decisions.

One approach to ecological assessment is knowledge-based decision support systems (Reynolds et al., 1996; Schmoldt and Rauscher, 1996). A knowledge base is an organized body of knowledge that provides a formal logical specification for the interpretation of information (Walters and Nielsen, 1988). In this paper, we develop a knowledge base to support sediment TMDL and related watershed assessment and monitoring. The knowledge base is designed in the NetWeaver knowledge base development system and evaluated in the Ecosystem Management Decision Support (EMDS) system developed by the US Forest Service (Reynolds, 1999). The result is the Watershed Assessment for Sediment (WAS) model. We also present an application of the WAS model to evaluation of the condition of a coastal watershed in northern California.

The rest of the paper is organized as follows. Section 2 describes the knowledge-based approximate reasoning methodology and the EMDS system. The WAS knowledge base is developed in Section 3, which defines the problem, describes the sources of knowledge acquisition and the criteria used for assessment, and illustrates the structure of the knowledge base. An application of the model to assess the condition of the Noyo River watershed in northern California is presented in Section 4. Finally, conclusions and discussion are provided in Section 5.

2. Methods

The WAS knowledge base was developed and evaluated using the EMDS system. EMDS is a general modeling tool that integrates knowledge-based reasoning into a GIS environment to provide decision support for ecological landscape assessment. EMDS consists of three major components: a knowledge base development tool (NetWeaver), a GIS application framework (ArcView), and an assessment system (Reynolds, 1999). NetWeaver is composed of a fuzzy-logic-based reasoning engine and a graphic user interface for knowledge base developers. The ArcView application provides database management, spatial analysis, system interface, and map display. The assessment system allows the user to evaluate the knowledge base for a specific spatial database and view the results. The methodology underlying EMDS is knowledge-based approximate reasoning. The knowledge base contains knowledge and experience for the subject domain (domain knowledge) and specifies the logical relations among topics of interest to an assessment. The inference engine performs knowledge-based approximate reasoning to draw conclusions about the state of the system.

The NetWeaver inference engine is based on the concept of approximate reasoning with fuzzy logic. Fuzzy set theory (Zadeh, 1965, 1968, 1979) resembles human reasoning in its use of approximate information and uncertainty to generate decisions. The theory is designed to mathematically represent uncertainty and vagueness and provide formalized tools for dealing with imperfect information and knowledge intrinsic to many problems. Fuzzy logic is concerned with quantification of set membership and associated set operations (Kaufmann, 1975). It is a departure from classical Boolean logic and provides a metric that expresses partial membership in a set. Because fuzzy logic can handle approximate information in a systematic way, it is a useful tool for modeling complex, abstract topics such as habitat suitability and watershed condition that depend on numerous, diverse subordinate conditions and where imprecise information is common. In multicriteria assessment, the fuzzy logic model provides a means to combine the scores of the individual factors into an overall ranking. It is more flexible and robust than the traditional ranking methods such as ordinal combination (McHarg, 1969), weighted linear combination (Banai-Kashani, 1989; Pereira and Duckstein, 1993) and Boolean algebra (Hall et al., 1992) (see reviews in Jiang and Eastman, 2000; Stoms et al., 2002).

Fuzzy logic has been applied to a variety of problems in environmental sciences and management, including ecosystem modeling (Salski and Sperlbaum, 1991), environmental assessment (Booty et al., 2001; Holland, 1994; Smith, 1997), land suitability assessment (Davidson et al., 1994; Kollias and Kalivas, 1998; Stoms et al., 2002; Van Ranst et al., 1996), natural hazard analysis (Chen et al., 2001), natural resources management (Mays et al., 1997; McBratney and Odeh, 1997), watershed assessment (Reynolds et al., 1996, 2000) and water pollution modeling (Kuncheva et al., 2000).

A NetWeaver knowledge base consists of dependency networks, data links, and logic operators (Stone et al., 1986). A dependency network is a hierarchical network that evaluates a proposition. The result of an evaluation is a score (the truth value) that expresses the degree to which the proposition is supported or refuted. The networks are recursive as a network may be evaluated by other networks. Data links are elementary networks used to read and evaluate data. A data link evaluates a proposition by comparing the value of an observation to an argument that defines the reference conditions. Calculated data links evaluate an expression, which may consist of mathematical operators, data links or calculated data links. Logic operators are used to specify the logical dependency of a network on its antecedents. The two most commonly used logic operators are “OR” and “AND”. The OR operator takes the maximum score of the antecedent networks or data links. The AND operator is similar to a minimum operator but adjusts for the scores of all antecedents to reduce the influence of missing data. Detailed information about NetWeaver and EMDS is available online from the EMDS website at http://www.fsl.orst.edu/emds/. Examples of dependency networks, data links, logic operators and their uses are given in Section 3.4.

3. The knowledge base

In the knowledge-based modeling approach, watershed assessment is a multicriteria evaluation in which knowledge of the experts is used to define the factors characterizing the watershed and the logic relations between the factors. The knowledge base encapsulates the assessment criteria and the relationships in an explicit form so that they can be easily examined, modified, or updated. This section describes the WAS knowledge base, including its objective and scope, the sources of knowledge acquisition, the assessment criteria, and the structure of the knowledge base.

3.1. Problem domain

The WAS knowledge base is designed for sediment related watershed assessment. The knowledge base describes watershed conditions that protect the beneficial uses of cold-water bodies such as cold-water fishery, migration, spawning, reproduction, and early development of cold-water fish. Native salmonids in Northern California and the Pacific Northwest are used as the index species. We define “healthy” watershed conditions to be those that are suitable for sustaining healthy populations of the native salmonids.

3.2. Knowledge acquisition

Knowledge about the problem domain was acquired primarily from the following sources:

- Review of the literature on sediment TMDL, water quality assessment, aquatic (especially salmonids) habitat protection, watershed protection and restoration. Special attention was given to information related to sediment TMDLs and Northern California. Major references are cited in Section 3.3. Reports of sediment TMDL for coastal watersheds in California by the US Environmental Protection Agency are available online at http://www.epa.gov/region09/water/tmdl/.
- Discussions with the staff of the California State Water Resources Control Board and the North Coast Regional Water Quality Control Board. Cooperation with the staff of the North Coast Watershed Assessment Program, which is a California state interagency program consisting of specialists from a number of disciplines including fishery biologists, foresters, geologists, hydrologists, water quality control engineers, water resources specialists, and GIS analysts. The program provided a forum for discussion and exchange of ideas on watershed assessment and modeling.

3.3. Assessment criteria

To assess watershed condition, we use indicators that represent characteristics of the watershed and sources of accumulative effects. These indicators are used with reference conditions as assessment criteria in the knowledge base. We consider two general categories of indicators: (1) those related to stream and salmonid habitat condition, and (2) those related to upland condition (Dai and Rocke, 2001; Mangelsdorf and Lundborg, 1997). We use six groups of indicators to assess stream condition: instream habitat, water quality, channel condition, riparian functioning, stream flow and fish passage. We employ four groups of indicators to evaluate upland condition: roads, land use, land cover, and slope instability. The indicators are briefly described below. Details of the assessment criteria are discussed in Rocke and Dai (2002). Definition of the terms is given in the TMDL reports available online at the USEPA website (see Section 3.2).

3.3.1. Stream indicators

Each of the six groups of stream indicators provides a suite of measures that reflect important characteristics of salmonid fitness and survival or that are related to sedimentation. The first group describes instream habitat condition and consists of three sub-groups of factors: pool habitat, substrate composition, and habitat complexity. The subgroup “substrate” is used to reflect the impact of fine sediment and substrate condition on salmonid life stages, especially spawning, embryo development, and fry emergence (Burns, 1970; Chapman, 1988; Flosi et al., 1998; Knopp, 1993; Lisle and Lewis, 1992; Lisle and Hilton, 1992; Tappel and Bjornn, 1983). It consists of indicators important to the spawning and early development of salmonids. For example, percent fines in sediment samples measure the abundance of fine sediment in the stream bottom. Embeddedness measures the degree to which the surface particles in the substrate are cemented. The subgroups “pool habitat” and “habitat complexity” consist of other important factors that limit the success of the native salmonids. Those factors include pool depth and frequency, large woody debris (LWD), instream cover, and off-channel habitat (Beechie and Sibley, 1997; Bilby and Ward, 1989; Fetherston et al., 1995; Flosi et al., 1998; NMFS, 1997).

The second group “channel stability” measures channel stability/complexity and is composed of three indicators: bank stability, width to depth ratio, and Thalweg variability (NMFS, 1997; Rosgen, 1996). The third group “water quality” consists of water quality indicators including water temperature, dissolved oxygen, suspended sediment, and water turbidity (NCRWQCB, 1994; Newcombe and MacDonald, 1991; NMFS, 1997). The fourth group “riparian vegetation” is concerned with riparian function, which provides shades and sources of LWD, and has two indicators, canopy density and the potential to provide LWD (Flosi et al., 1998; Keithley, 2000; NMFS, 1997). The last two groups of indicators consist of “stream flow” and “passage barrier”. Adequate stream flows are very important to the survival of fish. The indicators for flow condition are summer low flow and winter channel maintenance flow (Leland, 2001; Klamt, 2001a). The indicator of passage barrier describes the accessibility of the streams to anadromous fish (Mangelsdorf and Lundborg, 1997).

3.3.2. Upland indicators

The upland environment produces cumulative watershed effects and thus is an important part of watershed assessment. To reduce excessive sediment into the streams, it is necessary to control the sources of sediment. The upland sediment sources are primarily surface erosion, gully erosion, and mass wasting (e.g. landslides) (Dietrich et al., 1998; Mangelsdorf and Lundborg, 1997). There are two classes of factors in sediment production and delivery in the region (NCRWQCB, 1994; NCWAP, 2001):

- Natural factors such as properties of the bedrock, soil composition (depth, permeability, cohesion, and structure), slope steepness, rainfall intensity and duration, ground water levels;
- Human factors such as vegetation removal (e.g. timber harvest, livestock) and surface disturbances (e.g. road construction and drainage, urbanization).

We are primarily concerned with controllable or management-related sources of sediment, which are associated with human activities and will likely respond to land management and/or restoration measures. The upland indicators are selected to address issues on upland activities that are related to production or delivery of sediment to the watercourse. Four groups of upland indicators are used. Road-related run-off and sediment production are the dominant controllable sources of sediment in the coastal watersheds. The first group “road” consists of four road-related indicators: road density by road type, stream crossings by roads, road proximity to streams, and road use intensity (Cederholm et al., 1981; Flanagan et al., 1998; Leland, 2001; NCWAP, 2001; NMFS, 1997; Weaver and Hagans, 1994). The second group “land use” addresses other major issues on land use disturbance such as timber harvest, farming, ranching and urban land use (Lewis, 1998; NCWAP, 2001). The third group “land cover” deals with some of the issues related to land cover and consists of tow indicators: canopy density and forest seral stages (Keithley, 2000; NCWAP, 2001; White, 1982). The last group “slope instability” is concerned with management-related activities on unstable slopes (Dietrich et al., 1998; Klamt, 2001a; Leland, 2001; PWA, 1998).

3.4. Knowledge base structure

The knowledge base structure is a hierarchy of dependency networks. Each network evaluates a specific proposition about the state of watershed condition. All networks and their propositions in the WAS knowledge base are listed in Table 1. The knowledge base structure is designed to address the issues concerned by the watershed managers and to reflect their opinions on the importance of each issue. At the top of the hierarchy is the network *watershed condition* for the proposition that the overall condition of the watershed is suitable for sustaining healthy populations of the native salmonids. The *watershed condition* network depends on two lower-level networks: *stream condition* and *upland condition* (Fig. 1). Evaluations of stream and upland conditions depend on networks further down the hierarchy. For example, *stream condition* depends on six subordinate conditions: *instream habitat*, *channel stability*, *water quality*, *riparian vegetation*, *stream flow*, and *passage barrier* (Fig. 2). The networks used to evaluate upland condition are *road*, *land use*, *land cover*, and *slope instability*.

The knowledge base uses many features of NetWeaver. Logic operators such as “AND” and “OR” are used to specify the logical dependency of the networks. For a proposition to be true, the AND operator requires all antecedent networks or data links to be true, while the OR operator needs only any one of the premises to be true. For example, *watershed condition* depends on two antecedent networks *stream condition* and *upland condition* through the AND logic operator (Fig. 1). This means that, to be rated as highly suitable, a watershed must score reasonably high in both stream condition and upland condition. Data links are the terminal nodes in the knowledge base and are used to access data in the GIS database. Calculated data links can evaluate mathematical expressions and are used as

Table 1
Networks and propositions for assessment of watershed condition

Network name	Propositions evaluated by network	Sources of comparison
WATERSHED	Overall watershed condition is within a suitable range.	
Stream condition	Stream condition is within a suitable range.	
Instream habitat	Instream habitat condition is within a suitable range.	
Pool habitat	Pool habit condition is within a suitable range.	
Pool depth	Pool depth is within a suitable range.	TMDL targets
Pool frequency	Pool frequency is within a suitable range.	TMDL targets
Substrate	Substrate composition is within a suitable range.	
Percent fines	Percent fines are within suitable ranges.	
Spawning fines	Percent spawning fines are within suitable ranges.	TMDL targets
Emerging fines	Percent emerging fines are within suitable ranges.	TMDL targets
Embeddedness	Substrate embeddedness is within a suitable range.	TMDL targets
Habitat complexity	Habitat complexity is within a suitable range.	
LWD	Amount of large woody debris is within a suitable range.	Reference
Instream cover	Instream cover is within a suitable range.	Reference
Off-channel habitat	Off-channel habitat is within a suitable range.	Reference
Channel stability	Channel condition is within a suitable range.	
Bank stability	Bank stability is within a suitable range.	Reference
Width to depth ratio	Channel’s width to depth ratio is within a suitable range.	Reference
Thalweg variation	Channel’s thalweg variation is within a suitable range.	Reference
Water quality	Water quality condition is within a suitable range.	
Water temperature	Water temperature is within a suitable range.	Regulation
Dissolved oxygen	Dissolved oxygen is within a suitable range.	Regulation
Suspended sediment	Suspended sediment is within a suitable range.	Regulation
Water turbidity	Water turbidity is within a suitable range.	Regulation
Riparian vegetation	Riparian condition is within a suitable range.	
Riparian canopy	Riparian canopy is within a suitable range.	Reference
LWD potential	LWD potential is within a suitable range.	Reference
Stream flow	Stream flow condition is within a suitable range.	
Summer flow	Summer low flows are within a suitable range.	Reference
Winter flow	Winter maintenance flows are within a suitable range.	Reference
Passage barrier	Barriers to fish migration are within a suitable range.	TMDL targets
Urban condition	Upland condition is within a suitable range.	
Road	Road condition is within a suitable range	
Road density	Road density is within a suitable range.	Reference
Road crossings	Road crossings are within a suitable range.	Reference
Road proximity	Road proximity to streams is within a suitable range.	Reference
Road use	Road use intensity is within a suitable range.	Reference
Land use	Land use condition is within a suitable range.	
Timber harvest	Timber harvest disturbance is within an acceptable range.	Reference
Farming	Farming disturbance is within an acceptable range.	Reference
Ranching	Ranching disturbance is within an acceptable range.	Reference
Urban	Urban disturbance is within an acceptable range.	Reference
Land cover	Land cover condition is within a suitable range.	
Canopy	Canopy density is within a suitable range.	Reference
Forest	Forest seral stage is within a suitable range.	Reference
Slope instability	Activities on unstable slopes are within an acceptable range.	Reference

alternative ways to combine scores of individual factors. Dynamic data links are employed to define the fuzzy curves at the running time.

We describe the structure of one dependency network, *instream habitat*, as a typical example of the knowledge base structure. *Instream habitat* is one of the networks used to evaluate stream condition (Fig. 2). The suitability of instream habitat condition depends on three subordinate conditions of *pool habitat*, *substrate* and *habitat complexity* (Fig. 3). Use of the AND operator means that all of the three subordinate conditions must be suitable to the fish for instream habitat condition to be rated suitable. The suitability of pool habitat is evaluated by two data links, *pool depth* and *pool frequency*. A data link is a terminal node in the knowledge base, which reads a data value from the GIS database and compares the value to a fuzzy curve that defines the reference condition. The data links are implemented as dynamic data links so that the reference values are read and can be modified easily at the running time. Almost all data links in the knowledge base are defined as dynamic data links. The AND logic operator is used for evaluation of *pool*

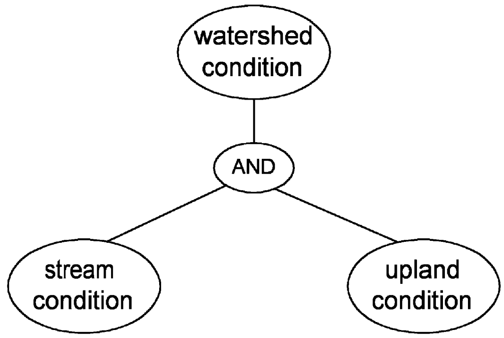

Fig. 1. The network for evaluation of overall watershed condition.

habitat condition because a good salmon habitat requires many deep pools in the streams. The *substrate* network depends on *percent fines* and *embeddedness*. *Embeddedness* is a data link. The *percent fines* network is evaluated by two data links, *spawning fines* (percent fines <0.85 mm) and *emerging fines* (percent fines <6.5 mm). Habitat complexity is evaluated by the calculated data link *habitat complexity calc* as a sum of products on three data links *LWD*, *instream cover*, and *off-channel habitat*. The use of calculated data link and the sum operation assumes that the factors for evaluating habitat complexity are compensatory to each other to some degree.

4. Application

4.1. The Noyo River watershed

The Noyo River watershed is a forested, coastal watershed in Mendocino County, California, which encompasses approximately 430 km^2 (106,168 acres). Like many other rivers in the region, the Noyo River and its tributaries are impacted by elevated sedimentation due to inherent geologic instabilities, past and present land use practices, and other characteristics of the watershed (NCRWQCB, 1994). The Noyo River watershed has been listed as water-quality impaired due to sedimentation on the State's list of impaired waters as required by Section 303(d) of the Clean Water Act. Sedimentation is impacting the cold-water fishery, a beneficial use of the Noyo River watershed, including the migration, spawning, reproduction, and early development of cold-water fish such as coho salmon and steelhead. The US Environmental Protection Agency (USEPA) has established sediment TMDL for the Noyo River (USEPA, 1999). The TMDL proposes a number of numeric targets for fish habitat protection and sediment control, which could be used as the reference conditions for assessment of the watershed condition. The watershed managers have suggested the planning watershed

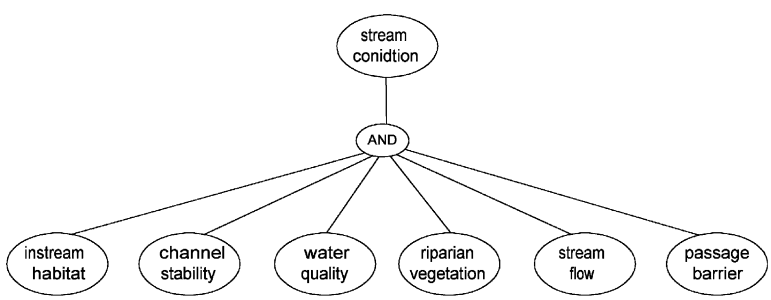

Fig. 2. The network for evaluation of stream condition.

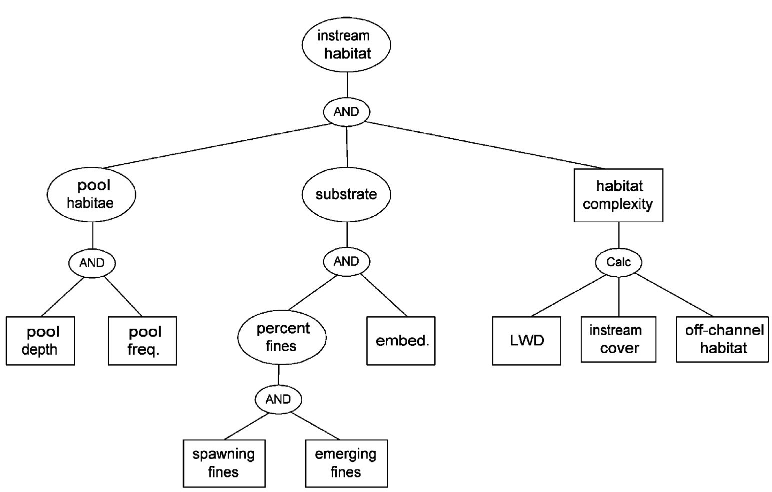

Fig. 3. The network for evaluation of instream habitat condition. Networks are shown as ovals and data links are rectangles.

(a unit in the California's standard watershed map) to be the unit of the watershed assessment. The Noyo River watershed consists of 22 planning watersheds (Fig. 4).

4.2. Data and processing

Data for the application were primarily from the Klamath Resource Information System for the Noyo River Watershed (KRIS Noyo, 2000). KRIS Noyo was developed by the Institute for Fisheries Resources and funded by the California Department of Forestry and Fire Protection and the National Fish and Wildlife Foundation. It contained GIS data and other useful information concerning the factors that affected fish resources and water quality of the watershed. Given the general lack of available data for watershed assessment in the region (Klamt, 2001b), KRIS Noyo provided the best data available to the public. The GIS data were in the form of ArcView shape files, ArcInfo coverages and grid files. We derived a number of data layers for analysis from the original files. The values for reference conditions (criteria) were mainly obtained from TMDL reports (e.g. USEPA, 1999), regional water quality control plans (e.g. NCRWQCB, 1994), and other scientific literature (e.g. NMFS, 1997; NCWAP, 2001).

The original data were in various spatial units. For example, data on land use and land cover were in area units; data on roads and streams were in line units; and some were point data such as road crossings of streams. A GIS provides a powerful tool for processing spatial data and performing spatial operations (Dai and Rocke, 2000). We performed three types of spatial overlays to derive data in the target unit from the original units: point in polygon operation (e.g. count the number of road crossings in a watershed), polygon on polygon operation (e.g. calculate land use densities), and line in polygon operation (e.g. estimate road densities). Most instream habitat data were collected in stream reach surveys and the reach unit was used for the habitat data.

The EMDS software, described in Section 2, was used to develop and run the Noyo application. The software integrates a knowledge base development tool (NetWeaver) and an assessment tool (ArcView extension) in a GIS environment (ArcView). We developed the WAS knowledge base in NetWeaver and the Noyo GIS database in ArcView; and run the application using the assessment tool. The software was easy to learn and to use, and had the ability to handle missing data. It used zero score to indicate that the suitability of a condition could not be determined due to missing data. It allowed the user to run the whole model, or select a subset of the knowledge base for analysis, an option of evaluation with missing data. One might also build the switch nodes into the knowledge base and turn some of them off when data were not available for evaluating some of the networks. In the application, three of the networks, channel stability, stream flow, and passage barrier, could not be evaluated because there were no data on them. We evaluated instream habitat condition at the reach level and aggregated the reach scores into the planning watersheds.

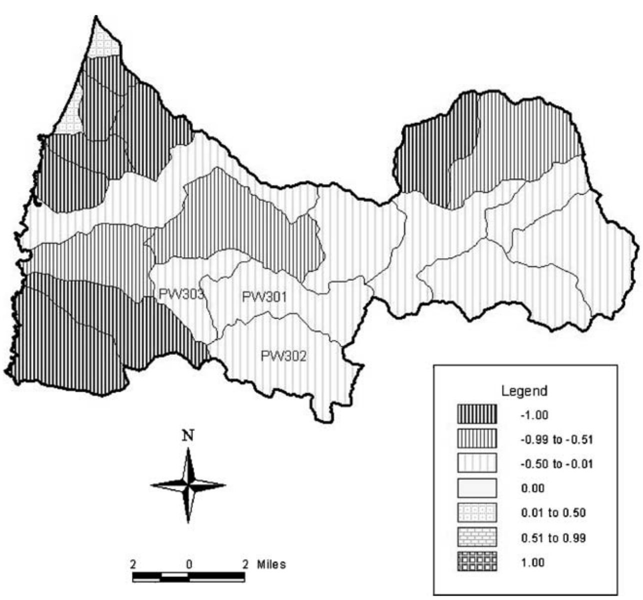

Fig. 4. Map of suitability scores of the Noyo River watershed. The proposition is that the overall condition of a planning watershed is suitable for sustaining healthy populations of native salmonids.

4.3. Results

Outcomes from the model are scores that express the degree to which the data support or refute the propositions about the suitability of various watershed conditions. We call them suitability scores. The scores range from −1 to 1. A value of 1 means a fully suitable condition and −1 implies a totally unsuitable condition. Scores between 0 and 1 indicate a partially suitable condition, and values between 0 and −1 suggest unsuitable conditions. A zero value indicates that the suitability cannot be determined due to missing data. The results of the model should be interpreted qualitatively. To facilitate discussion of the results, we classify the scores into seven suitability levels (Table 2). Results of evalu-

Table 2
Scores and levels of suitability

Score	Level of suitability
−1.00	Fully unsuitable
−0.99 to −0.51	Moderately unsuitable
−0.50 to −0.01	Somewhat unsuitable
0.00	Undetermined
0.01 to 0.50	Somewhat suitable
0.51 to 0.99	Moderately suitable
1.00	Fully suitable

ation from the model suggest that almost all planning watersheds in the Noyo River watershed are impaired to some degree. The mean suitability score for the overall condition is −0.59 with a range from −1.00 (fully unsuitable) to 0.10 (somewhat suitable). There are many variations in the suitability scores. Map of the scores shows the spatial variation in conditions of the planning watersheds (Fig. 4). The network structure of the model and its flow diagrams help to track down the possible sources of the problems.

We illustrate it with selected results on three planning watersheds, PW301, PW302 and PW303 (Part 1 of Table 3). The results suggest that the three planning watersheds are partially impaired, indicated by the negative overall suitability scores (−0.27, −0.45, −0.36). To identify causes of the problems, one simply traces the hierarchical network structure top down. The score on watershed condition depends on the evaluation of two subordinate conditions: stream condition and upland condition (Fig. 1). The scores on them indicate if anything is wrong with those subordinate conditions. For example, PW301 has a negative score on stream condition (−0.36, somewhat unsuitable) but a positive score on upland condition (0.20, somewhat suitable), suggesting that its main problem is in stream condition. PW303, on the contrary, has a negative score on upland condition (−0.44, somewhat unsuitable) and a positive score on stream condition (0.14, somewhat suitable). PW302 has negative scores on stream (−0.49) and upland (−0.14) conditions, suggesting that both are impaired. To identify specific problems in stream and upland, one needs to go down further the hierarchy and examine the antecedents of those two conditions. For example, the network diagram shows that stream condition depends on six subordinate conditions (Fig. 2): instream habitat, channel stability, water quality, riparian vegetation, stream flow, and passage barrier. Inspection of suitability scores on those subordinate conditions will provide specific information on what might be wrong in the streams. For example, one of the stream problems in PW301 and PW302 appears to be impaired instream habitat condition, indicated by their negative scores on instream habitat suitability (Part 1 of Table 3). One can examine the scores all the way down to the data links (the elementary networks) and identify all problems in each and every planning watershed.

Table 3
Suitability scores for selected networks and planning watersheds

WSID	Overall	Stream	Upland	Instream	Road
1. Current condition					
PW301	−0.27	−0.36	0.20	−0.44	0.10
PW302	−0.45	−0.49	−0.14	−0.46	−0.19
PW303	−0.36	0.14	−0.44	0.05	0.27
2. Scenario condition					
PW301	0.09	−0.03	0.48	0.02	0.47
PW302	−0.04	−0.07	0.08	0.01	0.24
PW303	−0.33	0.17	−0.41	0.28	0.05
3. Increase in scores					
PW301	0.36	0.33	0.27	0.46	0.38
PW302	0.41	0.42	0.22	0.47	0.43
PW303	0.04	0.03	0.03	0.23	0.32

As a decision support tool, the model can be further used to assess the potential impact of management plans or strategies on watershed conditions. We illustrate it with the evaluation of a simple management scenario on the three planning watershed. Roads, especially unpaved roads, have been a significant source of fine sediment in the Noyo River basin. Reduction of roads (e.g. closure of certain logging roads) is a management option that can reduce sediment production in the upland and sediment loads in the streams. The scenario assumes to close 40% of roads in the watershed and reduce the fine sediments in the streams by 20%. Results of evaluating the scenario suggest that the proposed action can have direct impact on instream habitat condition and road condition of the three planning watersheds. The suitability scores on those conditions are increased notably, ranging from 0.23 to 0.47 (Part 3 of Table 3), which suggests significant improvement of the conditions. Stream and upland conditions are improved as well although the degree of improvement varies among the planning watersheds. Overall, PW301 and PW302 would benefit more than PW303 under the scenario, indicated by the difference in increase of suitability scores. It is also noted that, even with the improvement, none of the suitability scores (Part 2 of Table 3) are close to 1.0 (fully suitable condition) and some are still in negative values (unsuitable conditions). It reflects the fact that there are many factors affecting watershed condition and this simple scenario addresses only a few of the problems. It suggests that more comprehensive plans and greater efforts are needed to make the watershed condition highly healthy. The model provides a ready means for assisting development of comprehensive plans for watershed protection and restoration.

5. Conclusions and discussions

A watershed is a complex ecosystem. Assessment of watershed condition entails consideration of numerous issues and factors. The problem is complex, the issues are not well defined, and data are often lacking. These characteristics suggest that a knowledge-based approximate reasoning approach is especially useful for watershed evaluation. In this paper, we have developed a knowledge base for watershed assessment of sediment. The WAS knowledge base is designed to evaluate watershed condition and to protect fish habitats, especially those of the native salmonids in the coastal regions of Northern California and the Pacific Northwest. WAS is implemented and evaluated in the EMDS system. The Noyo application illustrates that the WAS model is a useful tool for watershed assessment and for supporting management decisions.

The WAS model has a number of desirable features. It has the ability to address complex and abstract topics such as fish habitat suitability and stream condition. It can deal with watershed issues at different spatial scales from stream reaches to the entire watershed. It provides a means to assemble key pieces of information and reasoning that support sediment related land use or regulatory decisions. It encapsulates the assessment criteria and their logic relations in an explicit form so that they can be easily examined, explained, and modified. The model provides not only a tool for assessment of watershed condition, but also a ready means for watershed managers to depict assessment results and to explain the basis for their decisions.

The model has its limitations. First, it is a qualitative model. Results of the model only indicate the quality of watershed conditions, given the criteria of assessment, the structure of the knowledge base, and the data available. The model cannot and is not intended to make quantitative predictions. For example, the model cannot predict the amount of sediments in the streams. The suitability score only indicates the degree of suitability of the habitat condition to the fish given the data on sediment and other factors. Second, it is a knowledge-based model in which knowledge of the experts is used to define the topics, the assessment criteria, and the structure of the knowledge base. Different experts might have different opinions on the same issues. The knowledge of the experts could be interpreted in different ways. WAS is limited to the knowledge we acquired in this study and our interpretation of the knowledge.

An advantage of knowledge-based modeling is that it allows incremental, evolutionary development of a large and complex model. The WAS knowledge base is evolving and there is a lot to be done to improve it in further studies. Watershed assessment is often context-dependent. The current knowledge base is primarily based on knowledge and experience from the coastal region of Northern California and to some degree the Pacific Northwest of the US Applications of WAS to other regions may require modification of the knowledge base. Even for the region that the model currently applies to, better knowledge and data will improve the model. For instance, the fuzzy curves are defined by reference values that specify suitable and unsuitable conditions. Not all reference values are readily available. We are working with the scientists and watershed managers to develop and refine the reference values. Also better knowledge will help develop better indicators describing the ecosystem under study. Ultimately, it is the user who will have to decide what issues are important to them and what reference values are appropriate to use for their watersheds. This requires the system developer and the user to work closely in modeling and application. Data availability and accuracy are two issues important in applications. We are collecting more data for the Noyo River watershed so that we can test all elements of the model and evaluate the sensitivity of the outputs of the model to data inputs with various accuracy levels.

Acknowledgements

We gratefully acknowledge the assistance of the staff in the California State Water Resources Control Board and the North Coast Regional Water Quality Control Board, especially, David Leland, Bob Klamt, Dave Evans and Linda Rao. We are grateful to the staff of the California North Coast Watershed Assessment Program for discussions and cooperation. We thank the California Department of Forestry and Fire Protection and Institute for Fisheries Resources, especially Chris Keithley and Patrick Higgins, for assistance with the data. We wish to thank the referees for helpful comments and suggestions.

References

Banai-Kashani, R., 1989. A new analytical method for site suitability analysis: the hierarchical process. Environmental Management 13, 685–693.

Beechie, T.J., Sibley, T.H., 1997. Relationships between channel characteristics, woody debris, and fish habitat in northwestern Washington streams. Transactions of the American Fisheries Society 126, 217–229.

Bilby, R., Ward, W., 1989. Changes in characteristics and function of woody debris with increasing size of streams in western Washington. Transactions of the American Fisheries Society 118, 368–378.

Booty, W.G., Lam, D.C.L., Wong, I.W.S., Siconolfi, P., 2001. Design and implementation of an environmental decision support system. Environmental Modelling and Software 16, 453–458.

Burns, J.W., 1970. Spawning bed sedimentation studies in north California streams. California Fish and Game 56, 253–279.

Cederholm, C.J., Reid, L.M., Salo, E.O., 1981. Cumulative effects of logging road sediment on salmonid populations in the Clearwater River, Jefferson County, Washington. In: Proceedings of the Conference on Salmon Spawning Gravel: A Renewable Resource in the Pacific Northwest, Washington State University, Washington Water Research Center, pp. 38–74 Report 39.

Chapman, D.W., 1988. Critical review of variables used to define effects of fines in reds of large salmonids. Transactions of the American Fisheries Society 117, 1–25.

Chen, K., Blong, R., Jacobson, C., 2001. MCE-RISK: integrating multicriteria evaluation and GIS for risk decision-making in natural hazards. Environmental Modelling and Software 16, 387–397.

Dai, J.J., Rocke, D.M., 2000. A GIS-based approach to spatial allocation of area source solvent emissions. Environmental Modelling and Software 15, 293–302.

Dai, J.J., Rocke, D.M., 2001. A Knowledge Base for Watershed Assessment for Sediment (WAS). Technical Report, CIPIC, University of California, Davis, CA.

Davidson, D.A., Theocharopoulos, S.P., Bloksma, R.J., 1994. A land evaluation project in Greece using GIS and based on Boolean and fuzzy set methodologies. International Journal of Geographical Information Systems 8, 369–384.

Dietrich, W.E., Real de Asua, R., Coyle, J., Orr, B., Trso, M., 1998. A validation study of the shallow slope stability model, SHALSTAB. In Forested Lands of Northern California. Prepared for Louisiana-Pacific Corporation, Calpella, CA.

Fetherston, K.L., Naiman, R.J., Bilby, R.E., 1995. Large woody debris, physical process, and riparian forest development in montane river networks of the Pacific Northwest. Geomorphology 13, 133–144.

Flanagan, S.A., Furniss, M.J., Ledwith, T.S., Love, M.A., Moore, K., Ory, J., 1998. Methods for Inventory and Environmental Risk Assessment of Road Drainage Crossings. Water/Road Interaction Technology Series. USDA Forest Service, Technology and Development Program.

Flosi, G., Downie, S., Hopelain, J., Bird, M., Coey, R., Collins, B., 1998. California Salmonid Stream Habitat Restoration Manual. California Department of Fish and Game, Inland Fisheries Division, Sacramento, CA.

Hall, G.B., Wang, F., Subaryono, J., 1992. Comparison of Boolean and fuzzy classification methods in land suitability analysis by using geographical information systems. Environment and Planning A 24, 497–516.

Holland, J.M., 1994. Using fuzzy logic to evaluate environmental threats. Sensors 11, 57–61.

Jiang, H., Eastman, J.R., 2000. Application of fuzzy measures in multicriteria evaluation in GIS. International Journal of Geographical Information Science 14, 173–184.

Kaufmann, A., 1975. Introduction to the theory of fuzzy subsets. Fundamental Theoretical Elements, vol. 1. Academic Press, New York.

Keithley, C., 2000. Evaluating Stream and Watershed Conditions in Northern California. California Department of Fish and Games, Sacramento, CA.

Klamt, R.B., 2001a. Personal communications with the staff of the North Coast Regional Water Quality Control Board (NCRWQCB), Santa Rosa, CA.

Klamt, R.B., 2001b. Methods Manual for Water Quality Data Gathering and Analysis for the North Coast Watershed Assessment Program. North Coast Regional Water Quality Control Board, 5550 Skylane Boulevard, Suit A, Santa Rosa, CA 95403.

Knopp, C., 1993. Testing Indices for Cold Water Fish Habitat. Final Report for the North Coast Regional Water Quality Control Board, Santa Rosa, CA.

Kollias, V.J., Kalivas, D.P., 1998. The enhancement of a commercial geographical information system (ARC-INFO) with fuzzy processing capabilities for the evaluation of land resources. Computers and Electronics in Agriculture 20, 79–95.

KRIS Noyo 2000. The Klamath Resources Information System (KRIS) for the Noyo River Watershed in Mendocino County, California. Institute for Fisheries Resources, San Francisco, CA.

Kuncheva, L.I., Wrench, J., Jain, L.C., Al-Zaidan, A.S., 2000. A fuzzy model of heavy metal loadings in Liverpool bay. Environmental Modelling and Software 15, 161–167.

Leland, D., 2001. Personal communications with the staff of the North Coast Regional Water Quality Control Board (NCRWQCB), Santa Rosa, CA.

Lewis, J., 1998. Evaluating the impacts of logging activities on erosion and suspended sediment transport in the Caspar Creek watersheds. In: Proceedings of the Conference on Coastal Watersheds: The Caspar Creek Story, May 6, 1998, USDA Forest Service Gen. Technical Report PSW-GTR-168.

Lisle, T.E., 1989. Sediment transport and resulting deposition in spawning gravels, north coastal California. Water Resources Research 25, 1303–1319.

Lisle, T.E., Hilton, S., 1992. The volume of fine sediment in pools: an index of sediment supply in gravel-bed streams. Water Resources Bulletin 28, 371–383.

Lisle, T.E., Lewis, J., 1992. Effects of sediment transport on survival of salmonid embryos in a natural stream: a simulation approach. Canadian Journal of Fisheries and Aquatic Sciences 49, 2337–2344.

Mangelsdorf, A., Lundborg, H., 1997. Garcia River Watershed Water Quality Attainment Strategy for Sediment. Prepared for the North Coast Regional Water Quality Control Board, 5550 Skylane Boulevard, Santa Rosa, CA 95403.

Mays, M.D., Bogardi, I., Bardossy, A., 1997. Fuzzy logic and risk-based soil interpretations. Geoderma 77, 299–303.

McBratney, A.B., Odeh, I.O.A., 1997. Application of fuzzy sets in soil science: fuzzy logic, fuzzy measurements and fuzzy decisions. Geoderma 77, 85–89.

McHarg, I., 1969. Design with Nature. Natural History Press, New York.

Newcombe, C.P., MacDonald, D.D., 1991. Effects of suspended sediments on aquatic ecosystems. North American Journal of Fisheries Management 11, 73–82.

NCRWQCB 1994. Water Quality Control Plan for the North Coast Region. North Coast Regional Water Resources Control Board (NCRWQCB), Santa Rosa, CA.

NCWAP 2001. North Coast Watershed Assessment Program Methods Manual (draft). North Coast Watershed Assessment Program (NCWAP), Sacramento, CA.

NMFS 1997. Biological Opinion and Conference Opinion on Implementation of Land and Resource Management Plans and Land Use Planning Documents. National Marine Fisheries Service (NMFS), Southwest Region, 501 West Ocean Blvd, Long Beach, CA.

Pereira, J.M.C., Duckstein, L., 1993. A multiple criteria decision-making approach to GIS-based land suitability evaluation. International Journal of Geographical Information Systems 7, 407–424.

PWA, 1998. Sediment Source Investigation for the Lower Eel River (draft). Prepared by Pacific Watershed Associates (PWA) for the US Environmental Protection Agency, San Francisco, CA.

Reynolds, K.M., 1999. EMDS Users Guide (version 2.0): Knowledge-Based Decision Support for Ecological Assessment. General Technical Report PNW-GTR 470. US Department of Agriculture, Forest Service, Pacific Northwest Research Station, Portland, OR. http://www.fsl.orst.edu/emds/.

Reynolds, K., Saunders, M., Foster, M., Olson, R., Schmoldt, D., Latham, D., Miller, B., Steffenson, J., 1996. A knowledge-based information management system for watershed analysis in the Pacific Northwest US. AI Applications 10, 9–22.

Reynolds, K.M., Jensen, M., Andreasen, J., Goodman, I., 2000. Knowledge-based assessment of watershed condition. Computers and Electronics in Agriculture 27, 315–333.

Rocke, D.M., Dai, J.J., 2002. A Knowledge-Based Decision Support System for Watershed Assessment for Sediment. Final report prepared for the California State Water Resources Control Board, Sacramento, CA.

Rosgen, D., 1996. Applied River Morphology. Wildland Hydrology, Pagosa Springs, CO.

Ruffolo, J., 1999. TMDLs: The Revolution in Water Quality Regulation. CRB-99-005, California Research Bureau, 900 N Street, Sacramento, CA 94237.

Salski, A., Sperlbaum, C., 1991. Fuzzy logic approach to modeling in ecosystem research. Lecture Notes in Computer Science XX, 520–527.

Schmoldt, D.L., Rauscher, H.M., 1996. Building Knowledge-Based Systems for Natural Resource Management. Chapman and Hall, New York.

Smith, P.N., 1997. Environmental project evaluation: a fuzzy logic based method. International Journal of System Science 28, 467–474.

Stoms, D.M., McDonald, J.M., Davis, F., 2002. Fuzzy assessment of land suitability for scientific research reserves. Environmental Management 29, 545–558.

Stone, N.D., Coulson, R.N., Frisbie, R.E., Loh, D.K., 1986. Expert systems in entomology: three approaches to problem solving. Bulletin of Entomology Society of America 32, 161–166.

Tappel, P.D., Bjornn, T.C., 1983. A new method of relating size of spawning gravel to salmonid embryo survival. North American Journal of Fisheries Management 3, 123–135.

USEPA, 1999. Noyo River Total Maximum Daily Load for Sediment. US Environmental Protection Agency (USEPA). http://www.epa.gov/region09/water/tmdl.

Van Ranst, E.H., Tang, H., Groenemans, R., Sinthurahat, S., 1996. Application of fuzzy logic to land suitability for rubber production in peninsular Thailand. Geoderma 70, 1–9.

Waters, T.F., 1995. Sediment in streams: sources, biological effects, and control. American Fisheries Society Monograph 7.

Walters, J.R., Nielsen, N.R., 1988. Crafting Knowledge-Based Systems. Wiley, New York.

Weaver, W.E., Hagans, D.K., 1994. Handbook for Forest and Ranch Roads: A Guide for Planning, Designing, Constructing, Reconstructing, Maintaining and Closing Wild-land Roads. Prepared for the Mendocino County Resources Conservation District, Ukiah, CA in cooperation with the California Department of Forestry and Fire Protection and the USDA Soil Conservation Service.

White, W.R., 1982. Sedimentation problems in river basins. In Studies and Reports in Hydrology. Project 5.3 of the International Hydrological Program, UNESCO.

Zadeh, L.A., 1965. Fuzzy sets. Information and Control 8, 338–353.

Zadeh, L.A., 1968. Probability measures of fuzzy events. Journal of Mathematical Analysis and Application 23, 421–427.

Zadeh, L.A., 1979. A theory of approximate reasoning. In: Hayes, E., Michie, D., Mikulich, L.I. (Eds.), Machine Intelligence 9. Hatsted Press, New York.