

Case Studies of Canadian Environmental Decision Support Systems

William Booty and Isaac Wong
Environment Canada
Canada

1. Introduction

This chapter will discuss two different decision support systems that we have developed for Canadian environmental applications. We will first discuss how these systems utilize data and models to solve domain-specific problems and focus on effectiveness rather than efficiency in the decision making processes. In particular we will discuss how they are useful in better understanding the complex interaction between land and water and how they also provide a method to make informed resource management decisions and that they require the integration of scientific data, information, models and knowledge across multi-media (air, land and water), multi-disciplines and diverse landscapes. We will discuss how modelling is an important asset of any environmental decision support system (EDSS), particularly considering the high cost of full scale field work. Modelling presents a cost effective approach to assess the impact on the environment.

We will discuss how a typical EDSS needs to be developed to address the issues of linking multi-media models at different geospatial scales, how it provides interfaces that can accept, select, link and recalibrate discipline-specific component models, and how it can seek optimal solutions for a given domain problem.

We will also discuss that very often the EDSS is built around the concept of a management user interface to assist policy makers in their decision making. The technical users employ other tools to build model inputs, execute, and calibrate and validate the models while the management or policy makers view the inputs and outputs of the system that the technical users have built. This will allow management to investigate the analytical results based on robust science built by the researchers. Key functionality includes mapping and visualization of the results, scenario gaming and key statistical analyses of the results.

The first example we will discuss is the Environmental Effects Modelling Statistical Assessment Tools Decision Support System. We will discuss how it provides a user-friendly data analysis, display and decision support tool for Canada's federal environmental effects monitoring program for pulp and paper and mining industries. We will describe how the tool allows the assessment of the effects of effluent from industrial or other sources on fish and benthic populations. We will explain that in many of our EDSS systems, it is coupled with artificial intelligence such as expert systems to guide the users in the right direction. We will explain how the results are used in assessing the adequacy of existing regulations for protecting aquatic environments. We will explain how the design of such an EDSS has benefited from significant input from scientists, researchers, other end-users, system

Source: Decision Support Systems, Book edited by: Chiang S. Jao,
ISBN 978-953-7619-64-0, pp. 406, January 2010, INTECH, Croatia, downloaded from SCIYO.COM

developers, modellers and Geographic Information System (GIS) specialists. The integration includes data, maps and models with user-friendly tools, including data input/output views, map input/output views, and modelling result views for interpretation, further analysis, conclusion and recommendation with the aid of the expert system approach.

The second example of an EDSS is one that has been developed by Environment Canada to provide policy makers with a tool to help in examining management options for dealing with the impacts of land use on water for agricultural issues in Canada. The system deals with both temporal and spatial consistency among component models, where the output from one model is used as input to another in a sequence of linked calculations. In this example, the dynamic landscape model generates land use maps for various land use scenarios that can be used either in a single storm event non-point source pollutant model such as the Agricultural Non-Point Source Pollutant Model (AGNPS) (Young et al., 1987), or in a continuous time non-point source pollutant model such as the Soil and Water Assessment Tool (SWAT) (Arnold et al., 1998). It also provides the ability for scenario gaming, and testing, pollutant source tracing and the determination of optimal solutions. Examples will be provided of its application to a watershed in Ontario Canada.

Finally we will summarize the effectiveness of these systems and some insights as to how they might be improved and future directions.

Understanding complex environmental problems and making informed resource management decisions requires the integration of scientific data and knowledge across multiple disciplines and diverse landscapes. Ever increasing demands for timely, accurate and spatially explicit information require environmental modellers to deploy the latest information technology to provide decision support for various departmental priorities, such as global climate change, point source and non-point source pollution, lake eutrophication, biodiversity and ecosystem sustainability.

Decision Support Systems (DSS) (Alter, 1980) are computer-based interactive human-computer decision-making systems that assist policy makers in decision making processes. These systems utilize data and models to solve domain-specific problems and focus on effectiveness rather than efficiency in decision making processes. They also make informed resource management decisions and require the integration of scientific data, information, models and knowledge across multi-media (air, land and water), multi-disciplines and diverse landscapes in better understanding complex environmental issues.

In order for any decision support system to be a success, a proper design process is critical. The design of any environmental decision support system (EDSS) should come from a diversified functional group. They are scientists, environmental modellers, decision support system developers, computer programmers and component specialists such as Geographic Information Systems. Each of them contributes certain aspects of the system and how all the pieces fit into the system seamlessly. The system's blueprint should come from scientists who understand what is most required. Figure 1 depicts a schematic of an EDSS concept. At a glance, one can see that this kind of system offers a generic framework to integrate data, text, maps, objects, images, videos, environmental models and knowledge with user-friendly tools, including database management systems, mapping systems, visualization, advanced statistics, analytical functions and expert systems/artificial intelligence tools to produce outputs for interpretation, integration, post analysis and recommendation.

An EDSS should be able to handle data rich and poor situations. It should be functionality rich so the EDSS users not only perform the existing analytical routines but also expand to incorporate new ideas. Thus, a good EDSS should consist of the following functions and features:

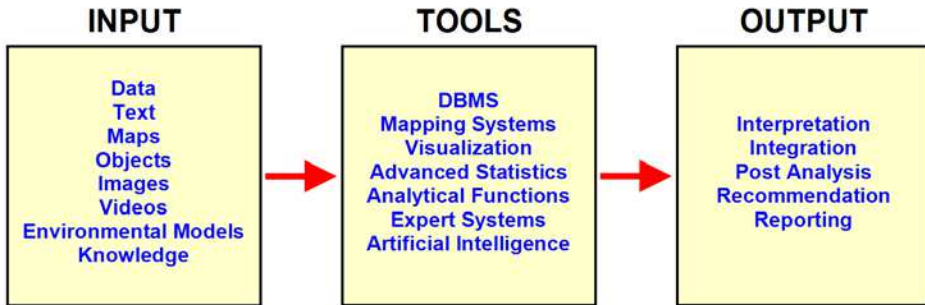


Fig. 1. Conceptual diagram of an environmental decision support system

1. User-friendly interface
2. Integrate and apply multiple tools such as environmental models, database management system, geographical information system, advanced statistics, artificial intelligence and, visualization in the analytical process
3. Assist users in integrating and inspecting complex and multiple input data and output results files that are at different spatial and temporal scales
4. Simplify the existing analytical process and allow room for any expansion
5. Provide an avenue to explore “what-if” option to compare scenarios

2. Case 1: environmental effects monitoring decision support system

2.1 Background

In Canada, pulp and paper and mining industries are currently required to conduct EEM under the Pulp and Paper Effluent Regulations (PPER) and Metal Mining Effluent Regulations (MMER). The Canadian EEM program is the first regulated, mandatory program of this magnitude in the world. The EEM involves iterative phases of monitoring and reporting. Each monitoring cycle of EEM requires submission of data and reports by pulp and paper mills and mines across Canada (Walker et al. 2002). This paper uses pulp and paper mill data as an example but the expert system also applies to the metal mining industry.

The key components of the EEM program include a fish survey to assess effects on fish and a benthic invertebrate community survey to assess effects on fish habitat. For the purposes of EEM, an effect is generally defined as a statistically significant difference in fish or benthic invertebrate community endpoints measured between an area exposed to effluent and a reference area or a statistically significant gradient in these endpoints from the exposure to reference areas. For the EEM, certain specified data (the effect endpoints, see Table 1 for fish endpoints) that are generated from the fish survey, benthic invertebrate community survey and fish usability studies were designated to assess the presence and level of effects.

The current objective of the EEM program is to evaluate the effects of pulp and paper or mining effluents on the aquatic environment. The program uses both field monitoring and laboratory approaches to directly or indirectly assess the health of fish, habitat impairment and concerns relating to human consumption of fish. It is also being used to discriminate between pulp and paper mill or mine-related effects and other natural or anthropogenic stressors. The program is designed to generate all relevant data and knowledge which may

Type of response	Effect endpoint	Statistical procedure
Survival	Age	ANOVA
Energy use	Size-at-age (body weight against age) Relative gonad size (gonad weight against body weight)	ANCOVA ANCOVA
Energy storage	Condition (body weight against length) Relative liver size (liver weight against body weight)	ANCOVA ANCOVA

Table 1. Endpoints to be used for determining exposure-associated effects on fish as designated by statistically significant differences between exposure and reference areas.

be used for analysis of both spatial and temporal trends in a way that can be reliably interpreted.

It should be pointed out that not all effects identified in the EEM will represent damage to fish, fish habitat, or the usability of the fisheries resources, but the effects can represent differences or gradients that may reflect changes to the ecosystem associated with the effluent. Detailed information on the effects, including the magnitude, geographic extent, and possible cause are used in the management of the aquatic resources.

Early in the EEM program it became clear to the EEM office that the assessment results coming back from the various industries were not always being carried out correctly or consistently. This was due to several factors. The first factor is that the collected data needs to be processed into a form that the statistics packages can accept, as well as carrying out log transformations if necessary and looking for outliers. Another factor is that the outputs of the results are often scattered because of multiple endpoints that need to be examined. It is also difficult to prepare the data for each statistical end point test, to make a decision as to whether there is a significant effect, and then to be able to move on to another end-point test procedure, if it is required. Finally, the results of all of the statistical tests need to be retrieved from various files and reviewed. The EEM Statistical Analysis Tool Decision Support System (EEM-SAT DSS) was conceptualized and developed to solve these problems with the aid of an expert system module. This chapter focuses on the development of the EEM-SAT DSS as well as examples of outputs from the system rather than the development of the EEM approach itself. The reader is directed to the following publications for details on the EEM approaches and designs (Lowell et al., 2002; Walker et al., 2002; Environment Canada, 2000a, 2002).

2.2 Key design considerations of the EEM-SAT decision support system

The flow schematic of the EEM-SAT procedures is shown in Figure 2. The EEM-SAT Expert System was developed to solve the problems of inconsistent data entry, data sampling collection, human induced errors, and the misinterpretation of the EEM statistical functions that were detected by the EEM office. There are many statistical treatments and routines required for testing the endpoint effects. In addition, reference sites must be established with domain experts to ensure the appropriateness and quality of these sites. The EEM procedures are tedious, time consuming and error prone if they are not automated. In the wake of these shortcomings, an EEM-SAT DSS was developed to fully integrate with the EEM statistical database. Since the test data is collected on site, it is more appropriate to

allow the consultant at each company to use the automated procedure in the DSS to analyze the most recent test data. Therefore, the development of the DSS takes great care in terms of defining the correct reference data for each company, the data validation and outlier screening, the implementation of rigorous statistical routines for the endpoint effect analysis and the site assessment report which indicates the overall condition of the fish and benthic invertebrate communities. When an extreme effect is detected, a company can take the proper course of action for remediation.

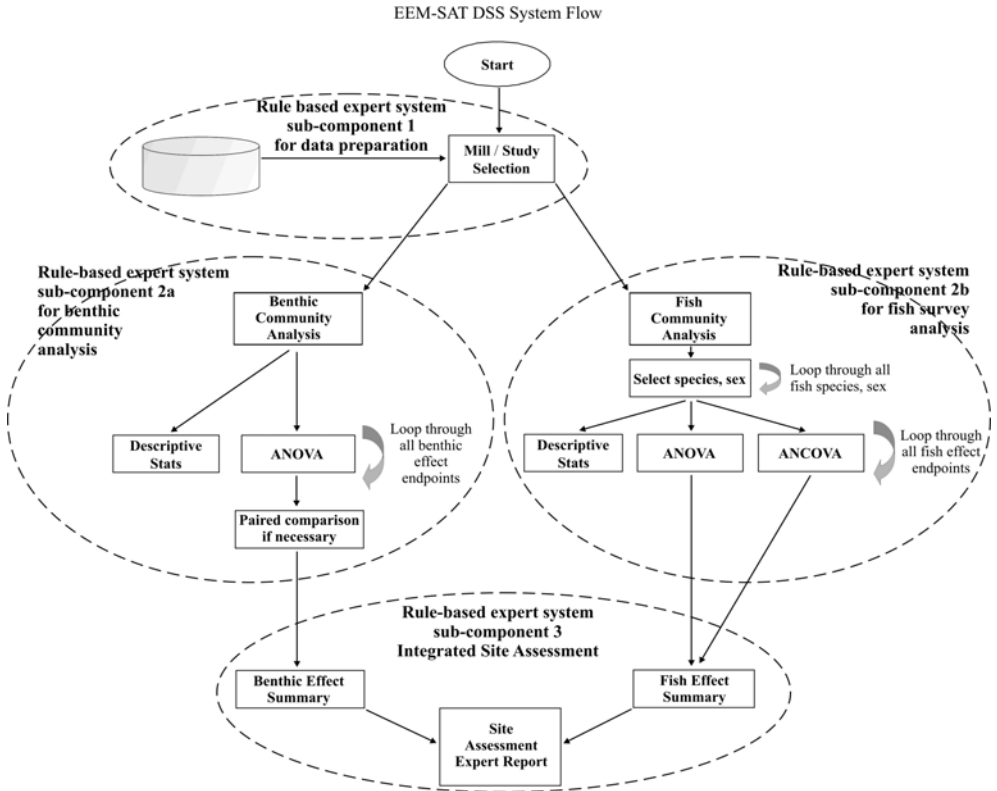


Fig. 2. EEM- SAT schematic diagram

The main driver of the EEM-SAT DSS is its rule-based expert system. The expert system provides control in the automation of the EEM-SAT procedures, linking appropriate test datasets with the reference datasets, utilizing the appropriate statistical procedures for each of the endpoint effects and determining the magnitude of the effect if it is found to be present. It has been developed based upon the RAISON Decision Support System framework (Booty et al., 2001; Lam et al., 2004), with interfaces constructed using Visual Basic. Expert systems, a form of artificial intelligence, are human computer systems that perform problem-solving tasks in a specific domain (Ignizio, 1991; Buchanan and Shortliffe, 1985). The systems are useful tools in numerous application areas including environmental domains. They apply heuristic rules to encode domain knowledge, together with inference engines, in order to deduce conclusions from information that the users provide. Decision

support systems are systems that employ various techniques that include artificial intelligence. In particular, expert systems can be integrated with more classical techniques of functionality such as statistics, mapping and/or data retrieval to form systems that provide more effective decision support in a study domain. The expert system module of the EEM-SAT system acts as a wizard. It has several expert system components including rule bases about the limits of its applicability, what kind of input data it requires, how to remove outliers, how to estimate its parameters from available information, how to extract data from the database, how to execute the statistical routines and correctly interpret the statistical results. In general, the expert system component assists users with less expertise in EEM to better perform their job. The EEM-SAT DSS is an example of a typical environmental decision support system. Figure 3 illustrates the user interface of the EEM-SAT DSS. The EEM-SAT expert system components contain only small knowledge bases, but they dramatically improve the functionality of the EEM-SAT DSS. The three major components of the expert system are described below.

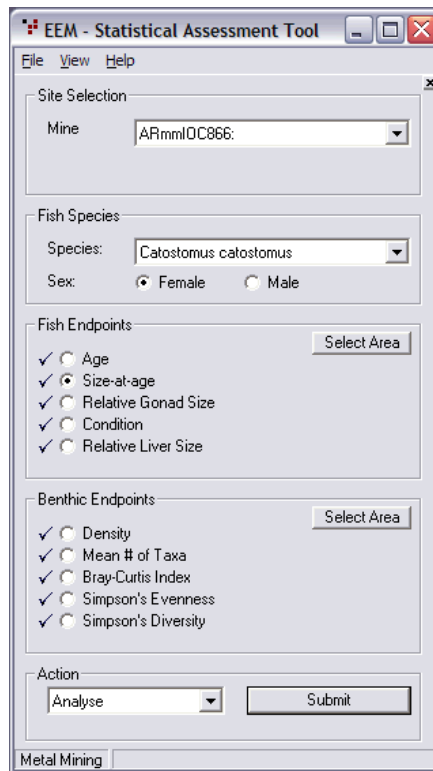


Fig. 3. EEM-SAT DSS user interface

2.3 EEM-SAT DSS expert system components

2.3.1 Intelligent data preparation and screening process

The expert system module guides the users to provide the appropriate data and information. The users can either submit the information interactively or via the database.

The system builds rules based on expert knowledge to return the pulp and paper mill and mill IDs that have data for certain types of fish species and invertebrate analysis. It provides an effective means to identify outliers based on robust statistical techniques. Finally, it also examines the nature of the data and determines if it is necessary to perform data transformations before submitting the data for the statistical routines. The rules used for fish are as follows:

1. If a pulp and paper mill is selected, then access the database to return available fish species and determine if there are enough benthic data for the EEM statistical analyses.
- 2a. If enough data is available, then the system displays the data and asks the user to select species and appropriate statistical tests.
- 2b. If not enough data is available, then it advises the user of the data gap, which data areas are lacking and signals to the DSS that statistical routines cannot proceed.
3. If fish species and benthic data are selected, then it advises the user of any potential outliers.
4. If fish species and benthic data are selected and the outlier detection procedure is complete, then examine the data to determine if a logarithmic transformation of data is required.
- 5a. If data transformation is required, then perform the necessary data transformation and submit the processed data to the fish and benthic statistical routines.
- 5b. If data transformation is not necessary, then submit the data to the fish and benthic statistical routines.

The rules for the benthic organisms are as follows:

1. Isolate data: data is selected on the basis of study ID, species and gender/sex. If there is not enough data for the study, the system will reject the request and inform the user to supply the missing information for the analysis.
2. Log₁₀ transform dependent and independent variables (if necessary): logarithmic transformations are preferred in the endpoint effect analysis simply because biological measures are often considered to be logarithmic or exponential scale. One of the criteria of the ANOVA and ANCOVA procedures is that the data should be normal. This includes body weight, total length, fork length, standard length, age, gonad weight and liver weight.
3. Checking for outliers: scatter plots of fresh weight vs. age, gonad weight vs. fresh weight, fresh weight vs. length, and liver weight vs. fresh weight (all variables are in logarithmic scale) are presented. When a scatterplot illustrates outliers, the user should be given an opportunity to identify, modify, and/or delete data. Although there is no formal guidance on screening data on the basis of studentized residuals, a rule of thumb is that when the studentized residual exceeds 4, this indicates that the observation may be unusual and the observation should be removed and the analysis should be rerun.

In addition, the user should pay attention to the high "leverage" values. Those observations tend to potentially skew the observed relationship in one direction or another. There is professional judgment used to determine whether the data with high leverage should be excluded. A common approach is that if they grossly skew the expected relationship, then exclusion should be considered.

2.3.2 Integrated advanced statistics for fish community end-point analyses

The EEM DSS uses both field monitoring and laboratory approaches to directly or indirectly assess the health of fish, habitat impairment, and concerns relating to human consumption

of fish. The objective of assessing impairment is met through assessment of the endpoint effects in both fish and benthic communities. The overall fish analysis pathway and its endpoint effects are shown in Figure 4. Fish monitoring for the EEM DSS involves monitoring both sexes of two sentinel species at reference and exposure areas to assess if there are differences in the growth, reproduction, survival or condition of fish populations. Sex differences are common due to differences in overall energetic requirements between male and female fish. Effect endpoints include weight at age, size-at-age, relative gonad size, liver weight, and condition factor for fish and taxon richness, Simpson’s diversity and evenness indices, and Bray–Curtis index for benthos, as shown in Table 1. Simpson’s diversity index is a measure of the character of a community that takes into account both the abundance patterns and the taxonomic richness of the benthic invertebrate community. It is

Fish Details Program Flow

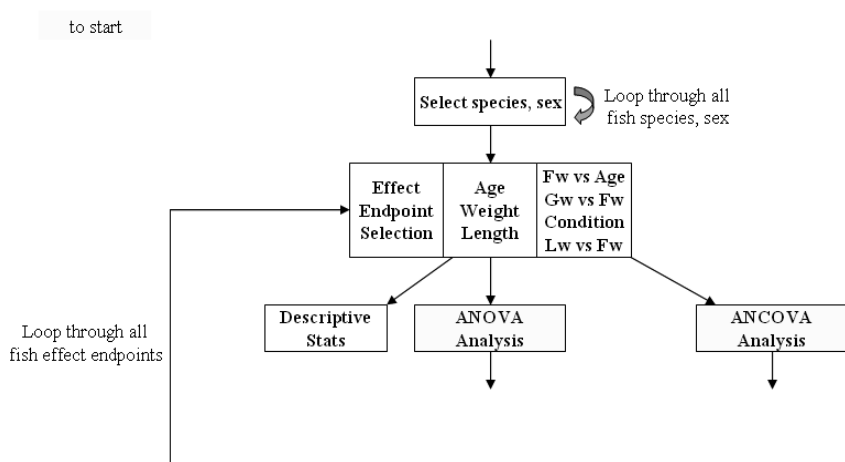


Fig. 4. Overall fish analysis pathway and its endpoint effects schematic.

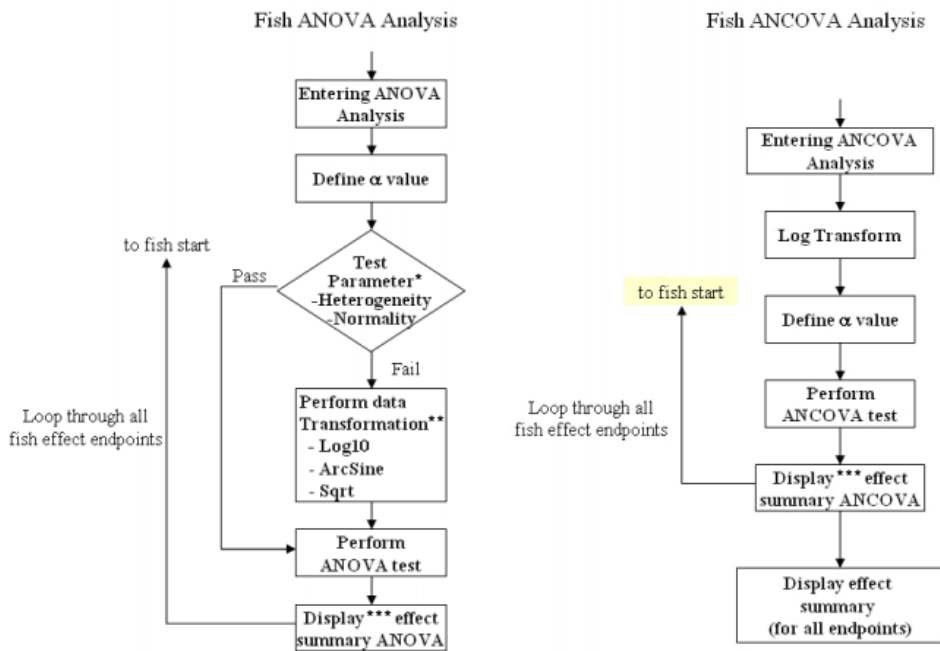
Type of response	Effect endpoint	Statistical procedure
Survival	Age	ANOVA
Energy use	Size-at-age (body weight against age)	ANCOVA
	Relative gonad size (gonad weight against body weight)	ANCOVA
Energy storage	Condition (body weight against length)	ANCOVA
	Relative liver size (liver weight against body weight)	ANCOVA

Table 1. Endpoints to be used for determining exposure-associated effects on fish as designated by statistically significant differences between exposure and reference areas.

calculated by determining, for each taxonomic group, the proportion of individuals that it contributes to the total sample. For the Simpson’s evenness index, evenness measures how similar the frequencies of the species are. When all the frequencies are equal, evenness is one. Frequency imbalance lowers the Simpson’s evenness index. The Bray–Curtis index is an index that measures the degree of difference in community structure (especially community composition) between sites. This measure helps to evaluate the amount of dissimilarity between benthic invertebrate communities at different sites.

A summary of the EEM-SAT fish analysis expert system rules is as follows:

1. Analyses of variance (ANOVA) test (Figure 5): carry out ANOVA to test for differences between areas, and calculate means and standard deviations for each key variable for both areas (reference and exposure). Once the groups are identified to be significantly different, the user needs to determine which pairs differ. This is done using Tukey’s HSD post hoc test (SYSTAT 11 Statistics II, 2005). Assumptions for ANOVA are that the data for reference and exposure populations are normally distributed; the variances are



- *Note:
 - 1) View residual plots for reference, exposed
 - 2) Check for patterns to determine if transformation is required. If so, which transformation is appropriate.
 - 3) Screen or indicate outlier, if required.
- **Note: The type of transformation will be based on the recommendation of the science committee
- ***Note: All results convert back to non-transformed state for display

Fig. 5. Fish ANOVA and ANCOVA Analysis schematic diagrams

equal between the reference and exposure populations and the error terms are independently distributed.

- Analyses of covariance (ANCOVA) and regression (Figure 5): this is done using the General Linear Model (GLM) (Environment Canada, 2000a). In particular, a contrast statement is used to test relationships among reference vs. exposure. The test is composed of two parts. It is carried out first to determine whether the slopes are approximately parallel. If the slopes are parallel, it then requires determining if the elevations of the regressions are significantly different. ANCOVA combines the features of ANOVA and regression, and can be used to compare regressions among treatments (i.e. reference vs. effect areas). Assumptions of ANCOVA are that the residuals are normally and independently distributed with zero mean and a common variance; the independent variable (covariate) is fixed and measured without error; the relationship has the form specified (linear regression) and the slopes of regression lines among areas are equal.

2.3.3 Integrated advanced statistics for benthic community end-point analyses

For the purposes of the EEM program an “effect on the benthic invertebrate community” means a statistical difference between benthic invertebrate community measurements taken in an exposure area and a reference area (e.g., control/impact design) or a statistical difference between measurements taken at sampling areas in the exposure area that indicate gradually decreasing effluent concentrations (e.g., a gradient design). The EEM-SAT DSS program only pertains to control/impact analyses at this stage of development. This design uses ANOVA to detect differences between reference and exposure areas. The six basic study designs and their associated statistical procedures are shown in Table 2.

Study Design	Statistical procedure
Control/impact	ANOVA/ANCOVA
Multiple Control/Impact	ANOVA/ANCOVA
Simple Gradient	Regression/ANOVA/ANCOVA
Radial Gradient	Regression/ANOVA/ANCOVA
Multiple Gradient	ANCOVA
Reference Condition Approach	Multivariate/ANOVA/ANCOVA

Table 2. Statistical procedure used to evaluate exposure-associated effects on benthic invertebrates for each of the six basic study designs.

The benthic analysis procedures are represented in Fig. 6. The expert system rules for the benthic invertebrate are the same as the fish analyses with different effect endpoints. The effect endpoints for benthic invertebrate analyses are abundance (density), mean number of taxa, Bray–Curtis index, evenness and Simpson’s diversity index. These descriptors are largely summary metrics selected to encompass the range of effects, which may be a result of mine or pulp and paper mill effluents.

The α value is set to 0.10 for both fish and benthic invertebrate in determining effects. This is because the fish and benthic invertebrate community survey should minimally have sufficient statistical power analysis to detect an effect size of two standard deviations. Table 3 indicates that α and β should be able to achieve both Type I and II errors at 0.10 since sample size is usually about 5 for both fish survey and benthic community studies.

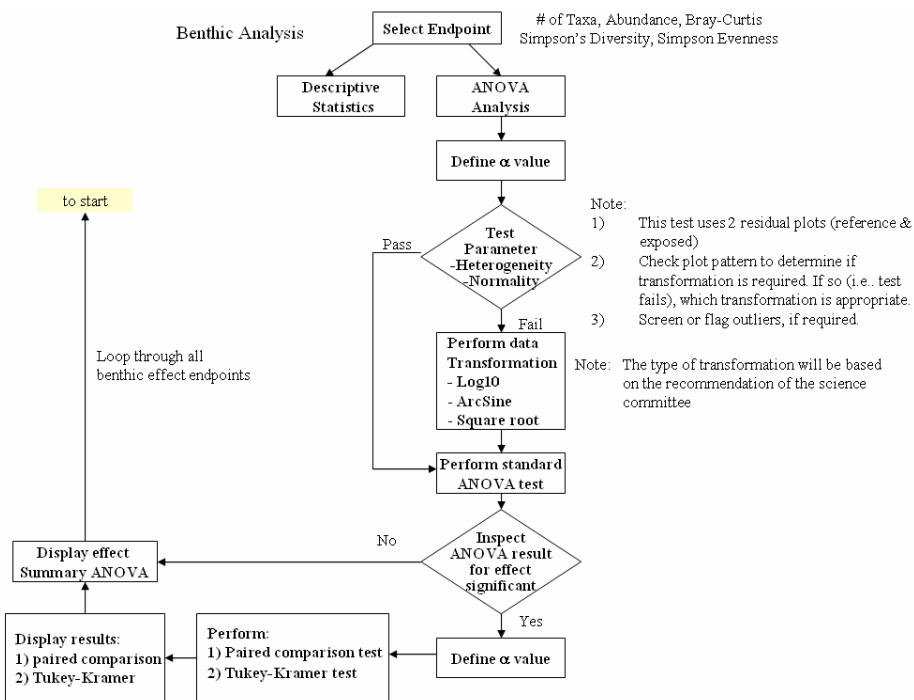


Fig. 6. Benthic Analysis schematic diagram.

A	1-β			
	0.99	0.95	0.90	0.80
0.01	14	11	10	8
0.05	11	8	7	5
0.10	9	7	9	4

Table 3. Sample size required to detect an effect size of ± 2 standard deviations

2.3.4 Integrated site assessment expert report to interpreting the analysis results

The expert system module ingests all the relevant information and the outputs for the current analysis. Specifically, a detailed summary of the current state of the analyses is provided. It will highlight and provide detailed information of potential troubled areas of the statistical analyses. For example, what tests do the fish species fail and by how much? This report allows the users to deal with the problem properly. The rules used in this subcomponent of the system include:

1. If the statistical routines are successfully completed, then extract metadata information of the pulp and paper mill, the name of the statistical routines and the summary results of each test.
2. If some of the results are outside of the normal bounds, then extract the detailed information of those tests and include in the summary results. The detailed information

includes the name of the test and all statistical data such as the degree of freedom, the test value, the critical value and the confidence level. The system will proceed to produce a complete and integrated report to the user to pinpoint the critical results of the full analyses.

2.4 An EEM-SAT DSS example

2.4.1 Fish size-at-age (effect endpoint)

In this example of the decision support process, the rates of growth are described by the relationship of size (as weight or length) to age. Over the entire life span of the fish, this relationship is curvilinear, with the rate of increase declining as fish approach the limit of their life span. Size-at-age may be estimated by calculating the regression relationship between body size (weight or length) and age for each sampling area (reference and exposure). Calculation of mean age is meant as a gross reflection of the age distribution of adult fish collected from each area. The EEM-SAT DSS automatically \log_{10} transforms all ANCOVA based analyses. It also uses weight as the covariate. Body weight is corrected by subtracting the gonad weight and liver weight from the body weight prior to analysis. The EEM-SAT DSS follows the two-step (slopes, then intercept) analyses. If slopes are not significantly different at the alpha value specified by the user, then the test for intercepts (least squared means) proceeds automatically. If slopes are significantly different, the software will not test for differences in intercepts. Endpoints analyzed by ANCOVA (size-at-age, relative gonad size, condition, and relative liver size) have only one component in the EEM-SAT: 1) Effect Analyses. Any descriptive measures associated with the analyses are included in the ANCOVA result table. An example is shown in Figure 7. In this example, the effect endpoint "size at age" is examined. Using ANCOVA testing the slope difference between the "Reference" and "Effects" data shows that test p-value is 0.854 and is larger than the α , (0.10), therefore, there is no significance difference. However, testing difference of the means indicates that the test p is almost zero and is statistically significant against α , (0.10).

2.4.2 Benthic endpoints

All benthic endpoints are analyzed by ANOVA and have two components in the EEM-SAT: 1) Descriptive Statistical Analyses and 2) Effect Analyses. Examples of an effects analysis of a site (name deleted for privacy) are presented below. In Figure 8 for mean number of taxa, the test p-value is less than α (0.10) indicating that there is a significant difference (magnitude of -45%) between the reference and the near field sites. The Bray-Curtis Index results are shown in Figure 9 where it can be seen that the test p-value is less than α (0.10), with a magnitude of +83.8%, again indicating a significant effect.

2.4.3 Integrated site assessment expert report

The system summarizes the results of the analyses for each site as shown in Figure 10. It can be seen that for fish species *Catostomus catostomus*, all of the endpoints show an effect. For benthos, all but the Simpsons Evenness endpoint indicates a significant effect.

The EEM-SAT DSS file open and save options allow users to generate scenarios for comparison with different levels of significance. In this program, a key point is to maintain high probabilities of correctly identifying areas that are actually impaired (high statistical power), while still maintaining low probabilities of falsely concluding that impairment has occurred in nonimpacted areas (low α values).

Effect Results: Fish

Species: *Catostomus catostomus* - Female adult

Mine: ARmmIOC866:

Data: Phase 1

Transform: Log10

ANCOVA: Size-at-age

Area	n	Log Transformed		Slopes Different?		Log Transformed		Means Different?		AntiLog Adjusted Means	Magnitude	α
		Slope	R ²	(p value)	sig @ p< α	Adjusted Means	SD (Adjusted Means)	(p value)	sig @ p< α			
R	20	0.434	0.428	-	-	3.218	0.079	-	-	1.653E+3	-	0.10
E	20	0.483	0.187	0.854	No	3.046	0.079	1.363E-7	Yes	1.111E+3	-32.8%	0.10

Size at age: *Catostomus catostomus* - Female Adult

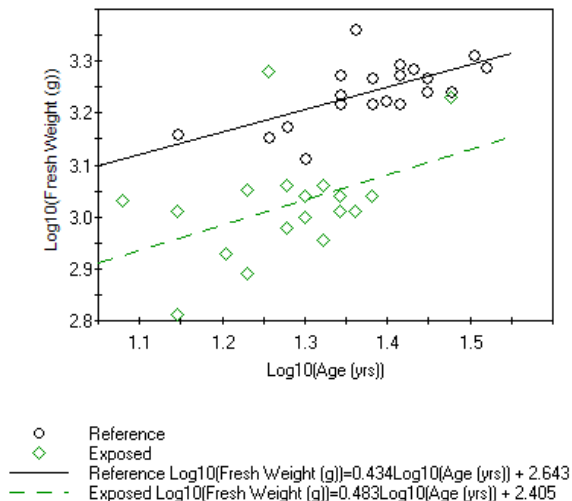


Fig. 7. ANCOVA fish size-at-age analysis example results

Effect Results: Benthic Invertebrate

Mine: ARmmIOC866:

Data: Phase 1

Transform: None

ANOVA: Mean # of Taxa

Source of Variation	SS	df	MS	F	p value	sig @ p < α			Ref	Exp	Direction	Magnitude
									Mean	SD		
Model	32.400	1	32.400	81.000	1.853E-5	Yes	0.10	8.000	4.400			
Error	3.200	8	0.400					0.707	0.548			
Total	35.600	9						0.316	0.245			
											Ref > NF	- 45.0%

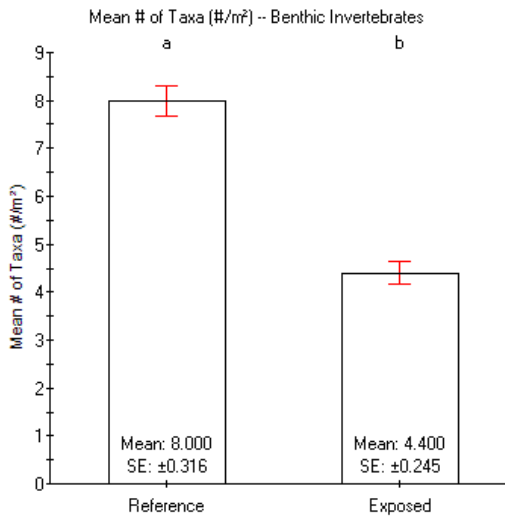


Fig. 8. Benthic Mean number of taxa example results.

Effect Results: Benthic Invertebrate

Mine: ARmmIOC866:

Data: Phase 1

Transform: None

ANOVA: Bray-Curtis Index

Source of Variation	SS	df	MS	F	p value	sig @ p < α	α			Direction	Magnitude	
								Ref	Exp			
								Mean	0.217	0.400		
								SD	0.085	0.115		
								SE	0.038	0.051		
Model	0.083	1	0.083	8.135	0.021	Yes	0.10	n	5	5	Ref < NF	+ 83.8%
Error	0.082	8	0.010					Min	0.136	0.263		
Total	0.164	9						Max	0.316	0.533		

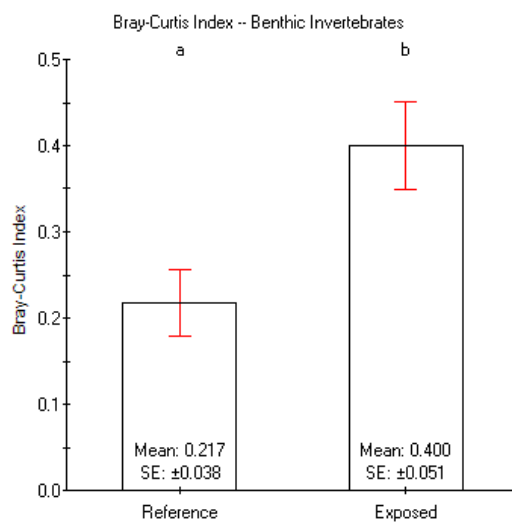


Fig. 9. Benthic Invertebrate example Bray-Curtis Index results.

Site Effect Summary

Mine: ARmmIOC866:

Data: Phase 1

Trophic Level	Species	Sex	Endpoint	Effect?	Direction	Magnitude	Transform
Fish	Catostomus catostomus	Female	Age	Yes	Ref > Exp	- 20.7%	None
			Size-at-age	Yes	Ref > Exp ²	- 32.8% ²	Log10
			Relative Gonad Size	Yes	Ref > Exp ¹	- 62.7% ¹	Log10
			Condition	Yes	Ref < Exp ¹	+ 70.9% ¹	Log10
			Relative Liver Size	Yes	Ref > Exp ¹	- 45.8% ¹	Log10

Trophic Level	Endpoint	Effect?	Direction	Magnitude	Transform
Benthos	Density	Yes	Ref > Exp	- 55.1%	Log10
	Mean # of Taxa	Yes	Ref > Exp	- 45.0%	None
	Bray-Curtis Index	Yes	Ref < Exp	+ 83.8%	None
	Simpsons Evenness	No			None
	Simpsons Diversity	Yes	Ref > Exp	- 43.7%	None

Fig. 10. Effect summary results for fish and benthic invertebrates

3. Case 2: land and water integration decision support system

3.1 Background

Agricultural activities such as animal farming, grazing, plowing, pesticide spraying, irrigation and fertilizer applications can cause non-point source or diffuse pollution. Nutrients and sediment are two of the main agricultural pollutants affecting water quality that result from these activities. Nutrients such as phosphorus and nitrogen are minerals that can be applied to enhance plant growth and crop production. When they are applied in excess of crop needs, the excess nutrients are often attached to soil particles that can be carried by overland water runoff from land into the aquatic ecosystems. The nutrients can cause excessive algae and aquatic plant growth in rivers and streams; cloud the water; reduce the amount of sunlight reaching aquatic plants; cover fish spawning areas and food supplies; greatly increase the costs of water treatment; reduce swimming and water recreation activities; create a bad smell; kill fish; and accelerate aging of rivers and lakes. Besides fisheries and recreation effects, these pollutants also have harmful effects on drinking water supplies and wildlife. Thus, there exists an important linkage between the land and the water. Environmental performance standards for floral and faunal communities in terrestrial ecosystems are based on assessments and forecasts of land cover and land use in agricultural regions. Performance standards for aquatic community

structure in streams are based on assessments and forecasts of flow regime, sediment levels and nutrient concentrations. However, the physico-chemical condition of a stream is strongly affected by catchment characteristics, including land cover and land use, but also by basin shape, surficial geology and soil structure. Thus, land cover and land use patterns will have profound impacts on both water quantity and quality, and aquatic biodiversity.

3.2 Key design considerations of the LWIDSS

Modelling is an important asset of any environmental decision support system. With the high cost of full scale field work, modelling presents a cost effective approach to assess the impact on the environment. In the LWIDSS, the emphasis is on the terrestrial and aquatic models that are commonly used to assess agricultural impacts.

Dynamic landscape modelling helps decision makers assess the consequences of alternative management scenarios at the landscape scale. It simulates land use scenarios that are characterized by different assumptions about management practices. The results are in the form of GIS spatial layers that can be evaluated at the landscape level or can be fed into other component models such as non-point source pollutant models to evaluate the impact of land to water quality. This approach differs from other modelling efforts in that it does not confine itself to just one model or a given set of models. Rather, it provides an open architecture framework that accepts any component model within the system that can be linked to other component models in the causal chain, be it a dynamic landscape model for land use scenario creation or a non-point source pollutant model for sediment and nutrients assessment.

Land use scenarios are integrated with watershed hydrology models to develop flow, sediment and nutrient performance standards in streams to protect aquatic biodiversity. In addition to the scenario representing the present day (current), others are developed to explore different land use cases (e.g., agricultural intensification). Validated and calibrated hydrologic models use these scenarios to estimate water quantity and quality parameters. These parameters are then used to forecast aquatic biodiversity according to empirically-derived relationships between stream flow, sediment and nutrient regimes and biotic condition. Benthic algal and invertebrate communities, as well as fish communities, function as the biotic endpoints of streams and rivers to gauge ecosystem integrity.

Non-point source pollutant modelling in general can be a large and complex process requiring great quantities of input and generating vast amounts of output because of its nature of trying to simulate real-world processes. Dealing with such sums of data, both inputs and outputs, can be daunting to those who are trying to understand and extract knowledge and information from them. Not all modelling programs contain tools for visualizing the results, comparing multiple sets of results, performing post-analysis or managing/organizing the data from different model runs. Typically, after executing the models, different software are employed to look at the output and to perform further statistical analysis and these can be time consuming procedures by themselves. It is apparent that it would be very useful to have an integrated set of tools in a single software system that performs these tasks (Lam et al., 1998). This would make those doing modelling more productive by allowing them to examine the results in a more efficient manner. More time can be spent on modelling and less time on manipulating data to move it into software programs.

The LWIDSS addresses the issues of linking multi-media models at different geospatial scales. It provides interfaces that can accept, select, link and recalibrate discipline-specific

component models. It can seek optimal solutions for best management practices such as buffer strip widths for sediment and nutrients reduction based on feedback of individual component models. The design of the LWIDSS has benefited from significant input from scientists, researchers, other end-users, system developers, modellers and GIS specialists. Figure 11 illustrates the schematic diagram of the LWIDSS. At a glance, the LWIDSS offers a medium to integrate the data and information of the land and water by providing a number of necessary functionalities. The integration includes data, maps and models with user-friendly tools, including data input/output views, map input/output views, and modelling result views for interpretation, further analysis, conclusion and recommendation. The LWIDSS is designed as a framework and can be easily adapted to any watershed as portability is an important aspect of design consideration.

The LWIDSS is built around the concept of a management user interface to assist policy makers in their decision making. The technical users employ other tools to build model inputs, execute, and calibrate and validate the models; the technical aspect of the modelling process is beyond the scope of this paper. The friendly interface provides a platform for the management or policy makers to view the inputs and outputs of the system that the technical users have built. This will allow management to investigate the analytical results based on robust science built by the researchers. Key functionality includes mapping of the results, scenario gaming and key statistical analyses of the study area.

The LWIDSS design also calls for both temporal and spatial consistency among component models as the output from one model is used as input to another in a sequence of linked calculations. For example, the dynamic landscape model generates land use maps for various land use scenarios that can be used either in a single storm event non-point source pollutant model such as Agricultural Non-Point Source Pollutant Model (AGNPS), which is grid based or in a continuous time non-point source pollutant model such as Soil and Water Assessment Tool (SWAT), which is vector based.

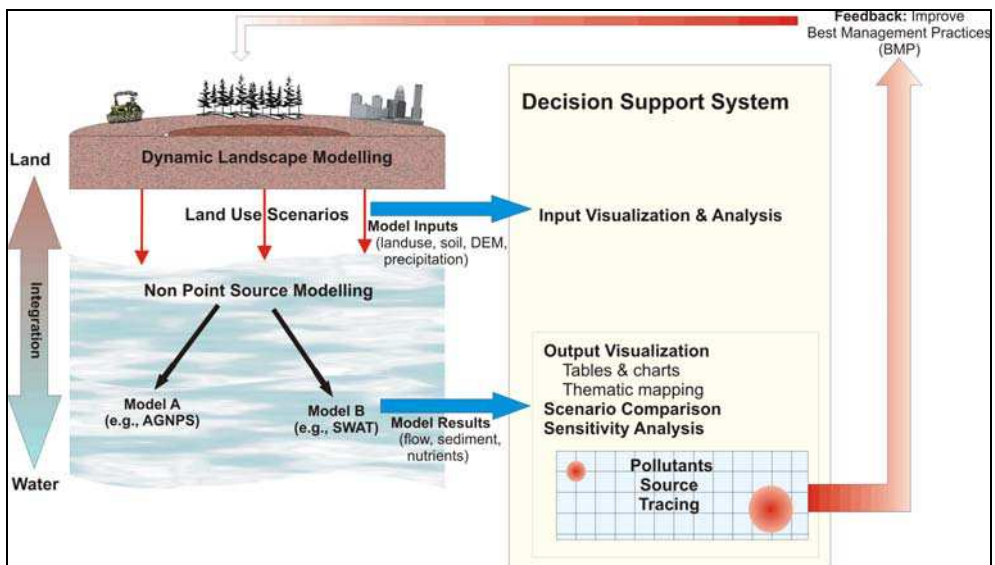


Fig. 11. Schematic diagram of the LWIDSS.

A major requirement for policy makers is to use the LWIDSS to predict results, including evaluation of different scenarios on land and water and optimization of BMP. These should be obtained in a relatively short computational timeframe with a friendly user interface. Using the modelling approach to understand the non-point source pollutant problem is important for providing the assessment of the impacts of land-water integration. In addition, implementing a scenario gaming approach would allow decision makers an opportunity to understand the problem based on different possible scenarios and to make viable decisions to manage the problem more effectively and to minimize the impacts. The feedback among component models is critical to the whole integration process. Different models can complement each other with their strengths. For example, the AGNPS model excels in identifying pollutant “hot spots” and the SWAT model can take the information on the “hot spots” to further evaluate the optimal solution of BMP.

3.3 Integrated analysis approach

The integrated analysis approach will be described through an example which decision makers setting environmental policies could potentially face. The problem is to investigate a study area in Ontario, Canada to assess its current state in terms of aquatic ecosystem health and to forecast future conditions if agriculture practices need to be intensified because of increasing demands for crop production. We identify the Raisin River Watershed, an agricultural watershed in Ontario, Canada as a pilot example for the development of the LWIDSS. It is selected to be the primary focus because agricultural activities can have an impact on the environment, in particular on the aquatic ecosystem. Current watershed conditions can be evaluated by gathering empirical data such as that done through water quality sampling. This usually involves sending people out into the field across the watershed at regular time intervals, collecting the water samples and analyzing the samples back at the laboratory. This method can be quite costly both in terms of time and money due to amount of labour and transportation and equipment requirements. In fact, it may not be feasible at all because of budget constraints or watershed size. Also, sites located in difficult to reach areas and poor weather can complicate the process. An alternate and more cost effective method is to make use of computer models to try to predict the water quality by simulating real-world physical processes occurring in nature (Leon et al., 2004). More specifically, some type of non-point source pollutant model is required to predict the level of pollutants (e.g., sediments) in the water system. Non-point source pollutant models such as AGNPS and SWAT, by themselves, are useful tools in aiding decision makers with setting up best management practices to improve water quality. But, when these models are coupled to a decision support system such as the LWIDSS, their effectiveness is enhanced. The LWIDSS provides value-added decision support functions to the models which otherwise may be lacking or deficient. In addition, it can make the overall decision support process easier and more productive for decision makers. Figure 12 illustrates the user interface of the LWIDSS. We will explain the development of these decision support processes in detail next.

3.4 Land water integrated decision support system components

3.4.1 Review of inputs

One of the first steps in any modelling work is to identify all of the input that is required by the model. In the case of watershed-based non-point source models, the physical characteristics of the watershed need to be described to the model for it to simulate the physical processes such as the hydrology. GIS map layers are used to define the soil texture, the land use and the surface topography (Digital Elevation Model) of the Raisin River

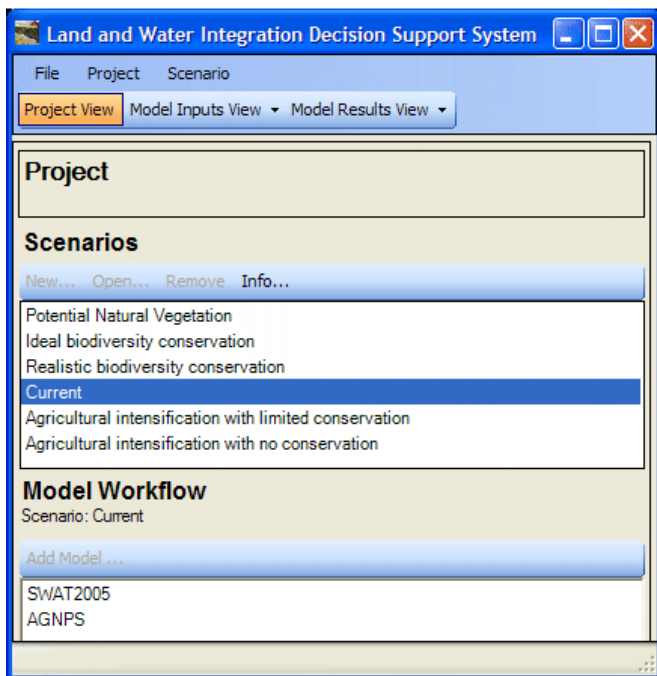


Fig. 12. User Interface of the LWIDSS

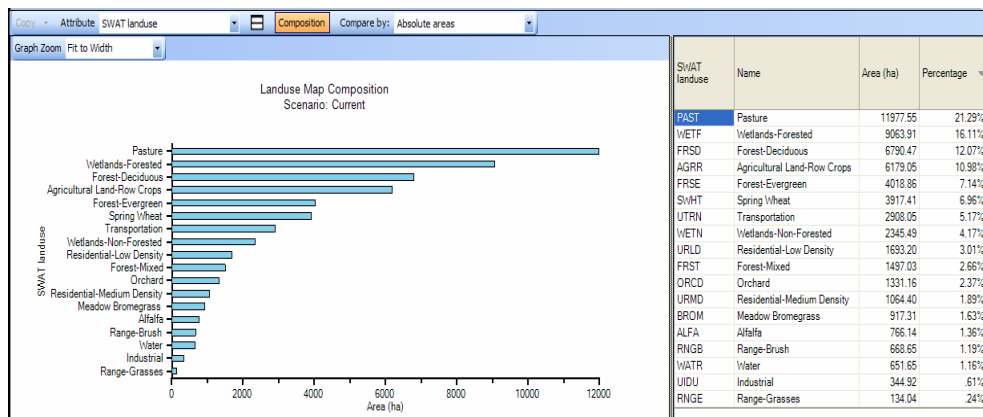


Fig. 13. Reviewing model inputs in the LWIDSS. A chart and table of land use composition in the watershed is presented to identify the primary uses of the land.

Watershed. In addition, non-spatial data such as climate data including precipitation and temperature may also be necessary to simulate rainfall and evapotranspiration (transport of water from the surface to the air). Reviewing the input data of the models within the LWIDSS using multiple formats including maps, graphs and tables (Figure 13) is an important aspect in the overall decision support process. It allows policy makers an opportunity to check and possibly confirm whether or not the results produced from the

model seem reasonable. Going back and examining the input data can also help to get some insight into why one is getting certain results when they may be expecting something different. The ability of the LWIDSS to overlay data from several GIS layers allows policy makers the ability to identify areas where data quality may be poor and thus require more attention. The ability to integrate data allows more valuable information to be generated.

3.4.2 Model calibration and validation

Non-point source models are run for the first time using current watershed conditions to assess the present situation (i.e., the current water quality). The accuracy of the model prediction is evaluated by comparing the results to known observation data (water quantity and quality). If the predictions are poor, then model variables are adjusted and the model is re-run until the predictions are satisfactory (unique watershed characteristics not reflected by the input data and data accuracy can cause the model to behave differently). This iterative process is referred to as model calibration and validation. The LWIDSS assists in this procedure by providing a platform for examining and visualizing the model outputs in a variety of different ways such as through graphs and tables and for comparing the outputs to the observation data. Model output from non-point source models are typically expressed as amounts of pollutants in the water (i.e., concentration). Results can be quickly accessed because they are organized by location within the watershed and by predicted parameter (e.g., stream flow and sediment). Results can also be filtered by time period and can be summarized to a broader timeframe. Predicted model outputs are compared to observations using plots of observations over predicted results (Figure 14) or using statistics such as regression coefficient R or Nash Sutcliffe coefficient (Nash & Sutcliffe, 1970) for quantitatively assessing the prediction accuracy.

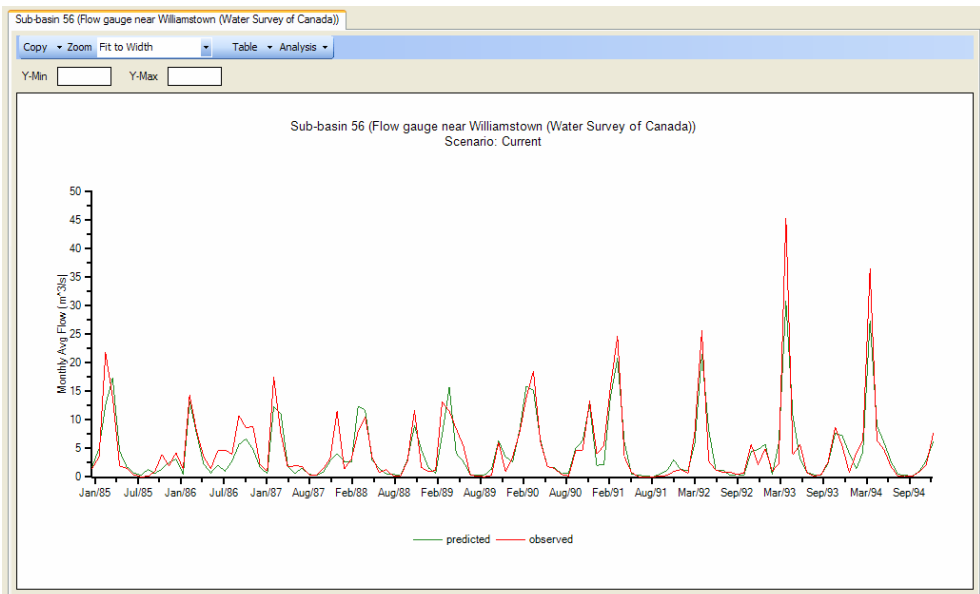


Fig. 14. Reviewing model results in the LWIDSS for the purpose of model calibration. A plot of predicted flow (model) and observed flow (from a monitoring station) is examined to gauge prediction accuracy.

3.4.3 Scenario gaming and comparison

After a model is calibrated, it is now ready to be used for prediction in other situations, real or hypothetical. These are commonly referred to as scenarios. Suppose, in the future, there will be a need to increase the amount of agriculture to produce higher crop yields. The question that needs to be addressed is: how much of an impact will this agricultural intensification scenario have on the environment? The only feasible and cost effective means of arriving at an answer is through modelling. This scenario requires that a new hypothetical land use map be developed by experts using dynamic landscape models. This land use map will undoubtedly involve increasing the amount of land available for farming. The LWIDSS should provide the ability for policy makers to review and compare the land use maps from multiple scenarios (Figure 16), thus allowing them to provide comments and feedback to the landscape modellers. The non-point source model is run using the new agricultural land use map (and with the other input data staying constant). Policy makers then use the LWIDSS to analyze the results for the new agricultural intensification scenario and to also perform a scenario comparison (Figure 15) against the current scenario (present day land use map) through scenario gaming (Wong et al., 2007). This allows them to assess how much the pollution is predicted to increase when compared to the current environment and for them to plan the next course of action.

3.5 Integrated modelling assessment

If the policy makers, after examining the model results, determine (based on current standards or guidelines) that the agricultural intensification scenario produces too much pollution, then what can be done to reduce the environmental impact (while still maintaining similar levels of agriculture needs)? An LWIDSS which has the capability of locating sources of pollutants through a source tracing analysis would be valuable to a policy maker in finding areas ("hot spots") whose pollution contribution exceeds a set threshold. These hot spots can then be targeted for change. They can be mapped and this information feeds back to the dynamic landscape model so that it uses this knowledge to update the existing agricultural land use map by applying best management practices to the hot spot areas. For example, filter strips could be added near water bodies to prevent agricultural pollutants from entering water resources. The non-point source model is then re-run and the results are compared to the previous set in the LWIDSS (Figure 16). If the amount of pollution reduction is unacceptable, then the set of target areas are updated and the whole feed back mechanism between the non-point source model and the landscape model is repeated until an "optimal" solution for best management practices is found.

4. Conclusions

Environmental decision support systems have been implemented in Environmental Canada. They are proved to be practical. Not only the modellers increase their productivity with the assist of the EDSSs, government officials and policy makers also gain valuable insights by using these EDSSs to set policy guidelines and objectives.

The EEM-SAT DSS has been tested by the EEM National office as well as by external users and has proven to substantially improve the ability of the user to generate accurate and consistent analyses that are required under the Canadian EEM program. The responses from external users have been positive. A significant proportion of the time and effort spent on

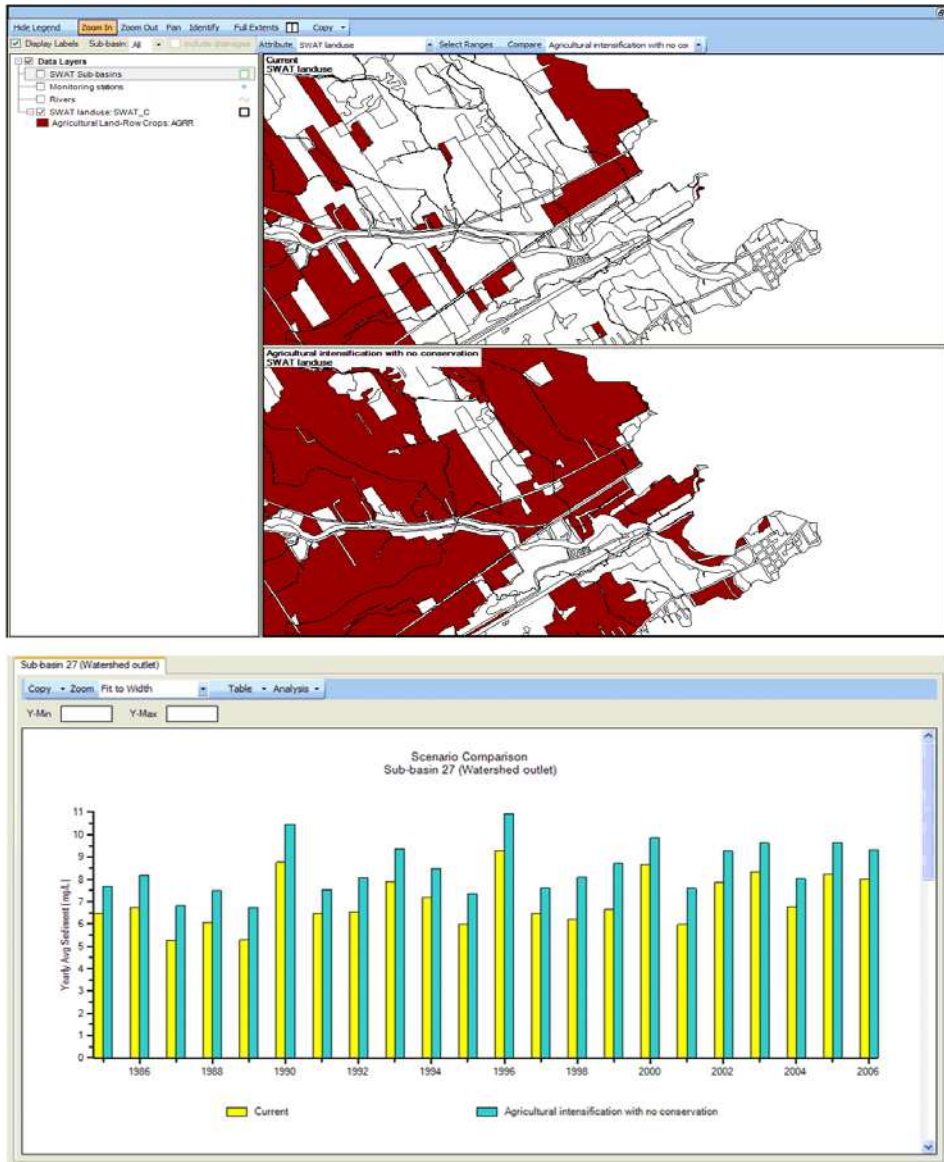


Fig. 15. Reviewing model input (top) and results (bottom) in the LWIDSS for the purposes of scenario comparison. Top: Maps comparing land use for the two scenarios (top map: current; bottom map: agricultural intensification). Shaded areas are agricultural lands. Bottom: a chart comparing sediment concentrations predicted at the watershed outlet between two land use scenarios is examined to investigate the impact of land use to water quality.

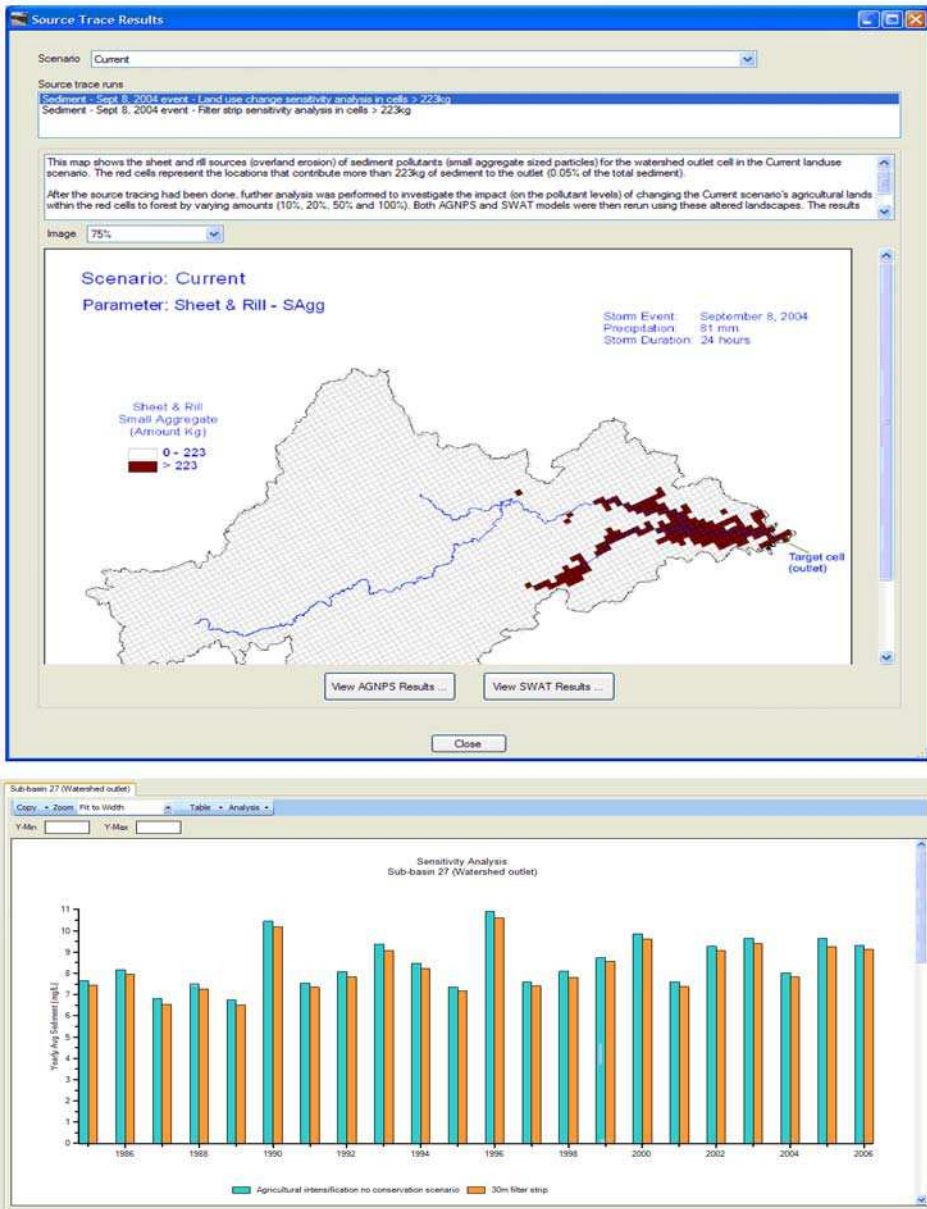


Fig. 16. Reviewing results from source tracing analysis. Top: map displaying locations in the watershed that contribute higher amounts of sediment to the watershed outlet is analyzed to find potential target areas for best management practices. Bottom: a chart comparing sediment concentrations before and after best management practices (application of filter strips).

developing the EEM-SAT DSS involved the data entry design, expert system rulebase for data preprocessing and QA/QC, expert system rulebase for logic control of the endpoint analyses for both fish and benthic invertebrates, as well as the rulebase for the site effects analyses and current status reporting.

The release of EEM-SAT DSS makes the statistics requirements of the Canadian EEM robust and transparent. It reduces the likelihood of incorrect data retrieval, improper data manipulation, misuse of statistics procedures and misinterpretation of the statistics results. With repeated beta testing and user usability analysis, the system has become a seamless and friendly user interface for its users based on sound science to ensure accuracy and effectiveness for users of all levels. The system addresses three key improvement areas that otherwise requires a considerable amount of time and effort by the user. These areas include proper data manipulation, the statistics analyses and results interpretation. Efforts were also made to ensure that the design would be generic enough to allow the system to work for other sector data analyses such as metal mining.

Designed for policy makers, the LWIDSS system is an effective analytical, planning and management tool to interpret and report modelling and scenario gaming results of the watershed-based modelling. Its framework can be used to determine the impact assessment of various land use scenarios on sediment and nutrients. The integrated decision support processes are captured in the LWIDSS and can be used for other similar watersheds if appropriate data is available. The results of the modelling and scenario gaming can be linked with other systems using a common data exchange interface.

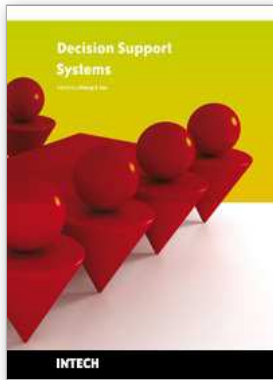
The design of the LWIDSS facilitates integration of diverse information, ranging from various land use scenario map layers, to soil texture map layer, to Digital Elevation Model data, and to water quality data. It is also designed to provide an opportunity to query a variety of databases, to visualize spatial and/or temporal patterns, and to analyze model input data and output results using the DSS customized tools. Inherent to the LWIDSS architecture are relational databases with common design structures for data integration to be used in the modelling and scenario gaming framework. The use of modelling and scenario gaming will allow decision-makers to explore potential responses of land and water integration to hypothetical situations, i.e. to answer the "what-if" question. The feedback loop among the models is important for the policy makers to explore the best possible management options and the course of action for pollution control.

We are continuing to strive to improve the development of future EDSSs. The use of automation and expert system rules to help guide modellers in the decision process will greatly reduce the uncertainty by using the appropriate tools and approaches. In addition, bringing the EDSSs to the web will allow timely update of models, data and information thus increases the usability of the EDSSs by removing data and model inconsistency and duplication.

5. References

- Alter, S.L. 1980. *Decision support systems: current practice and continuing challenges*, Addison-Wesley, Reading, MA.
- Arnold J.G., Srinivasan R., Muttiah R.S., Williams J.R. 1998. Large area hydrologic modeling and assessment. Part I, model development. *Journal of the American Water Resource Association* 34(1):73-89.

- Booty, W.G., Lam, D.C.L., Wong, I.W.S., and Siconolfi, P. 2001. Design and implementation of an environmental decision support system. *Environmental Modelling and Software* 16, pp. 453-458.
- Buchanan, B.G. and Shortliffe, E.H. 1985. *Rule-Based Expert Systems: The MYCIN Experiments of the Stanford Heuristic Programming Project*, Addison-Wesley, Reading, MA.
- Environment Canada. 2000a. Data Interpretation Guidance for Environmental Effects Monitoring. *National EEM Office*, Environment Canada, Ottawa, Ontario.
- Environment Canada 2002. Metal Mining EEM Guidance Document for Aquatic Environmental Effects Monitoring, *National EEM Office*, Science Policy and Environmental Quality Branch, Ottawa Canada.
- Ignizio, J.P. 1991. *Introduction to Expert Systems: The Development and Implementation of Rule-Based Expert Systems*, McGraw-Hill, 402 pp.
- Lam, D.C.L., Puckett, K.J., Wong, I., Moran, M.D., Fenech, G., Jeffries, D.S., Olson, M.P., Whelpdale, D.M., McNichol, D., Mariam, Y.K. and Minss, C.K. 1998. Anintegrated acid rain assessment model for Canada: from source emission to ecological impact, *Water Quality Res. J. Canada*, 33, 1-17.
- Lam, D.C.L., Leon, L.F., Hamilton, S., Crookshank, N., Bonin, D. and Swayne, D.A. (2004). Multi-Model Integration In A Decision Support System: A Technical User Interface Approach For Watershed And Lake Management Scenarios. *Environmental Modelling and Software* 19, pp. 317-324.
- Leon, L.F., Booty, W.G., Bowen, G.S., and Lam, D.C.L. 2004. Validation of an agricultural non-point source model in a watershed in southern Ontario, *Agricultural Water Management*, 65, 59-75.
- Lowell, R.B., Hedley, K, and Porter, E. 2002. Data Interpretation Issues for Canada's Environmental Effects Monitoring Program. *Water Qual. Research J. Canada* 37(1), pp.101-117.
- Nash, J.E., and Sutcliffe, J.V. 1970. River flow forecasting through conceptual models: Part I - A discussion of principles, *Journal of Hydrology*, 10, 282-290.
- SYSTAT 11 Statistics II, 2005. "General Linear Models", Pages 139-205, Richmond, CA
- Walker, S.L., Hedley, K, and Porter, E. 2002. Pulp and Paper Environmental Effects Monitoring in Canada: An Overview. *Water Qual. Research J. Canada* 37(1), pp. 7-19.
- Wong, I., Bloom, R., NcNicol, D.K., Fong, P., Russell, R. and Chen X. 2007. Species at risk: Data and knowledge management within the WILDSPACE decision support system, *Environmental Modelling and Software*, 22, pp. 423-430.
- Young, R., Onstad, C., Bosch, D. and Anderson, W. 1987. *Agricultural non-point source pollution model: A watershed analysis tool*, Conservation Research Report 35, Agricultural Research Service, U.S. Department of Agriculture, Washington, D.C.



Decision Support Systems

Edited by Chiang S. Jao

ISBN 978-953-7619-64-0

Hard cover, 406 pages

Publisher InTech

Published online 01, January, 2010

Published in print edition January, 2010

Decision support systems (DSS) have evolved over the past four decades from theoretical concepts into real world computerized applications. DSS architecture contains three key components: knowledge base, computerized model, and user interface. DSS simulate cognitive decision-making functions of humans based on artificial intelligence methodologies (including expert systems, data mining, machine learning, connectionism, logistical reasoning, etc.) in order to perform decision support functions. The applications of DSS cover many domains, ranging from aviation monitoring, transportation safety, clinical diagnosis, weather forecast, business management to internet search strategy. By combining knowledge bases with inference rules, DSS are able to provide suggestions to end users to improve decisions and outcomes. This book is written as a textbook so that it can be used in formal courses examining decision support systems. It may be used by both undergraduate and graduate students from diverse computer-related fields. It will also be of value to established professionals as a text for self-study or for reference.

How to reference

In order to correctly reference this scholarly work, feel free to copy and paste the following:

William Booty and Isaac Wong (2010). Case Studies of Canadian Environmental Decision Support Systems, Decision Support Systems, Chiang S. Jao (Ed.), ISBN: 978-953-7619-64-0, InTech, Available from: <http://www.intechopen.com/books/decision-support-systems/case-studies-of-canadian-environmental-decision-support-systems>

INTECH
open science | open minds

InTech Europe

University Campus STeP Ri
Slavka Krautzeka 83/A
51000 Rijeka, Croatia
Phone: +385 (51) 770 447
Fax: +385 (51) 686 166
www.intechopen.com

InTech China

Unit 405, Office Block, Hotel Equatorial Shanghai
No.65, Yan An Road (West), Shanghai, 200040, China
中国上海市延安西路65号上海国际贵都大饭店办公楼405单元
Phone: +86-21-62489820
Fax: +86-21-62489821

© 2010 The Author(s). Licensee IntechOpen. This chapter is distributed under the terms of the [Creative Commons Attribution-NonCommercial-ShareAlike-3.0 License](#), which permits use, distribution and reproduction for non-commercial purposes, provided the original is properly cited and derivative works building on this content are distributed under the same license.