

Combining Scientific Facts and Significance Criteria to Predict the Result of an Environmental Impact Assessment Review

K. F. R. Liu^{1,*}, H. H. Liang², C. W. Chen³, J. S. Chen¹, and Y. S. Shen⁴

¹Department of Safety, Health and Environmental Engineering, Ming Chi University of Technology, New Taipei City, Taiwan 24301, ROC

²Department of Architecture, National United University, Miaoli, Taiwan 36003, ROC

³Institute of Maritime Information and Technology, National Kaohsiung Marine University, Kaohsiung 80543, Taiwan, ROC

⁴Holistic Education Center, Mackay Medical College, New Taipei City, Taiwan 25245, ROC

Received 25 July 2011; revised 25 January 2012; accepted 9 April 2012; published online 22 June 2012

ABSTRACT. The purpose of determining impact significance is to place value on impacts. Environmental impact assessment review is a process that judges whether impact significance is acceptable or not in accordance with the scientific facts regarding environmental, ecological and socio-economical impacts described in environmental impact statements or environmental impact assessment reports. The first aim of this paper is to summarize the criteria of significance evaluation from the past review results and accordingly utilize fuzzy rule-based system to incorporate these criteria into scientific facts. The second aim is to employ data mining technique to construct an environmental impact statement or environmental impact assessment report prediction model for reviewing results which can assist developers to prepare and revise better environmental management plans in advance. The validity of the previous prediction model proposed by authors in 2009 is 92.7%. The enhanced validity in this study can attain 100.0% after taking significance criteria into account. In the case study, some indicators are over-standard in first and second-order significance but all indicators ultimately fall into the category of conditional approval in third-order significance.

Keywords: environmental impact assessment review, impact significance, fuzzy rule-based system, data mining, classification tree

1. Introduction

Environmental impact assessment (EIA) stands for an environmental management plan based on scientific, objective and comprehensive surveys, forecasting, analyses and evaluations conducted prior to project implementation in order to determine the degree and scope of the potential impact of development activity or government policy on the environment, society, economy, culture and ecology, and the public explanation and review of such a plan. In Taiwan, development projects for which there is concern of adverse impact on the environment should prepare an environmental impact statement for the phase-I EIA, and then transfer the environmental impact statement to the competent authority for review. The developer should edit an environmental impact assessment report for the phase-II EIA for those circumstances in which the review result of the environmental impact statement is concerned with a significant impact on the environment. The review results of environmental impact statements or environmental impact assessment reports can be classified into three categories: conditional approval, phase-II EIA, or dis-

approval on the development project. A major concern for developers is the possible review results of projects. The aim of this paper is to assist developers with predictions of review results in order to facilitate early preparation and all them to devise better environmental management plans which can ameliorate highly risky nuisances in advance; thereby, the probability of passing review is enhanced.

The prediction of a review results is feasible if there are sufficient previous review cases. From 1981 to present, the Taiwan Environmental Protection Administration has been collecting environmental impact statements and environmental impact assessment reports and their associated review results and opinions, and disclosing these documents to the public on an online basis. Based on these past cases, the authors and their colleagues have proposed an integrated prediction model consisting of case-based reasoning and fuzzy rule-based system to qualitatively forecast possible review results, which presents an overall prediction validity of 92.7% (Liu and Yu, 2009b). Our previous attempt was unable to obtain higher validity because the model is based on past environmental impact statements and environmental impact statements. In fact, the environmental information provided in environmental impact statements and environmental impact statements is primarily related to scientific facts (magnitudes of impacts induced by a development project); however, the review process can be viewed as highly subjective judgment because it has to ruminate over the scientific facts and subjective values (judgment,

* Corresponding author. Tel.: +886 2 29089899 4697; fax: +886 2 29041914.
E-mail address: kevinliu@mail.mcut.edu.tw (K. F. R. Liu).

preference, value and concern) (Liu et al., 2010). More specifically, review work determines the significance of impacts (Beattie, 1995; Kontic, 2000; Wilkins, 2003; Lawrence, 2007; Wood et al., 2007; Liu et al., 2010).

It is true that synthesizing several techniques can produce effective hybrid systems by reaping the advantages from each component method. The combination of (two or more) different techniques is a very active research area in artificial intelligence (Prentzas and Hatzilygeroudis, 2007). This paper intends to integrate two artificial intelligence techniques to predict EIA review results with higher validity as described below.

- Fuzzy rule-based system for evaluating impact significance. In Taiwan, EIA review committees composed of experts and scholars carry out the determination of impact significance and accordingly conduct the review results. In other words, the significance of environmental impacts is subjectively appraised by experts and scholars. In artificial intelligence, fuzzy logic (Zadeh, 1996) is a mechanism with the ability of computing with words to mimic human reasoning. Therefore, this paper intends to utilize fuzzy rule-based system to model experts' judgment in EIA reviews.

- Significance transformation for incorporating significance thresholds. Thresholds can be described in terms of legal requirements (e.g. pollution standards), ecological and social-economic tolerance standards. They can provide a means of making policy and legislation more rational, predictable and scientific by examining the effects of a project with reference to thresholds (Haug et al., 1984). The transformation divides the values resulting from fuzzy rule-based systems by thresholds to produce significance scores which are easy to interpret: an impact is unacceptable when its significance is larger than 100%.

- Data mining for predicting EIA review results. This paper intends to use data mining technique to foresee three possible review results for a development project with higher validity. This is because the technique is capable of inducing classification knowledge through a learning mechanism.

The validity of the previous prediction model proposed by authors in 2009 is 92.7%. The enhanced validity in this study can attain 100.0% after taking significance criteria into account. The organization of this paper is presented as follows. In Section 2, we review background and methods of applying fuzzy rule-based system and data mining to environmental fields, especially to EIA. Section 3 discusses environmental factors and their significance criteria for road construction. In addition, the proposed prediction model, which integrates fuzzy rule-based system for evaluating impact significance and data mining for forecasting EIA review results, is introduced and verified in Section 3. In Section 4, a case study is used to illustrate this model and, finally, conclusions and future work are delineated in Section 5.

2. Background Literature

This paper attempts to take advantages of artificial inte-

lligence to help developers. In environmental areas, artificial intelligence has become an important discipline because its branch techniques such as expert systems, artificial neural networks, fuzzy logic, Bayesian networks, evolutionary algorithms, data mining, etc. have been successfully applied to environmental studies. For example, several researchers have been endeavoring to use fuzzy logic methods to infer environmental impacts or significances. For instance, Borri et al. (1998) introduced a fuzzy rule-based methodology for environmental evaluation which provided a robust tool to directly cope with linguistic models of human interpretation of environmental systems. Van der Werf and Zimmer (1998), as well as Roussel et al. (2000), endeavored to use fuzzy expert systems to calculate an indicator 'IPEST' which reflects an expert perception of the potential environmental impact of the application of a pesticide in a crop field. Enea and Salemi (2001) developed a matrix method to calculate the total environmental impact and the percentage of impact on every environmental component in fuzzy form. Phillis et al. (2001) and Andriantiatsaholiniaina et al. (2004) used fuzzy logic reasoning and basic indicators of environmental integrity, economic efficiency, and social welfare to derive measures of human, ecological, and overall sustainability. González et al. (2002) utilized fuzzy logic to avoid the need for in-depth environmental knowledge and extremely accurate data to perform assessment, thus making life-cycle assessment more applicable to small and medium-sized enterprises. Smith (2002) used fuzzy terms to express the performance and importance of impacts and applied the terms to assess road infrastructure projects. de Siqueira Campos Boclin and de Mello (2006) used a fuzzy logic computational approach to operate fuzzy and crisp variables and make inferences from resultant values of the systemic indicator as well as environmental, cultural, social and economic thematic indicators. Wood et al. (2007) explored the potential of fuzzy-set theory and simulations of environmental change for delineating evaluations of impact significance made by a range of stakeholders. Liu et al. (2007, 2009a, 2010) combined fuzzy logic with other multi-criteria decision analysis such as fuzzy analytic hierarchy process and fuzzy analytic network process to appraise environmental impact. Morón et al. (2009) presented a new fuzzy EIA model which incorporates both qualitative and quantitative information. Peche and Rodríguez (2009, 2011) used fuzzy weighted summation to consider positive and negative impacts at the same time in EIA. Lv et al. (2010) proposed an interval fuzzy bi-level programming approach for planning water resources management system. Tennakoon et al. (2010) used fuzzy logic programming for indoor air and temperature quality monitoring and a fuzzy inference system for hazard detection.

Ekasingh et al. (2005) proposed a data mining approach to simulating farmers' crop choices for integrated water resource management. Le Ber et al. (2006) used a data mining software to study crop sequences and to understand the recent changes of land uses and to forecast the future of new land uses. Dixon et al. (2007) incorporated inducted rules into the supervision of an anaerobic wastewater treatment process. Ekasingh and Ngamsomsuke (2009) used the C4.5 data mining

Table 1. Criteria and Decision Techniques Used to Evaluate Impact Significance

Authors	Scales for significance	Significance criteria	Scales for criteria	Evaluation techniques
Duinker and Beanlands (1986)	Major Moderate Minor Negligible	Magnitude Importance Distribution of change in time Distribution of change in space Prediction reliability	Not available	Direct expert assessment
Bojórquez-Tapia et al. (1998)	Very high (0.75 - 1.00) High (0.50 - 0.75) Moderate (0.25 - 0.50) Low (0 - 0.25)	Magnitude Spatial extent Duration Synergy Cumulative effects Controversy Mitigation measures	0 - 9	Nonlinearly weighted sum
Antunes et al. (2001)	Very significant negative impact (-5) ~ Very significant positive impact (+5)	Affected area Affected population Affected sensitive ecosystems	0 - 10	Linearly weighted sum
Cloquell-Ballester et al. (2007)	Very high (0.75-1.00) High (0.50-0.75) Low (0.25-0.75) Very low (0-0.25)	Magnitude Spatial extent Duration Synergy Cumulative effects Controversy Mitigation measures	0 - 9	Linearly weighted sum Nonlinearly weighted sum Electre TRI Expert assessment
Wood (2008)	Very substantial Substantial Moderate Slight None	Magnitude Sensitivity/Importance	High Medium Low	Decision matrix

algorithm to produce simplified models of farmers' crop choices in two watersheds in Thailand. Gibert et al. (2010) used clustering based on rules by states to permit qualitative analysis of the dynamics of a wastewater treatment plant. Gross et al. (2010) developed Enchilada (Environmental Chemistry through Intelligent Atmospheric Data Analysis) to carry out a variety of clustering methods on mass-spectral data. Birant (2011) compared many decision tree algorithms for predicting potential air pollutant emissions with data mining models and found C4.5 algorithm has the highest accuracy value.

3. Materials and Methods

3.1. Environmental Factors and Their Significance Criteria

The three steps of EIA can be summarized as follows (Rossouw, 2003): (1) identify the potential impacts of a development project on the environment; (2) quantify and predict the likely nature of such impacts; and (3) evaluate the significance of the potential impacts. The last step, evaluation of impact significance, is the core of EIA review - the EIA review committee makes judgments about which impacts are important, undesirable or unacceptable. Despite of its importance, impact significance is one of the most complex and least-understood concepts in EIA. It is sometimes directly linked to the concept of impact magnitude (Wood et al., 2007) or the concept of importance (Antunes et al., 2001; Peche and Rodriguez, 2009, 2011), or their combination (UNU et al., 2010);

on the other hand, it can also be related to a set of criteria (Cloquell-Ballester et al., 2007; Wood, 2008).

Impact significance places value on impact. In this paper, impact significance is defined as the effects of impacts on what stakeholders are concerned about. Evaluation of impact significance is a complicated decision-making process because it is subjective, political, normative, value-dependent and often controversial (Lawrence, 2007). The complexities of impact significance determinations spur the developments of criteria and evaluation techniques. As shown in Table 1, the significance criteria include magnitude, importance, sensitivity/affected ecosystems or resources, spatial extent/affected area or population, duration/persistence, synergy, cumulative effects, controversy, mitigation measures, prediction reliability, etc., and the evaluation techniques are comprised of direct expert assessment, linearly or nonlinearly weighted sums, Electre TRI, decision matrix, etc.

The indicators considered in EIA for road construction projects, as listed in Table 2 (Liu et al., 2009b), contains air (I1), water (I2), soil (I3), solid waste (I4), noise (I5), terrestrial (I6), aquatic (I7), economics (I8), society (I9) and culture (I10); the subindicators include: carbon monoxide (I11), sulfur dioxide (I12), nitrogen dioxide (I13), total suspended particulates (I14), dissolved oxygen (I21), biochemical oxygen demand (I22), suspended solids (I23) and ammonia nitrogen (I24), heavy-metal pollution (I31), noise (I41), vibration (I42), rubbish (I51), industrial waste (I52), terrestrial animals (I61), terrestrial plants (I62), terrestrial endangered species (I63),

Table 2. Factors for the Assessment of Environmental Impact and Information for the Case Study

Indicator	Sub-indicator	Unit	SV ^a	Study Case		
				BC ^b	PIWOM ^c	PIWM ^d
Air (I ₁)	CO (I ₁₁)	ppm	35.0	0.7	0.9	0.8
	SO ₂ (I ₁₂)	ppb	250.0	6.8	39.5	16.6
	NO ₂ (I ₁₃)	ppb	250.0	24.5	94.6	66.5
	TSP (I ₁₄)	g/m ³	250.0	71.8	136.6	97.7
Water (I ₂)	DO (I ₂₁)	mg/L	6.5	7.1	7.1	7.1
	BOD (I ₂₂)	mg/L	3.0	4.1	4.1	4.1
	SS (I ₂₃)	mg/L	20.0	5.3	5.3	5.3
	NH ₃ -N (I ₂₄)	mg/L	0.5	4.7	4.7	4.7
Soil (I ₃)	Heavy metal (I ₃₁)	0-100	100.0	66.9	66.9	66.9
Solid waste (I ₄)	Rubbish (I ₄₁)	0-100	20.0	4.0	9.0	5.0
	Construction waste (I ₄₂)	0-100	30.0	0.0	33.0	3.0
Noise (I ₅)	Noise (I ₅₁)	dB	60.0	59.8	71.5	68.0
	Vibration (I ₅₂)	dB	60.0	33.2	60.1	52.0
Terrestrial (I ₆)	Threatened terrestrial animals (I ₆₁)	%	20.0	3.0	20.0	10.0
	Threatened terrestrial plants (I ₆₂)	%	20.0	3.0	20.0	10.0
	Threatened endangered terrestrial species (I ₆₃)	%	10.0	1.0	3.0	2.0
Aquatic (I ₇)	Threatened aquatic animals (I ₇₁)	%	20.0	15.0	20.0	18.0
	Threatened aquatic plants (I ₇₂)	%	20.0	10.0	15.0	12.0
	Threatened endangered aquatic species (I ₇₃)	%	10.0	0.0	0.0	0.0
Economics (I ₈)	Land-use and development obstacle (I ₈₁)	0-100	20.0	2.0	3.0	3.0
	Life-quality decline (I ₈₂)	0-100	10.0	2.0	3.0	3.0
	Economic activity disturbance (I ₈₃)	0-100	10.0	0.0	1.0	1.0
Society (I ₉)	Public facility inaccessibility (I ₉₁)	0-100	20.0	0.0	0.0	0.0
	Transportation inaccessibility (I ₉₂)	0-100	50.0	10.0	50.0	30.0
	Community disconnection (I ₉₃)	0-100	20.0	5.0	10.0	6.0
Culture (I ₁₀)	Cultural heritage destruction (I ₁₀₁)	0-100	20.0	0.0	0.0	0.0
	Landscape demolition (I ₁₀₂)	0-100	20.0	3.0	10.0	8.0

* ^a SV: standard value; ^b BC: baseline condition; ^c PIWOM: predicted pollution increment without mitigation measures; ^d PIWM: predicted pollution increment with mitigation measures.

aquatic animals (I₇₁), aquatic plants (I₇₂), aquatic endangered species (I₇₃), land-use and development (I₈₁), life quality (I₈₂), economic activities (I₈₃), inaccessibilities in public facilities (I₉₁) and transportation (I₉₂), disconnection in communities (I₉₃), cultural heritage (I₁₀₁) and landscapes (I₁₀₂). Some of these subindicators, such as I₁₁, I₁₂, I₁₃, I₁₄, I₂₁, I₂₂, I₂₃, I₂₄, I₃₁, I₅₁ and I₅₂ can be objectively measured by instruments. The rest are subjectively estimated by experts on the basis of information collected. To reduce the subjectivity in the assessment and provide a replicable and explicit framework, the scoring guideline is predefined for some subindicators, as shown in Table 3. All values of these subindicators are precise numbers with precision to the nearest tenth and their respective units of measurement and standard value (SV) are shown in the third and fourth columns of Table 2. It should be noted that the purpose of the proposed model is to help developers identify and deal with risky nuisances to the environment in their development projects, and then enable them to pass EIA reviews, rather than to help EIA review to make better decision-making. That's why positive impacts were not included in this study. As shown in Table 2, the descriptions of the subindicators tend to negative impacts.

Sensitivity, spatial extent, mitigation measure reliability and information integrity are four major significance criteria

used by the EIA review committee for reviewing road construction projects. These criteria are extracted from the EIA review opinions of 63 real cases from 1994 to 2009. The four criteria are discussed 28, 3, 43 and 84 times respectively within the period and their distributions over the ten indicators and their standard values (SV) are also shown in Table 4. Sensitivity refers to whether or not a project is located in air pollution control regions, water pollution control areas, drinking water source protection areas, a certain distance from drinking water intake points, the water catchment area of a dam or reservoir, soil pollution control sites, noise control zone, aquatic plant and animal breeding conservation areas, fishing areas, planned urban protection zones, heritage conservation areas, special landscape protection areas, etc. Spatial extent refers to the distribution of significant impacts. Although spatial extent was only used three times by the EIA review committee in Taiwan, it is still included as a criterion because it was suggested in most of the related studies, as shown in Table 1. Mitigation measure reliability represents the confidence in the effectiveness of proposed mitigation actions. Information integrity means the completeness of information that can help the committee make judgments. The rating guidelines for all indicators are defined in advance in order to clarify the rating process and an example of air pollution is illustrated in Table 5.

Table 3. Scoring Guidelines for Some Sub-Indicators and Scores for the Case Study

Sub-indicator	Info in EIA reports	Scoring guideline					Scores of case study (BC;PIWOM;PIWM)
		0 - 20	20 - 40	40 - 60	60 - 80	80 - 100	
I ₄₁	ton	0 - 0.5	0.5 - 1.0	1.0 - 1.5	1.5 - 2.0	2.0 - 2.5	0.01; 0.21; 0.02
I ₄₂	km ³	0.0 - 3.0	3.0 - 6.0	6.0 - 9.0	9.0 - 12.0	12.0 - 15.0	0.00; 4.99; 0.60
I ₆₁	Number	Numbers of threatened terrestrial animals/Numbers of total terrestrial animals×100*					3.0; 20.0; 10.0
I ₆₂	Number	Numbers of threatened terrestrial plants/Numbers of total terrestrial plants×100*					3.0; 20.0; 10.0
I ₆₃	Description	Subjective determination on degree (%) of impact on endangered terrestrial species*: (Slight One-third About half Severe Very severe)					1.0; 3.0; 2.0
I ₇₁	Number	Numbers of threatened aquatic animals/Numbers of total aquatic animals×100*					15.0; 20.0; 18.0
I ₇₂	Number	Numbers of threatened aquatic plants/Numbers of total aquatic plants×100*					10.0; 15.0; 12.0
I ₇₃	Description	Subjective determination on degree (%) of impact on endangered aquatic species*: (Slight One-third About half Severe Very severe)					0.0; 0.0; 0.0
I ₈₁	Description	Subjective determination on degree (%) of impact on land-use and development*: (Slight Low Moderate High Severe)					2.0; 3.0; 3.0
I ₈₂	Description	Subjective determination on degree (%) of impact on life-quality: (Slight Low Moderate High Severe)					2.0; 3.0; 3.0
I ₈₃	Description	Subjective determination on degree (%) of impact on economic activity: (Slight One-third About half Severe Very severe)					0.0; 1.0; 1.0
I ₉₁	Description	Subjective determination on degree (%) of impact on public facility accessibility: (Slight One-third About half Severe Very severe)					0.0; 0.0; 0.0
I ₉₂	Description	Subjective determination on degree (%) of impact on transportation accessibility: (Slight One-third About half Severe Very severe)					10.0; 50.0; 30.0
I ₉₃	Description	Subjective determination on degree (%) of impact on community connection: (Slight One-third About half Severe Very severe)					5.0; 10.0; 6.0
I ₁₀₁	Description	Subjective determination on degree (%) of impact on cultural heritage: (Slight One-third About half Severe Very severe)					0.0; 0.0; 0.0
I ₁₀₂	Description	Subjective determination on degree (%) of impact on landscape: (Slight One-third About half Severe Very severe)					3.0; 10.0; 8.0

Note: Rubbish (I₄₁), Construction waste (I₄₂), Threatened terrestrial animals (I₆₁), Threatened terrestrial plants (I₆₂), Threatened endangered terrestrial species (I₆₃), Threatened aquatic animals (I₇₁), Threatened aquatic plants (I₇₂), Threatened endangered aquatic species (I₇₃), Land-use and development obstacle (I₈₁), Life-quality decline (I₈₂), Economic activity disturbance (I₈₃), Public facility inaccessibility (I₉₁), Transportation inaccessibility (I₉₂), Community disconnection (I₉₃), Cultural heritage destruction (I₁₀₁), Landscape demolition (I₁₀₂).

* The score can be appropriately increased if some threatened species play key roles in food web.

3.2. Overall Prediction Model

The concept of impact significance underlies an integrated prediction model of EIA review results of road construction projects. Furthermore, the example of air pollution is depicted in Figure 1. As discussed in previous section, significance is a complex concept that relates to not only impact magnitude but also other considerations and its evaluation may be viewed as a highly subjective judgment. Therefore, in part (a) of Figure 1, fuzzy rule-based systems (FRBS) are employed to infer significances because they can imitate experts' thinking process. FRBSs are used to model experts' judgment in EIA reviews, including taking significance criteria into account. To better interpret the outputs of fuzzy rule-based systems, all outputs are transformed into impact significance by dividing with the outputs derived from their respective standard values. This step is termed significance transformation (ST). It should be noted that the value resulting from significance transformation is expressed as a percentage of thresholds (standard values), ranging from 0 (insignificant) to 200% (completely significant), for the sake of incorporating legal requirements, ecological and social-economic tolerance standards. The maximum of significance transformation, 200%, is a reasonable calibra-

tion because the outputs resulting from significance transformation in most of the cases are below this value. The determination of impact significances is developed into three tiers. The first tier produces first-order significance which mainly evaluates the impact magnitudes of an indicator according to the pollution levels of its subindicators. For example, as shown in Figure 1, air pollution evaluation refers to the appraisal of emission of carbon monoxide (CO), sulfur dioxide (SO₂), nitrogen dioxide (NO₂) and total suspended particulates (TSP). Second-order significance is an outcome of simultaneous consideration of baseline condition (BC), predicted pollution increment without and with mitigation measures (PIWOM and PIWM). The first and second-order significance are evaluated based on the objective measurements or predictions; however, the last tier, termed third-order significance, merges significance criteria which are extracted from the opinions given in EIA reviews. The use of fuzzy rule-based system to estimate the impact significances of the indicators is outlined in Section 3.3.

The second part of the model predicts EIA review results according to impact significance. A data mining technique is adopted for the task of prediction. First of all, the third-order

Table 4. Significance Criteria and Their Counts in the Conclusions of the EIA Reviews of 63 Road Construction Projects, from 1994 to 2009

	Significance criteria							
	Sensitivity		SE		MML		Info Integrity	
	Count	SV	Count	SV	Count	SV	Count	SV
I ₁	2	30	1	20	3	70	6	70
I ₂	9	30	0	20	8	70	9	70
I ₃	1	NA	0	0	3	70	3	70
I ₄	1	NA	0	NA	9	70	17	70
I ₅	5	25	1	20	5	70	8	70
I ₆	3	0	0	10	3	70	9	70
I ₇	4	0	0	10	2	70	9	70
I ₈	1	0	0	NA	2	70	3	70
I ₉	1	NA	0	NA	3	70	14	70
I ₁₀	1	0	1	NA	5	70	6	70
Total	28		3		43		84	

Note: Indicator: Air (I₁), Water (I₂), Soil (I₃), Solid waste (I₄), Noise (I₅), Terrestrial (I₆), Aquatic (I₇), Economics (I₈), Society (I₉), Culture (I₁₀); SE: Spatial extent, MML: Mitigation Measure Liability.

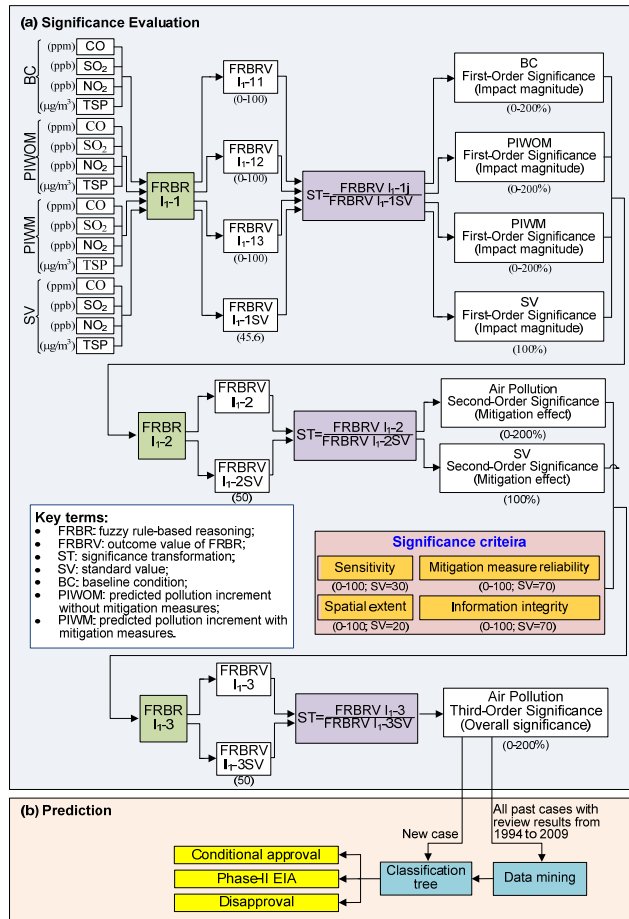


Figure 1. Integrated prediction model for the results of EIA reviews of road construction projects (air pollution).

significances of all EIA cases from 1992 to 2009 (i.e. 50 cases gained conditional approval, 4 cases were phase-II and ano-

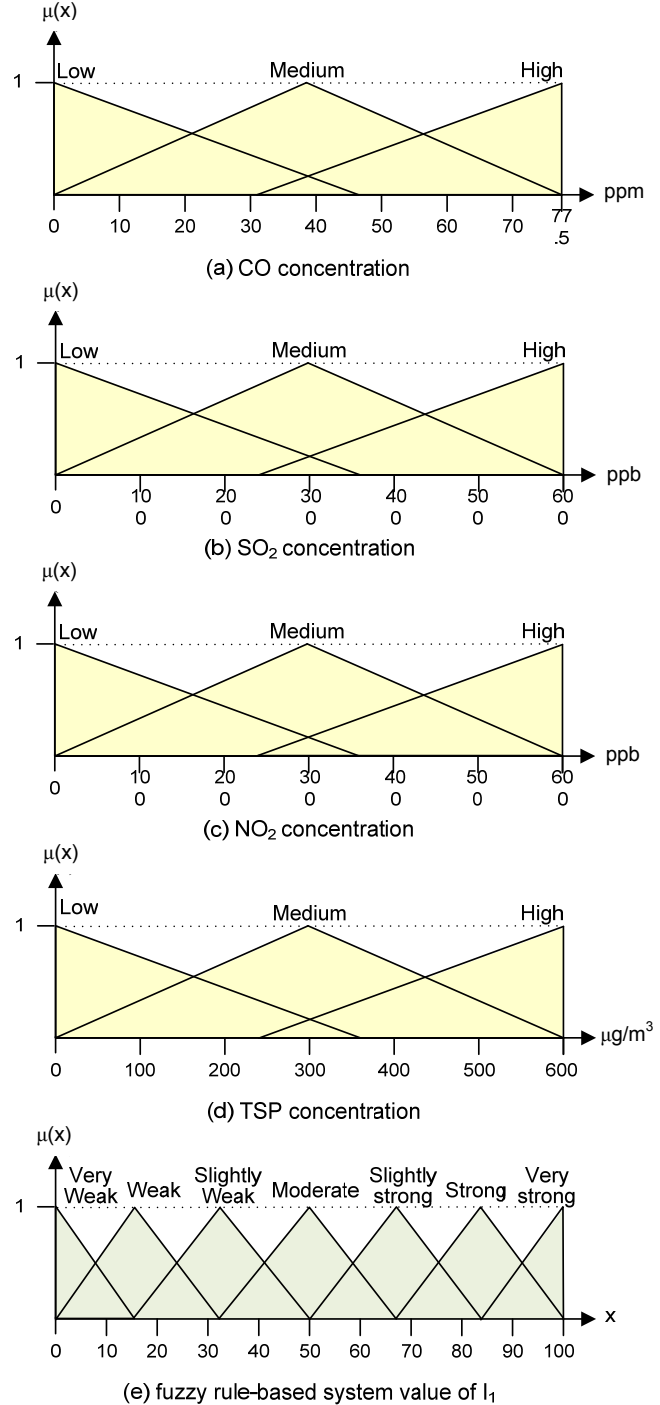


Figure 2. Membership functions of fuzzy values for linguistic variables (a) CO concentration, (b) SO₂ concentration, (c) NO₂ concentration, (d) TSP concentration and (e) fuzzy rule-based system value of I₁.

ther 9 cases gained disapproval) are used to derive a classification tree. This classification tree is then exploited to predict the possible review result for a new case, as shown in part (b) of Figure 1. The use of data mining to construct a classification tree for the prediction task is introduced in Section 3.4.

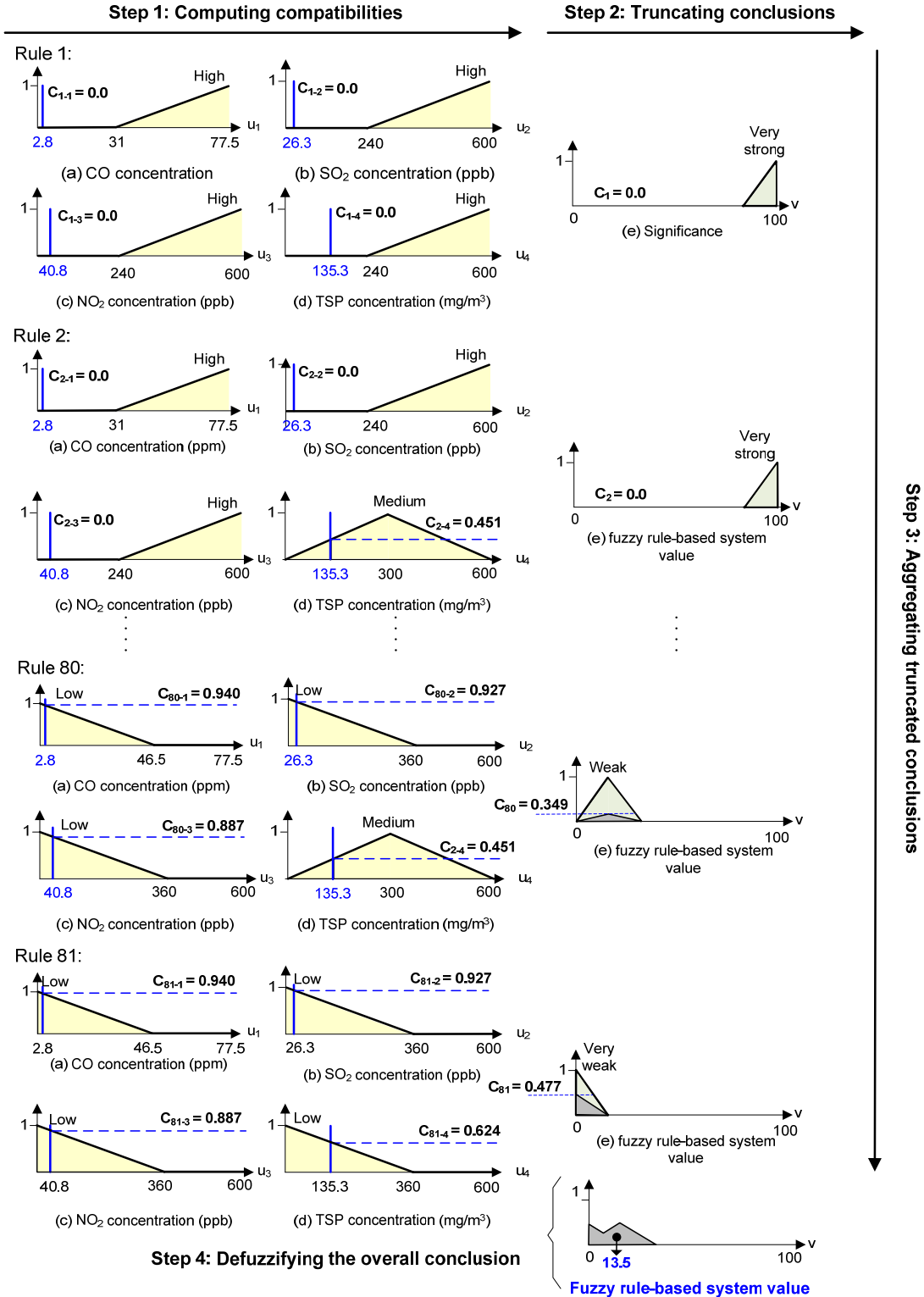


Figure 3. Graphical representation of fuzzy rule-based reasoning.

3.3. Fuzzy Rule-Based System (FRBS) for Evaluating Impact Significance

Fuzzy rule-based system (Zadeh, 1975) can be treated as a tool with the ability to compute with words for modeling

qualitative human thought processes in the analysis of complex systems and decisions. In fuzzy rule-based system, qualitative perception-based reasoning is represented by 'IF-THEN' fuzzy rules. The rule set of FRBS II-1 in Figure 1, concerning

the first-order significance of air pollution (I1) can be exemplified as:

Rule 1: IF CO concentration is high AND SO₂ concentration is high AND NO₂ concentration is high AND TSP concentration is high THEN the fuzzy rule-based system value of I1 is very strong.

Rule 2: IF CO concentration is high AND SO₂ concentration is high AND NO₂ concentration is high AND TSP concentration is medium THEN the fuzzy rule-based system value of I1 is strong.

...

Rule 80: IF CO concentration is low AND SO₂ concentration is low AND NO₂ concentration is low AND TSP concentration is medium THEN the fuzzy rule-based system value of I1 is weak.

Rule 81: IF CO concentration is low AND SO₂ concentration is low AND NO₂ concentration is low AND TSP concentration is low THEN the fuzzy rule-based system value of I1 is very weak.

where 'CO concentration,' 'SO₂ concentration,' 'NO₂ concentration,' 'TSP concentration' and 'fuzzy rule-based system value of I1' are linguistic variables; 'high,' 'medium,' 'low,' 'very strong,' 'strong,' 'weak' and 'very weak' are their possible fuzzy values, as defined by membership functions in fuzzy set theory (as shown in Figure 2).

For example, when four baseline conditions extracted from the environmental statement of the Dongshih-Chiayi line of the east-west expressway construction projects in the Taiwan western corridor (i.e., Fact 1: CO concentration is 2.8 ppm; Fact 2: SO₂ concentration is 26.3 ppb; Fact 3: NO₂ concentration is 40.8 ppb; Fact 4: TSP concentration is 135.3 g/m³) are fed into this inference mechanism (FRBS I1-1), fuzzy rule-based system proceeds. The theory of fuzzy rule-based system can be easily explained by a graphical representation as shown in Figure 3. In this figure, the four major steps in reaching a conclusion using fuzzy reasoning are described as follows.

Step 1: Computing compatibilities. Compatibility designates the similarity of an antecedent referring to a fact with the same linguistic variable or the suitability of a specific rule regarding several facts corresponding to the respective antecedents. For rule 80, the compatibility of Fact 1 with 'CO concentration is low' is 0.940; for Fact 2 with 'SO₂ concentration is low,' compatibility is 0.927; for Fact 3 with 'NO₂ concentration is low,' it is 0.887; for Fact 4 with 'TSP concentration is medium,' it is 0.451. It should be noted that 'product' is chosen as the t-norm operator instead of another more widely used t-norm operator, 'min,' because the t-norm operator 'product' makes the conclusion sensitive to every input; whereas, only one input controls the conclusion in the case of the t-norm operator 'min.' The overall compatibility of Rule 80 with the four facts is computed by $0.940 \times 0.927 \times 0.887 \times 0.451$, thereby obtaining 0.349. Similarly, the compatibilities of Rules 81 with the same facts are 0.477. The compatibilities of other rules are also calculated in the same way.

Step 2: Scaling conclusions. Once the compatibility for

each rule has been calculated, the degree to which the antecedents have been satisfied for each rule is known. The implication operator in fuzzy rule-based system is set as 'product' in this paper and therefore, as shown in Figure 3, a triangle conclusion is inferred by scaling the triangular conclusion of each rule with its corresponding compatibility.

Step 3: Aggregating scaling conclusions. Several inferred conclusions with the same linguistic variable should be aggregated. Aggregation is the process by which the fuzzy sets representing the scaled conclusions of triggered rules are combined into a single fuzzy set. In Figure 3, the final conclusion is aggregated by using the union of all scaled conclusions.

Step 4: Defuzzifying the overall conclusion. In many cases, the final output of an inference system should be a single number. Defuzzification is a method to justifiably convert a fuzzy set into a precise value. This study utilized the center-of-gravity method, which takes the center of the area under the curve of the membership function of a fuzzy set as the answer. Figure 3 indicates that the score of the fuzzy rule-based system value for the air baseline condition is 13.5.

For evaluating the significance of air pollution (I1), three fuzzy rule-based systems are required: FRBSs I1-1, I1-2 and I1-3. Because FRBS I1-2 is a common fuzzy rule-based system, 21 FRBSs containing 2,302 fuzzy rules were developed for evaluating the ten indicators and they are implemented with MATLAB Fuzzy Logic Toolbox.

3.4. Significance Transformation (ST) for Incorporating Significance Thresholds

A significance threshold represents the point at which the significance of an impact is unacceptable; that is, the impact is significant. This paper sets these thresholds in terms of legal requirements (e.g. pollution standards), ecological and social-economic tolerance standards. For example, the regulatory standards of CO, SO₂, NO₂ and TSP in Taiwan are 35.0 ppm, 250.0 ppb, 250.0 ppb and 250.0 g/m³, respectively; hence fuzzy rule-based system (FRBS I1-1) infers a value of 45.6. The process of significance transformation (ST) divides the value of 13.5 by 45.6 to produce the first-order significance value of 29.7%. The predicted concentrations of CO, SO₂, NO₂ and TSP in construction are 2.8 ppm, 30.6 ppb, 60.3 ppb and 836.7 g/m³, respectively; the first-order significance is 43.6% (19.9/45.6). With mitigation measures, the concentrations of the four subindicators CO, SO₂, NO₂ and TSP can be decreed to 2.8 ppm, 28.8 ppb, 46.6 ppb and 219.0 g/m³, respectively, and the first-order significance in this situation becomes 40.9% (18.7/45.6).

Similarly, 27 fuzzy rules for the second-order significance of air pollution (FRBS I1-2) are developed and infer the severity of 33.1 for the example case and thereby its second-order significance is 66.2%. This case is located in a class-II air pollution control region so its sensitivity is estimated at 30.0, according to Table 5. The spatial extent is 9.4 because six out of 64 measurements are beyond the standard values. The scores for mitigation measure reliability and integrity are 95.0 because all required information is reported and reliable.

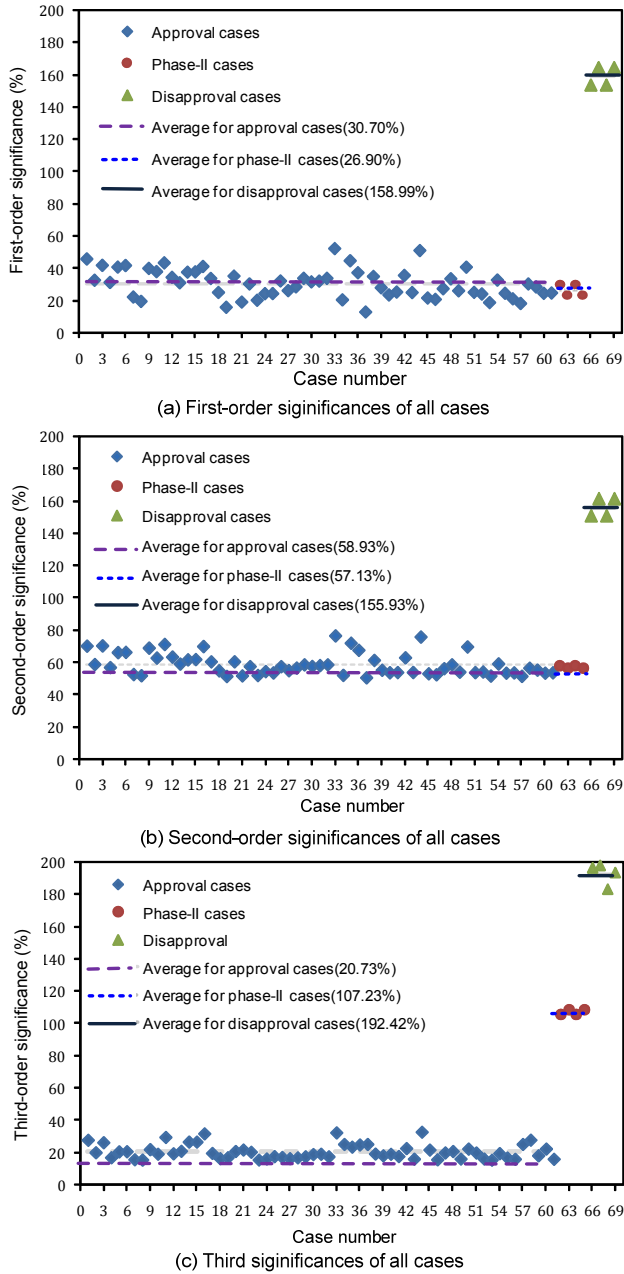


Figure 4. First, second and third-order significance and their distributions over all cases (air pollution).

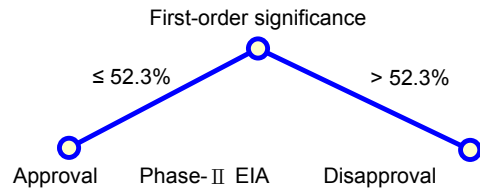
Based on these significance criteria scores and the second-order significance value, the fuzzy rule-based system value is derived as 10.1, and the third-order or overall significance of air pollution (FRBS I1-3) is 28.3%.

As for the air pollution indicator, there are 63 real and 6 artificial road construction projects consisting of 61 real cases (case no. 01 ~ 61) receiving conditional approval, two real cases (case no. 62 ~ 63) and another two artificial cases (case no. 64 and 65) are Phase-II EIA, and four artificial cases (case no. 66 ~ 69) received disapproval. The artificial cases are not real-world ones but were devised by the authors to allay the

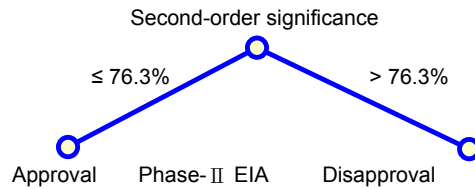
problem of the lack of real-world cases, especially the Phase-II EIA and disapproved projects. At this point, the 69 projects can be estimated for their first, second and third-order significances of air (I1), and their distributions are presented in Figure 4, respectively. Obviously, the Phase-II EIA projects cannot be distinguished from the pool of conditionally approved cases in the assessments of first and second-order significances until the third-order significances. Finally, through the same procedure, the other nine indicators are also used to compute the first, second and third-order significances for all projects.

3.5. Data Mining for Predicting Review Results

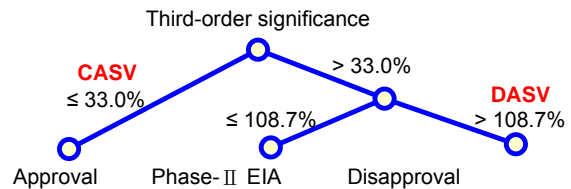
Classification trees are a common method in data mining. A set of pre-classified samples are used to create a tree-like model that can predict the class of a new sample based on its attributes. At each node of the tree, classification algorithms choose the next best attribute with the highest normalized



(a) 94.2% correctness of classification



(b) 94.2% correctness of classification



(c) 100.0% correctness of classification

Figure 5. Classification trees for the first, second and third-order significance (air pollution).

information gain (difference in entropy) to split the set of samples. This process is repeated on each derived subset in a recursive manner called recursive partitioning. The recursion is completed when the subset at a node all has the same value of the attribute, or when splitting no longer adds value to the

Table 5. Significance Criteria and Their Rating Guidelines for Air Pollution

Rating	Significance criteria			
	Sensitivity	Spatial extent	Mitigation measure reliability	Information integrity
Range	0 - 100	0 - 100	0 - 100	0 - 100
Guidelines	90: Class I control regions means such areas as national parks and nature protection and conservation areas. 60: Class III control regions means areas that do not meet air quality standards, with the exception of Class I control regions. 30: Class II control regions means areas that meet air quality standards, with the exception of Class I control regions.	Percentage of all predictions about air pollution exceeds the Taiwan standard values.	100: All mitigation measures are reliable. 70: Only unimportant mitigation measures are unreliable. 50: Secondary mitigation measures are unreliable. 30: Major mitigation measures are unreliable.	100: All information is reported. 70: Only unimportant information is lost. 50: Secondary information is lost. 30: Major information is lost.

Table 6. Splitting Values of Classification Trees and Their Validities

Indicator	Splitting values		Third-order significances for test cases						Validity of previous work (Liu and Yu, 2009b)	
	CASV	DASV	1	2	3	4	5	6		validity
Air (I ₁)	33.0%	108.7%	26.8%	32.8%	78.0%	79.2%	122.8%	113.0%	100%	75.5%
Water (I ₂)	86.0%	152.1%	48.1%	60.2%	133.0%	134.2%	207.7%	209.8%	100%	80.0%
Soil (I ₃)	49.4%	116.6%	38.1%	40.6%	114.7%	115.4%	147.7%	148.4%	100%	100%
Solid waste (I ₄)	65.1%	149.0%	40.1%	42.7%	120.7%	121.4%	155.4%	156.2%	100%	80.0%
Noise (I ₅)	71.5%	141.9%	46.1%	48.8%	122.0%	122.9%	171.1%	171.6%	100%	100%
Terrestrial (I ₆)	75.8%	170.3%	52.6%	60.8%	129.9%	136.7%	181.3%	180.8%	100%	100%
Aquatic (I ₇)	75.8%	159.7%	44.8%	70.8%	141.4%	134.5%	194.4%	201.1%	100%	100%
Economics (I ₈)	52.1%	130.8%	39.4%	42.8%	102.9%	103.9%	146.7%	147.4%	100%	100%
Society (I ₉)	49.6%	116.6%	39.2%	40.4%	95.2%	95.9%	128.8%	129.6%	100%	100%
Culture (I ₁₀)	54.0%	134.3%	39.2%	40.9%	109.6%	110.5%	147.1%	147.4%	100%	100%
Overall validity			100%							92.7%

Note: CASV: Conditional approval splitting value; DASV: Disapproval splitting value.

predictions. In this paper, See5 (Quinlan, 2010) is used to construct the tree.

The classification trees for the first, second and third-order significances of air (I₁) is demonstrated in Figure 5, which shows that the correctness of the classification trees for the first and second-order significances is 94.2% because they fail to designate phase-II EIA projects. Nevertheless, the classification tree for the third-order significance (Figure 5(c)) is able to identify three review results: a case will be predicted to receive conditional approval for air pollution if its third-order significance is lower 33.0% (conditional approval splitting value, CASV); if its value is between 33.0 and 108.7% it will be asked to fulfill phase-II EIA; disapproval will be suggested if the value is greater than 108.7% (disapproval splitting value, DASV). Similarly, the classification trees of the other nine indicators are also constructed and presented in Table 6, which shows that the CASV for the ten indicators are 33.0, 86.0, 49.4, 65.1, 71.5, 75.8, 75.8, 52.1, 49.6 and 54.0%, sequentially; 108.7, 152.1, 116.6, 149.0, 141.9, 170.3, 159.7, 130.8, 116.6 and 134.3%, respectively, are the DASV of the ten indicators.

3.6. Software

The software used to construct the integrated prediction model included the MATLAB Fuzzy Logic Toolbox (The

Mathworks Inc. USA) and the Data Mining Tools See5 (Rule-Quest Research Pty Ltd, Australia). The graphical user interfaces editors and viewers in the MATLAB Fuzzy Logic Toolbox make users easy to build the rules, define the membership functions, and analyze the behavior of a fuzzy inference system. It was used to develop the 21 rulebases containing 2,302 fuzzy rules and their corresponding fuzzy inference systems for the purpose of evaluating impact significance. The 21 fuzzy inference systems is available free of cost from the corresponding author. See5 is also easy to use and assist authors to establish the 30 classification trees for predicting review results.

3.7. Model Verification

A critical step of model development is to verify its feasibility or correctness. The classification tree for the third-order significances of air (I₁) can successfully categorize all existing cases; in other words, the correctness of the classification of air (I₁) based on the third-order significance is 100.0%. Besides, the classification trees of the other nine indicators can also reach 100.0% correctness for all existing cases.

Another six real cases, excluded in the previous 63 real projects, are reserved to test the validity of the prediction model. The six real cases are identical to those used in the authors' previous work (Liu and Yu, 2009b), with two cases of



Figure 6. Projected route for the case study.

conditional approval, two phase-II EIA cases and two disapproved projects. The final test results of this study are presented in Table 6 and are compared with the previous study (Liu et al., 2009b). Table 6 shows that the overall validity is enhanced from 92.7 to 100.0%.

4. Case Study

4.1. Case Description

Taiwan High-Speed Rail (THSR), commencing on January 5, 2007, is capable of running at up to 300 kilometers per hour and travels from Taipei City to Kaohsiung City in about 90 minutes, compared the 4.5 hours spent by trains on the conventional western trunk line of the Taiwan Railway Administration. It passes 13 major cities and runs 344.68 kilometers, including about 252 kilometers of overpasses. The spaces under the viaducts were designed as part of the transportation network and were successively constructed as connections between cities and THSR stations. As shown in Figure 6, the case study is a 12.147-kilometer extension of provincial highway no.31, which runs beneath the THSR bridges and is divided by piers as north-south bounds. This project intends to minimize the cost of land purchase and the demolition of existing buildings, and it also intends to achieve the goal of easing the traffic burden between Taoyuan and Hsinchu counties.

For preventing a lateral impact on the adjacent environment along the area of the case study within the construction and operation stages, an environmental impact statement concerning the natural, biological, social and economical impacts was submitted to the review committee in July 2009 and ultimately conditionally approved in December of the same year. According to the information provided in this environmental impact statement, as summarized in the last three columns of Table 2, the prediction model of the review result pro-

posed in this paper demonstrates its use. In Table 2, three conditions are discussed: the baseline condition (BC) before construction, prediction of the impact without mitigation measures (PIWOM) and prediction of the impact with mitigation measures (PIWM). Indicators I1, I2, I3 and I5 take the average of all measurement points and the others are subjectively estimated based on Table 3 as follows.

The case study is located in rural areas which are made up of mostly agricultural land, ponds and a few residences. Thereby, the rubbish load (I41) is minor for BC and is rated 4.0; for PIWOM, about 200 construction workers generate rubbish in the amount of 0.2 ton per day and thus, the I41 score is 9.0; the mitigation measure deals with the waste by government cleaning units and enables I41 to decrease to 5.0 in PIWM. As for construction waste (I42), BC does not produce industrial waste, therefore, it is rated as 0; in PIWOM, a large number of excavations are generated, so I42 waste is estimated at 33.0. After the mitigation measures for PIWM, almost all the excavations are reused on site, causing I42 to decrease to 3.0.

As the terrestrial animals (I61) in these areas include common native species and the terrestrial plants (I62) are mostly crops along the projected line, the impact of existing THSR on the surrounding terrestrial is minor, with about 3% of them being affected in BC. The construction will have a significant impact on them, about 20% in PIWOM. After taking some mitigation measures such as reducing air pollution to not cover surface of the plan leaves or reducing noise and vibration as to avoid disturbance of the nearby wildlife habitats, the affected percentage can be brought down to 10% in PIWM. The surrounding endangered species (I63) are all birds, they are four mynas, one falco tinnunculus, two serpent eagles and eight brown shrikes. This area involves 43 bird species and 1,621 birds; therefore, the endangered birds make up about 9.3% of all birds (15/1,621). The extensive surrounding area of the project provides similar habitat, so the impact of the existing THSR on these species is minor, about 1.0% in BC. The construction will not influence them much because it is certain distance away, about 3.0% in PIWOM. Through the mitigation measures, for example districted construction areas, the score may drop to 2.0% in PIWM.

The case study crosses over four slightly polluted rivers and nine ponds which are already heavily or moderately polluted, therefore the aquatic animals (I71) and plants (I72) have been serious affected, about 15 and 10%, respectively, in BC. In PIWOM, the slight leakage of oil and sewage brings the scores up to 20 and 15%, respectively. Mitigation measures such as collecting sewage and controlling oil leaks appropriately can make the affected ratios fall to 18 and 12%, respectively, in PIWM. The scores are all 0% under any conditions for indicator I73 since there are no endangered species in these ponds. As the projected line runs through agricultural land with small population, land-use and development obstacle (I81), life-quality decline (I82) and economic activity disturbance (I83) are only slight. Thus, these factors are rated as 2.0, 2.0 and 0, respectively, in BC. In PIWOM, these three indicators create a slight impact and are

Table 7. Significance Evaluation and Prediction of the Results of the Review for the Case Study

I	First-order significance			Second-order significance	Significance criteria				TOS	Review result prediction
	BC	PIWOM	PIWM		Sensitivity	SE	MMR	II		
I ₁	25.3%	35.6%	29.1%	45.6%	60.0	0	95.0	100.0	30.8%	CA (< 33.0%)
I ₂	115.2%	115.2%	115.2%	106.2%	30.0	100.0	85.0	85.0	80.9%	CA (< 86.0%)
I ₃	73.9%	73.9%	73.9%	87.4%	NA	0	100.0	100.0	31.3%	CA (< 49.4%)
I ₄	18.8%	97.6%	35.6%	66.6%	NA	NA	90.0	95.0	51.5%	CA (< 65.1%)
I ₅	83.7%	105.2%	99.1%	101.6%	46.2	50.0	90.0	85.0	69.4%	CA (< 71.5%)
I ₆	41.1%	74.6%	57.1%	77.4%	10	10.0	80.0	80.0	67.4%	CA (< 75.8%)
I ₇	54.5%	61.2%	58.5%	75.6%	5	10.0	80.0	85.0	66.7%	CA (< 75.8%)
I ₈	42.2%	45.1%	45.1%	60.4%	0	NA	100.0	100.0	27.8%	CA (< 52.1%)
I ₉	47.4%	74.4%	57.8%	77.8%	NA	NA	100.0	95.0	47.8%	CA (< 49.6%)
I ₁₀	38.6%	55.8%	51.6%	69.4%	0	NA	100.0	95.0	33.1%	CA (< 54.0%)

* BC: baseline condition; PIWOM: predicted pollution increment without mitigation measures; PIWM: predicted pollution increment with mitigation measures; SE: Spatial extent, MML: Mitigation Measure Liability; II: Information integrity; TOS: Third-order significance; NA: Not applicable; CA: Conditional approval; Air (I₁), Water (I₂), Soil (I₃), Solid waste (I₄), Noise (I₅), Terrestrial (I₆), Aquatic (I₇), Economics (I₈), Society (I₉), Culture (I₁₀).

Table 8. Verification of Predicted Significant Indicators by Comparison with the Review Committee’s Opinions and Their Improvements

I	(A)	(A)/CASV	RCO	Improvement in third-order significance				(B)	Improvement [(B) - (A)]/(A)	
				Second-order significance	Significance criteria					
					Sensitivity	SE	MMR	II		
I ₁	30.8%	93.3%	2	45.6%	60.0	0	100.0	100.0	25.3%	-17.8%
I ₂	80.9%	94.1%	7	106.2%	30.0	100.0	100.0	100.0	56.9%	-29.7%
I ₃	31.3%	63.4%	1	87.4%	NA	0	100.0	100.0	31.3%	0.0%
I ₄	51.5%	79.1%	2	66.6%	NA	NA	100.0	100.0	39.4%	-23.5%
I ₅	69.4%	97.1%	9	101.6%	46.2	50.0	100.0	100.0	43.9%	-36.8%
I ₆	67.4%	88.9%	7	77.4%	10	10.0	100.0	100.0	24.5%	-63.7%
I ₇	66.7%	88.1%	4	75.6%	5	10.0	100.0	100.0	23.7%	-64.5%
I ₈	27.8%	53.4%	1	60.4%	0	NA	100.0	100.0	27.8%	0.0%
I ₉	47.8%	96.4%	3	77.8%	NA	NA	100.0	100.0	45.1%	-5.6%
I ₁₀	33.1%	61.3%	3	69.4%	0	NA	100.0	100.0	27.8%	-15.9%

Note: (A): Predicted Third-order significance; CASV: Conditional approval splitting value; RCO: Review committee’s opinions on significance criteria; SE: Spatial extent, MML: Mitigation Measure Liability; II: Information integrity; (B): Improved third-order significance; Air (I₁), Water (I₂), Soil (I₃), Solid waste (I₄), Noise (I₅), Terrestrial (I₆), Aquatic (I₇), Economics (I₈), Society (I₉), Culture (I₁₀).

rated as 3.0, 3.0 and 1.0, respectively. Because of failure to take mitigation measures in PIWM, the scores are unchanged.

There are no accessibility problems for public facilities around the project. Thus, indicator I91 is rated as 0 under any condition. In addition, the existing THSR disconnects transportation and communities slightly, so indicators I92 and I93 are assessed as 10.0 and 5.0, respectively, in BC. The construction will produce impacts on the above two indicators, and thus, the scores are 50.0 and 10.0, respectively, in PIWOM. Taking mitigation measures, such as traffic management, can make them decrease to 30.0 and 6.0, respectively, in PIWM. Indicator I101 is rated as 0 in any condition because the project is in an area with no cultural heritage. The existing THSR slightly damages the landscape and the indicator I102 is thereby assessed as 3.0 in BC. The construction will heavily change the landscape, making the score go up to 10.0 in PIWOM. Mitigation measures, such as beautification of construction fences and management of construction layout, can drop the score to 8.0 in PIWM.

4.2. Evaluation through Fuzzy Rule-Based System and Prediction by Classification Tree

The above-mentioned case information provides the basis for first-order significance assessment of the ten indicators in the three conditions (BC, PIWOM and PIWM), the results of which are listed in Table 7. In this table, the expression of percentage, a ratio compared to standard values, is used to involve the concept of significance thresholds, as stated in Section 3.4. Table 7 shows, obviously, that first-order significances of water pollution (I2) in all conditions are beyond 100%, indicating they are over-standard, as a result of the slightly polluted rivers and heavily polluted ponds. In addition, noise (I5) in PIWOM is also over-standard due to excessive construction noise. The second-order significance, in the fifth column of Table 7, merges the first-order significances of BC, PIWOM and PIWM to produce a percentage score which, similarly, expresses a proportion to standard values. As expected, water pollution (I2) and noise (I5) are also assessed to be over-standard (beyond 100%) according to second-order signifi-

cance, which is represented by a single-underline in Table 7.

The third-order significance mingles values with second-order significance. The values coming from the EIA review committee are referred to as significance criteria. These criteria include sensitivity, spatial extent, mitigation measure reliability and information integrity. On the basis of rating guidelines (see Table 5), the entire line of the case study is located in a class-III air pollution control region and therefore its air pollution sensitivity is 60.0, as shown in Table 7. Water pollution has a sensitivity of 30.0 since it is located in a water pollution control zone. The case study obtains a sensitivity score of 46.2 for noise pollution, which is the average derived when crossing through two category-II and five category-III noise pollution control zones. Spatial extent refers to the proportion of measurement points which are over-standard. All rivers and ponds with at least slight pollution make the spatial extent of water pollution 100.0; noise pollution is over standard by 50.0% even in PIWM; terrestrial and aquatic are only over-standard by 10.0% of the spatial extent. As for the assessment of mitigation measure reliability and integrity, it requires professional and subjective judgments, the results of which are shown in Table 7. Compared to the predefined standard values of the four significance criteria, the over-standard scores are single-underlined in this table. The third-order significances, the penultimate second column of this table, are the final outcomes when combining these significance criteria with second-order significances; that is, they are 30.8, 80.9, 31.3, 51.5, 69.4, 67.4, 66.7, 27.8, 47.8 and 33.1%, respectively, for the ten indicators. Although several indicators such as air (I1), water (I2) and noise (I5) are over-standard in second-order significance or significance criteria, however, all indicators are ultimately under-standard and are also below conditional approval splitting values (referring to Table 6) in third-order significance. This is mainly because of the reliable mitigation measures and adequate information. Eventually, the review result for this case study is predicted to receive conditional approval, as shown in the last column of Table 7.

4.3. Discussion

Significant indicators. The purpose of predicting review results is to identify significant environmental factors and thereby improve the environmental management plan in advance even though the case is forecasted to receive conditional approval. In this case, the predicted review result is for conditional approval; nevertheless, water (I2), noise (I5), terrestrial (I6) and aquatic (I7) are the four relatively significant indicators. These indicators are double-underlined in the second column of Table 8. The former two have higher third-order significances because of their high second-order significances and spatial extent; the latter two are due to relatively low mitigation measure reliability and information integrity. If the third-order significances of the ten indicators are compared to the conditional approval thresholds, the additional two indicators, air (I1) and society (I9), are highlighted as factors of concern, as shown in the third column of Table 8.

Verification. Finally, the case study was reviewed by the

committee on July 27 and September 17 of 2009 and received the result of conditional approval on January 11 of the next year, which coincided with the prediction of this study. In the two review meetings, the number of committee review opinions was 67 in total and 58.2% of them (37 out of 67) were related to the four significance criteria selected in this paper, which confirms the applicability of these criteria. The distribution of the 37 review opinions over the ten indicators is delineated in the fourth column of Table 8 and it reveals that water (I2), noise (I5) and terrestrial (I6) are the relatively significant indicators, which concurred with the prediction, with the exception of aquatic (I7). Water (I2) received seven opinions and most of them are related to unreliable mitigation measures such as concern for downstream river water quality arising from the use of emulsifiers to remove oil. The mitigation measures for noise (I5) are also unclear in, for example, the impact and management of noise pollution for nearby residents and schools. Furthermore, lack of long-term vibration monitoring and misuse of revised noise control standards led the committee to ask the developer to provide more information. No details about how to avoid the impact on protected wild animals incurred the committee's worries about the terrestrial (I6) factor. Moreover, the committee was also concerned with the ecological investigation of wetlands and the impact on the separation of habitats by the project and recommended additional ecological monitoring.

Improvement. The genuine goal of assessment is to make improvement. For a developer, it is easier to improve an environmental impact statement or EIA report by enhancing mitigation measures and strengthening information integrity. Assume that these two criteria acquire perfect scores (100) for all indicators. Therefore their third-order significances can be ameliorated as shown in the last two columns of Table 8. Particularly, the predicted significant indicators of water (I2), noise (I5), terrestrial (I6) and aquatic (I7) are largely improved at rates of 29.7, 36.8, 63.7 and 64.5%, respectively.

5. Conclusions

The purpose of predicting EIA review results is to help developers identify and deal with risky nuisances to the environment in their development projects, and then enable them to pass EIA reviews. The prediction can be achieved by a comprehensive use of several artificial intelligence technologies. The authors' previous paper (Liu and Yu, 2009b) used case-based reasoning for retrieval of similar cases and fuzzy rule-based systems for qualitative risk forecast and presented an overall prediction validity of 92.7%. This paper proposed another comprehensive use of several artificial intelligence technologies: (1) fuzzy rule-based system for evaluating impact significance; (2) significance transformation for incorporating significance thresholds; and (3) data mining for predicting EIA review results. A case study of road construction in north Taiwan was used to illustrate the use of the integrated system. The major advantages of fuzzy rule-based system are that it can model the human thought process in analyzing complex systems and decisions, and it allows a human expert

to naturally express his knowledge. Therefore, it is easy to model experts' knowledge of significance evaluation. In this study, a total of 21 fuzzy rule-based systems containing 2,302 fuzzy rules were developed. For the case study, although several indicators such as air (I1), water (I2) and noise (I5) are over-standard in second-order significance or significance criteria, the third-order significances of the ten indicators were under-standard and were estimated as 30.8% for air (I1), 80.9% for water (I2), 31.3% for soil (I3), 51.5% for solid waste (I4), 69.4% for noise (I5), 67.4% for terrestrial (I6), 66.7% for aquatic (I7), 27.8% for economics (I8), 47.8% for society (I9) and 33.1% for culture (I10), sequentially. Nevertheless, water (I2), noise (I5), terrestrial (I6) and aquatic (I7) are the four relatively significant indicators. The main benefits of classification trees are their abilities to explore the most relevant attributes and their corresponding splitting values at each node, and to easily and explicitly explain classification knowledge. Ten classification trees were established for the ten indicators, which present the CASVs for the ten indicators as 33.0, 86.0, 49.4, 65.1, 71.5, 75.8, 75.8, 52.1, 49.6 and 54.0%, respectively; the DASVs of the ten indicators are 108.7, 152.1, 116.6, 149.0, 141.9, 170.3, 159.7, 130.8, 116.6 and 134.3%, respectively. For the case study, all indicators were ultimately below CASVs in third-order significance and therefore the case was predicted to receive conditional approval. The critical step of model development verifies feasibility and correctness. This study adopted six projects from the authors' previous work to test the validity of the prediction model. The validity of the previous prediction model is 92.7% and it can be enhanced to 100.0% in this paper. The genuine goal of assessment is to make improvement. If mitigation measures and information integrity is improved for this case study thereby the predicted significant indicators water (I2), noise (I5), terrestrial (I6) and aquatic (I7) can be largely improved 29.7, 36.8, 63.7 and 64.5%, respectively.

The techniques used here are artificial intelligence techniques. Artificial intelligence intends to simulate human thinking but cannot replace human in this case, of course including the task of predicting (or even guiding) the significance determination by systematic, explicit, but still qualitative, procedures. However, this paper still attempted to take advantages of artificial intelligence to the task. The model perhaps cannot produce better prediction than the results achieved with more qualitative and accessible procedures, but it will be more efficient if the database is implemented.

The proposed integrated model can be further improved in some directions. Firstly, although it serves as a decision support for developers, it can also become a quantitative tool which would assist the qualitative procedure in current EIA review if it can further involve a wider range of environmental considerations, greater involvement of interested and affected parties, consideration of indirect and cumulative effects, positive impacts and sustainability considerations. Secondly, one restriction of this model is that it have to summarize a score for each subindicator in a case, which might average out some severe points over time or space dimensions that they warrant project rejection, regardless of the aggregate score.

One possible solution uses interval values instead of precise numbers to include the range of worst and least impacts in rating subindicators.

References

- Andriantatsaholainaina, L.A., Kouikoglou, V.S., and Phillis, A.P. (2004). Evaluating strategies for sustainable development: fuzzy logic reasoning and sensitivity analysis. *Ecol. Econ.*, 48 (2), 149-172. <http://dx.doi.org/10.1016/j.ecolecon.2003.08.009>
- Antunes, P., Santos, R., and Jordao, L. (2001). The application of Geographical Information Systems to determine environmental impact significance. *Environ. Impact Assess. Rev.*, 21(6), 511-535. [http://dx.doi.org/10.1016/S0195-9255\(01\)00090-7](http://dx.doi.org/10.1016/S0195-9255(01)00090-7)
- Beattie, R.B. (1995). Everything you already know about EIA (but don't often admit). *Environ. Impact Assess. Rev.*, 15(2), 109-14. [http://dx.doi.org/10.1016/0195-9255\(95\)00001-U](http://dx.doi.org/10.1016/0195-9255(95)00001-U)
- Birant, D. (2011). Comparison of decision tree algorithms for predicting potential air pollutant emissions with data mining models. *J. Environ. Inf.*, 17(1), 47-53. <http://dx.doi.org/10.3808/jei.201100186>
- Bojórquez-Tapia, L.A., Ezcurra, E., and García, O. (1998). Appraisal of environmental impacts and mitigation measures through mathematical matrices. *J. Environ. Manage.*, 53(1), 91-99. <http://dx.doi.org/10.1006/jema.1998.0191>
- Borri, D., Concilio, G., and Conte, E. (1998). A fuzzy approach for modeling knowledge in environmental systems evaluation. *Comput., Environ. Urban Syst.*, 22 (3), 299-313. [http://dx.doi.org/10.1016/S0198-9715\(98\)00045-3](http://dx.doi.org/10.1016/S0198-9715(98)00045-3)
- Cloquell-Ballester, U.A., Monterde-Díaz, R., Cloquell-Ballester, V.A., and Santamarina-Siurana, M.C. (2007). Systematic comparative and sensitivity analyses of additive and outranking techniques for supporting impact significance assessments. *Environ. Impact Assess. Rev.*, 27(1), 62-83. <http://dx.doi.org/10.1016/j.eiar.2006.08.005>
- de Siqueira Campos Boclin, A., and de Mello, R. (2006). A decision support method for environmental impact assessment using a fuzzy logic approach. *Ecol. Econ.*, 58 (1), 170-181. <http://dx.doi.org/10.1016/j.ecolecon.2005.06.007>
- Dixon, M., Gallop, J.R., Lambert, S.C., and Healy, J.V. (2007). Experience with data mining for the anaerobic wastewater treatment process. *Environ. Model. Software*, 22(3), 315-322. <http://dx.doi.org/10.1016/j.envsoft.2005.07.031>
- Duinker, P.N., and Beanlands, G.E. (1986). The significance of environmental impacts: an exploration of the concept. *Environ. Manage.*, 10(1), 1-11. <http://dx.doi.org/10.1007/BF01866412>
- Ekasingh, B., and Ngamsomsuke, K. (2009). Searching for simplified farmers' crop choice models for integrated watershed management in Thailand: A data mining approach. *Environ. Model. Software*, 24 (12), 1373-1380. <http://dx.doi.org/10.1016/j.envsoft.2009.02.015>
- Ekasingh, B., Ngamsomsuke, K., Letcher, R.A., and Spate, J. (2005). A data mining approach to simulating farmers' crop choices for integrated water resources management. *J. of Environ. Manage.*, 77(4), 315-325. <http://dx.doi.org/10.1016/j.jenvman.2005.06.015>
- Enea, M., and Salemi, G. (2001). Fuzzy approach to the environmental impact evaluation. *Ecol. Model.*, 136(2-3), 131-147. [http://dx.doi.org/10.1016/S0304-3800\(00\)00380-X](http://dx.doi.org/10.1016/S0304-3800(00)00380-X)
- Gibert, K., Rodriguez-Silva, G., and Rodriguez-Roda, I. (2010). Knowledge discovery with clustering based on rules by states: A water treatment application. *Environ. Model. Software*, 25(6), 712-723. <http://dx.doi.org/10.1016/j.envsoft.2009.11.004>
- González, B., Adenso-Díaz, B., and González-Torre, P.L. (2002). A fuzzy logic approach for the impact assessment in LCA. *Resour. Conserv. Recycling*, 37(1), 61-79. [http://dx.doi.org/10.1016/S0921-3449\(02\)00069-1](http://dx.doi.org/10.1016/S0921-3449(02)00069-1)

- Gross, D.S., Atlas, R., Rzeszotarski, J., Turetsky, E., Christensen, J., Benzaid, S., Olson, J., Smith, T., Steinberg, L., Sulman, J., Ritz, A., Anderson, B., Nelson, C., Musicant, D.R., Chen, L., Snyder, D.C., and Schauer, J.J. (2010). Environmental chemistry through intelligent atmospheric data analysis. *Environ. Model. Software*, 25(6), 760-769. <http://dx.doi.org/10.1016/j.envsoft.2009.12.001>
- Haug, P.T., Burwell, R.W., Stein, A., and Bandurski, B.L. (1984). Determining the significance of environmental issues under the National Environmental Policy Act. *J. Environ. Manage.*, 18(1), 15-24.
- Kontic, B. (2000). Why are some experts more credible than others? *Environ. Impact Assess. Rev.*, 20(4), 427-34. [http://dx.doi.org/10.1016/S0195-9255\(00\)00057-3](http://dx.doi.org/10.1016/S0195-9255(00)00057-3)
- Lawrence, D.P. (2007). Impact significance determination-Back to basics. *Environ. Impact Assess. Rev.*, 27(8), 755-769. <http://dx.doi.org/10.1016/j.eiar.2007.02.011>
- Le Ber, F., Benoit, M., Schott, C., Mari, J.F., and Mignolet, C. (2006). Studying crop sequences with CarrotAge, a HMM-based data mining software. *Ecol. Model.*, 191(1), 170-185. <http://dx.doi.org/10.1016/j.ecolmodel.2005.08.031>
- Liu, K.F.R. (2007). Evaluating environmental sustainability: An integration of multiple-criteria decision-making and fuzzy logic. *Environ. Manage.*, 39(5), 721-736. <http://dx.doi.org/10.1007/s00267-005-0395-8>
- Liu, K.F.R., and Lai J.H. (2009a). Decision-support for environmental impact assessment: A hybrid approach using fuzzy logic and fuzzy analytic network process. *Expert Syst. Appl.*, 36, 5119-5136. <http://dx.doi.org/10.1016/j.eswa.2008.06.045>
- Liu, K.F.R., Liang, H.H., Yeh, K., and Chen, C.W. (2010). A qualitative decision support for environmental impact assessment using fuzzy logic. *J. Environ. Inf.*, 13(2), 94-104. <http://dx.doi.org/10.3808/jei.200900144>
- Liu, K.F.R., and Yu, C.W. (2009b). Integrating case-based and fuzzy reasoning to qualitatively predict risk in an environmental impact assessment review. *Environ. Model. Software*, 24(10), 1241-1251. <http://dx.doi.org/10.1016/j.envsoft.2009.04.005>
- Lv, Y., Huang, G.H., Li, Y.P., Yang, Z.F., Liu, Y., Cheng, G.H., 2010. Planning regional water resources system using an interval fuzzy bi-level programming method. *J. Environ. Inf.*, 16(2), 43-56. <http://dx.doi.org/10.3808/jei.201000177>
- Morón, A.B., Calvo-Flores, M.D., Ramos, J.M.M., and Almohano, M.P.P. (2009). AIEIA: Software for fuzzy environmental impact assessment. *Expert Syst. Appl.*, 36(5), 9135-9149. <http://dx.doi.org/10.1016/j.eswa.2008.12.055>
- Peche, R., and Rodríguez, E. (2009). Environmental impact assessment procedure: A new approach based on fuzzy logic. *Environ. Impact Assess. Rev.*, 29(5), 275-283. <http://dx.doi.org/10.1016/j.eiar.2009.01.005>
- Peche, R., and Rodríguez, E. (2011). Environmental impact assessment by means of a procedure based on fuzzy logic: A practical application. *Environ. Impact Assess. Rev.*, 31(2), 87-96. <http://dx.doi.org/10.1016/j.eiar.2010.03.006>
- Phillis, Y.A., and Andriantiatsaholiniaina, L.A. (2001). Sustainability: an ill-defined concept and its assessment using fuzzy logic. *Ecol. Econ.*, 37(3), 435-456. [http://dx.doi.org/10.1016/S0921-8009\(00\)00290-1](http://dx.doi.org/10.1016/S0921-8009(00)00290-1)
- Prentzas, J., and Hatzilygeroudis, I. (2007). Categorizing approaches combining rule-based and case-based reasoning. *Expert Syst.*, 24(2), 97-122. <http://dx.doi.org/10.1111/j.1468-0394.2007.00423.x>
- Quinlan, J.R. (2010). Data Mining Tools See5/C5.0. In <http://www.rulequest.com/see5-info.html>
- Rossouw, N. (2003). A review of methods and generic criteria for determining impact significance. *African Journal of Afr. J. Environ. Assess. Manage.*, 6, 44-61.
- Roussel, O., Cavelier, A., and Werf, H.M.G. (2000). Adaptation and use of a fuzzy expert system to assess the environmental effect of pesticides applied to field crops. *Agric. Ecosyst. Environ.*, 80(1), 143-158. [http://dx.doi.org/10.1016/S0167-8809\(00\)00142-0](http://dx.doi.org/10.1016/S0167-8809(00)00142-0)
- Smith, P.N. (2002). Linguistic evaluation method for the environmental assessment of infrastructure projects. *Int. J. Syst. Sci.*, 33(7), 567-575. <http://dx.doi.org/10.1080/00207720210123760>
- Tennakoon, M., Mayorga, R.V., and Shirif, E. (2010). A fuzzy inference system prototype for indoor air and temperature quality monitoring and hazard detection. *J. Environ. Inf.*, 16(2), 70-79. <http://dx.doi.org/10.3808/jei.201000179>
- Van der Werf, H.M.G., and Zimmer, C. (1998). An indicator of pesticide environmental impact based on a fuzzy expert system. *Chemosphere*, 36(10), 2225-2249. [http://dx.doi.org/10.1016/S0045-6535\(97\)10194-1](http://dx.doi.org/10.1016/S0045-6535(97)10194-1)
- Wilkins, H. (2003). The need for subjectivity in EIA: discourse as a tool for sustainable development. *Environ. Impact Assess. Rev.*, 23(4), 401-414. [http://dx.doi.org/10.1016/S0195-9255\(03\)00044-1](http://dx.doi.org/10.1016/S0195-9255(03)00044-1)
- Wood, G. (2008). Thresholds and criteria for evaluating and communicating impact significance in environmental statements: 'See no evil, hear no evil, speak no evil'? *Environ. Impact Assess. Rev.*, 28(1), 22-38. <http://dx.doi.org/10.1016/j.eiar.2007.03.003>
- Wood, G., Rodriguez-Bachiller, A., and Becker, J. (2007). Fuzzy sets and simulated environmental change: evaluating and communicating impact significance in environmental impact assessment. *Environ. Plann. A*, 39(4), 810-829. <http://dx.doi.org/10.1068/a3878>
- Zadeh, L.A. (1975). The concept of a linguistic variable and its application to approximate reasoning. *Inf. Sci.*, 8, 199-249. [http://dx.doi.org/10.1016/0020-0255\(75\)90036-5](http://dx.doi.org/10.1016/0020-0255(75)90036-5)
- Zadeh, L.A. (1996). Fuzzy logic=computing with words. *IEEE Trans. Fuzzy Syst.*, 4(1), 103-111. <http://dx.doi.org/10.1109/91.493904>