CHAPTER 4

KNOWLEDGE DRIVEN MODEL FOR SHRIMP FARMS EXPANSION

Introduction

Modeling plays an important role to help understand and predict behavior of the ecological system. A model to predict where shrimp farms would likely take place is a potential tool to study the impact of shrimp farms on the environment and to come up with suggested management systems to avoid problem.

This chapter will discuss the concept of knowledge driven models for shrimp farm expansion and the steps of model generation. It is then followed by a justification of the modeling results.

Knowledge Driven Model

A knowledge based system can be defined as a computer program designed to model the problem-solving ability of a human expert (Durkin, 1994). To develop a model representing shrimp farm expansion one must know the desirous conditions of the investors to convert the land to shrimp farm. From such knowledge, one can formulate *the production rules* which are identified as one of the knowledge representation schemes in a knowledge based system development (Rolston, 1988).

The rule-based representation generated in this study is expected to describe patterns and the chance in process of changing mangrove or any other land cover for shrimp farm expansion.

The rules have the following form:

IF { Set of conditions : Variables $i,...,n$ }

THEN { Set of conclusion : A chance of area conversion to shrimp farm is equal to ... }

The rules logically relate information contained in the 'set of conditions' portion to other information contained in the 'set of conclusion' portion. Since the rules are developed from acquired knowledge, they are viewed as *a knowledge driven model*.

Determining Model Variables

Shrimp farming is a human activity that the occurrence relates to human's decision. There are various conditions that a land owner or investor (or a farmer) has to consider before making a decision to convert an area into shrimp farm. Most of the conditions are physical factors (Boonchana Klankumsorn *et al.*,1989). Therefore, the shrimp farm conversion occurred where necessary physical factors were met. Although some socio-economic factors also influence land-use, in this study they are threaten in similarly manner.

There are several factors contributed to a decision making process of the investor which can be entered as variables of the model. In this study, the following variables were considered:

- Distance from source of water (sea, river and cannel)
- Land cover
- Distance from road
- Soil type
- Legal status of the area

The above variables were selected and listed by importance order followed the result of investigation on importance factors for shrimp farming by Boonchana

Klankumsorn *et al.*(1989 from interviewing people who lived in the same area as this study.

Extracting data for the model variables were done by using several softwares. Data at different scales must be rectified to a common reference. The software used for these purposes were summarized as following:

IDRISI., developed by a research team at Clark University, USA., provides facilities for image displaying and data manipulation, such as regrouping data and data conversion.

TOSCA., developed by Graduate school of Geography, Clark University. It is a feature oriented interactive digitizing and vector management package designed for use with IDRISI. This software can be an efficient means for capturing data, a method for data input and data set creation and a mean for data interchange between IDRISI and other systems. This study used TOSCA to capture data from hard copy format (topographical map and soil map) to a digital vector form.

Variables

Distance from source of water

Shrimp farming needs brackish water. Source of salt water as well as fresh water are essential. The longer distance from both type of water would lead to cost increasing. However, this study has not separated distance from source of water as salt or fresh water because there are no consistent boundary between them in estuary.

To obtain these data, manually digitized into a computer readable form from the 1:50,000 topographical map of the study area by program TOSCA was conducted. Digitizer used in the study was the HiSketch 1812/D of Genius. The map was prepared and published by the Defense Mapping Agency Topographic Center(in Washington DC) in cooperation with the Royal Thai Survey Department in 1981. The digitized data were transformed into raster form by program IDRISI. The final resolution for the digital data is 30-m. registered to the UTM reference (Fig. 4.1).

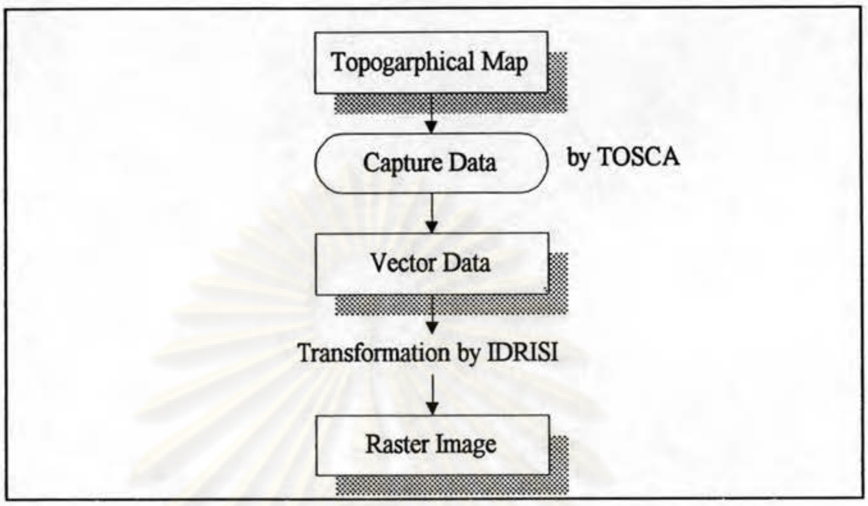

Figure 4.1 Data flow diagram of the process to generate variable data layer

Land cover .

Land cover is an important variable in decision process because it can visually recognized easily. Additionally, in business it is viewed as an original cost for evaluating benefit.

Land cover was classified from remotely sensed data as described in the previous chapter.

Distance from road

Road is one of the factors to be considered in selecting farm site for transportation of rearing materials, feed or yield after harvesting.

The procedure to extract these data is the same as the procedure of the preceding variables. It was captured through a digitizing process from the 1:50,000 topographical map of the study area.

Soil type

Although detection of soil properties required the use of specific instrument by technicians, but they could also be related to the land cover or the landuse since each soil series would be appropriated for different group of plants or landuse. Therefore considering on land cover of most investors is indirectly based on the properties of soil as well. In addition, some investors have been using some properties of soil which could obviously been recognized as considerate factors, e.g., texture and soil color.

Soil type data were digitized from the 1:100,000 detailed reconnaissance soil map of the study area. The map, prepared by Soil Survey Division, Department of Land Development, was registered to the UTM reference at 30-m. (the same scale of classified images).

Legal status of the area

December 15, 1987, the cabinet approved the mangrove forest land use zoning submitted by the National Committee on Mangrove Resources. The concept of zonation is based on the integration of physical and human environments. In detail, mangrove area could be classified into three principal zones:

Preservation Zone : This area is strictly protected from any activities in order to preserve the natural environmental system.

Economic Zone A : This area is allowable for forest utilization activities depending on a sustained yield basis.

Economic Zone B : This area allowed certain development which considered the impacts to the natural environment.

Most of the study area covered by mangrove are classified by zoning policy as Economic Zone A. The area the along coastal line which had shrimp farms for a long time at least prior than 1969 (observed in topographical map prepared since 1969) could be classified as Economic Zone B. Therefore, legal status of the study area should influenced decision making of the investors.

The zoning map was prepared by Royal Forestry Department on the scale 1:50,000. It was digitized and processed same as the preceding variables.

Using Knowledge Driven Model to Simulate Shrimp Farm Expansion Map

This section is concerned with describing the details of how models are developed. The process includes acquiring knowledge and classification data using the rules.

Developing Knowledge Driven Model

In the knowledge acquisition, there are several techniques for acquiring knowledge for knowledge base system, including interviewing techniques, protocol analysis, repertory grid technique, the expert-driven approach, machine-driven approach (Olson and Courtney, 1992). In this study the machine-driven approach was used. This technique let machine itself learn how problem are solved. As described by Olson and Courtney (1992), four types of its acquisition are available: learning by rote, learning by being told, learning from example and learning by analog. The most common approach is learning from example through induction and that was used in the study. Statistical analysis was used to objectively learn the relationship between variables. Although induction is objective, repeatable, indefatigable, consistent and easy to understand, but verification of induced rules is difficult. Problem resulted from the quality of induced results depending both on the algorithm used and the particular example set available.

Initially, formulating the rules requires the set of data called *the leaning data set* to represent area of shrimp farms expansion (in classified images) as well as for each variables of the model. To sample data sets, 7,885 pixels were randomly selected using module SAMPLE in program IDRISI.

The program KnowledgeSEEKER was used to formulate shrimp farms expansion rules from the training data. This program was developed by FirstMark

Technologies Ltd. in Canada. KnowledgeSEEKER is a menu driven system having the classification process in the form of *tree structure*. It can express the relation it discovers as a set of rules which can be stored as knowledge base. The program start from the root node, then partitions data of each independent variables into a number of homogenous groups. The number of groups depends on nature of data, e.g., ranges of data for continuous data, number of classes for categorical data, and the distribution of data themselves. User can view a list of variables having significant levels of separability. In order to decide whether variable in the list is suitable, users have alternative consideration either to use their background knowledge about the problem or let the program select the splitter automatically. After that, the program split the node into the number of branches equal to number of group partitioned in selected variable. At each branch end is child node that owns a part of data partitioned by its parent. Further partitioning was performed on data partitioned for each child node in the same manner of root node. The process will terminate when there is no more node which can be partitioned or when the user request after satisfying.

This study selected five variables including *land cover* , *distance from source of water*, *distance from road*, *soil type* and *legal status of the area* as the independent variables, and *landuse* as the dependent variable. The selection of splitting variables at each nods was considered by following the importance of variables as described in section 4.3. Finding the split was performed on each year in the same method while the land cover variable was organized in two ways. The first one, land cover variable of 1982 were used for all the following years in the classification process as an independent variable. For the other one, land cover variable was changed by relating to the following every other past year, for example, land cover of 1986 was used as independent variable in the model of 1988 and land cover of 1988 was used as independent variable in the model of 1990.

Figure 4.2 shows the classification tree developed for representing shrimp farms expansion in 1992 by using land cover 1982 as the independent variable. The tree can be converted into a set of rules by RULE command in KnowledgeSEEKER as shown in table 4.1. These rules can be used as knowledge in knowledge base system.

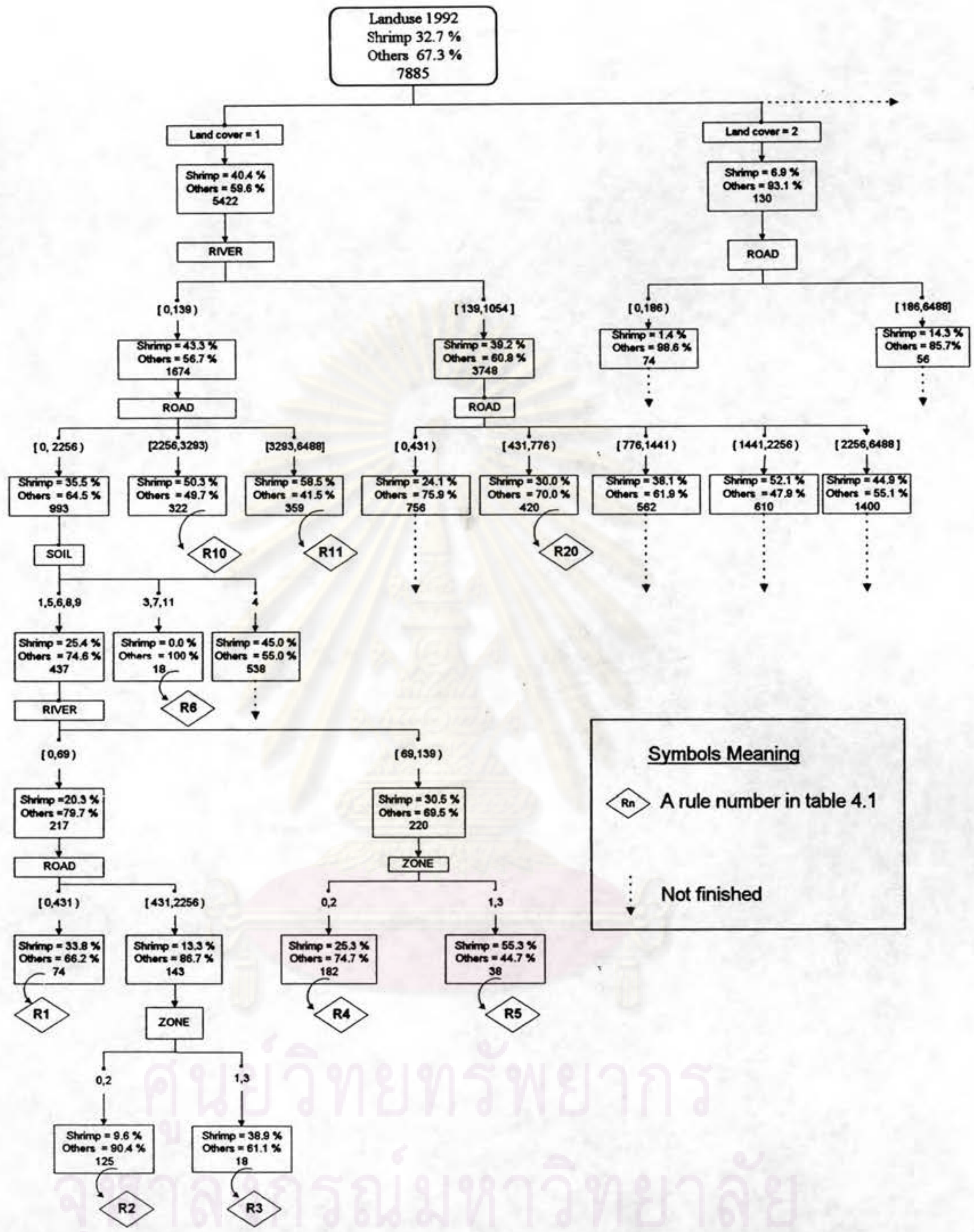

Figure 4.2 A classification tree for representing shrimp farm expansion

Table 4.1 Classification rules developed by the program KnowledgeSEEKER
('[' and ']' in an interval indicates that the range includes the number, while '(' and ')' indicates that the interval goes up to but does not include that number.)

Rule no.	Premise	% Occurrence of shrimp farms	No. of samples observed
1.	Land cover 1982 = 1 Soil = 1,5,6,8 or 9 River = [0,69) Road = [0,431)	33.8	74
2.	Land cover 1982 = 1 Soil = 1,5,6,8 or 9 River = [0,69) Road = [431,2256) Zone = 0 or 2	9.6	125
3	Land cover 1982 = 1 Soil = 1,5,6,8 or 9 River = [0,69) Road = [431,2256) Zone = 1 or 3	38.9	18
4	Land cover 1982 = 1 Road = [0,2256) Soil = 1,5,6,8 or 9 River = [69,139) Zone = 0 or 2	25.3	182
5	Land cover 1982 = 1 Road = [0,2256) Soil = 1,5,6,8 or 9 River = [69,139) Zone = 1 or 3	55.3	38
6	Land cover 1982 = 1 River = [0,139) Road = [0,2256) Soil = 3,7 or 11	0	18
7	Land cover 1982 = 1 River = [0,139) Soil = 4 Road = [0,186)	16.3	49
8	Land cover 1982 = 1 River = [0,139) Soil = 4 Road = [186,776)	57.1	147

Table 4.1 (cont.)

Rule no.	Premise	% Occurrence of shrimp farms	No. of samples observed
9	Land cover 1982 = 1 River = [0,139) Soil = 4 Road = [776,2256)	43.9	342
10	Land cover 1982 = 1 River = [0,139) Road = [2256,3293)	50.3	322
11	Land cover 1982 = 1 River = [0,139) Road = [3293,6488)	58.5	359
12	Land cover 1982 = 1 Road = [0,431) River = [139,225) Soil = 1,3,8 or 11	0.0	25
13	Land cover 1982 = 1 Road = [0,431) River = [139,225) Soil = 4,5,6,7 or 9	43.5	131
14	Land cover 1982 = 1 Road = [0,431) River = [225,473)	24.5	326
15	Land cover 1982 = 1 River = [473,1054] Road = [0,186) Soil = 1	20.7	29
16	Land cover 1982 = 1 River = [473,1054] Road = [0,186) Soil = 3,4,5,6,8,9 or 11	7.5	107
17	Land cover 1982 = 1 River = [473,1054] Road = [0,186) Soil = 7	100	2
18	Land cover 1982 = 1 River = [473,1054] Road = [186,431) Soil = 1,3,6,7,9 or 11	40.7	54

Table 4.1 (cont.)

Rule no.	Premise	% Occurrence of shrimp farms	No. of samples observed
19	Land cover 1982 = 1 River = [473,1054] Road = [186,431) Soil = 4,5 or 8	8.5	82
20	Land cover 1982 = 1 River = [139,1054] Road = [431,776)	30	420
21	Land cover 1982 = 1 Road = [776,1441) River = [139,649) Soil = 1,3,5,7,8,9 or 11 Zone = 0,1 or 3	31.1	74
22	Land cover 1982 = 1 Road = [776,1441) River = [139,649) Soil = 1,3,5,7,8,9 or 11 Zone = 2	12.5	96
23	Land cover 1982 = 1 Road = [776,144)] River = [139,649) Soil = 4	36.5	211
24	Land cover 1982 = 1 Road = [776,1441) Soil = 6 River = [139,225)	20.0	5
25	Land cover 1982 = 1 Road = [776,1441) Soil = 6 River = [225,649)	77.8	45
26	Land cover 1982 = 1 Road = [776,1441) River = [649,1054]	50.4	131
27	Land cover 1982 = 1 Road = [1441,2256) River = [139,473)	45.2	367

Table 4.1 (cont.)

Rule no.	Premise	% Occurrence of shrimp farms	No. of samples observed
28	Land cover 1982 = 1 Road = [1441,2256) River = [473,1054] Soil = 1,3,4,5,7,8,9 or11	54.7	179
29	Land cover 1982 = 1 Road = [1441,2256) River = [473,1054] Soil = 6 Zone = 0, 1 or 3	87.1	62
30	Land cover 1982 = 1 Road = [1441,2256) River = [473,1054] Soil = 6 Zone = 2	100	2
31	Land cover 1982 = 1 River = [139,341) Road = [2256,3293)	45.4	324
32	Land cover 1982 = 1 River = [139,341) Road = [3293,6488] Zone = 0	22.2	18
33	Land cover 1982 = 1 River = [139,341) Road = [3293,6488] Zone = 1,2 or 3	60.3	272
34	Land cover 1982 = 1 Road = [2256,6488] River = [341,649) Zone = 0 or 2	42.3	558
35	Land cover 1982 = 1 Road = [2256,6488] River = [341,649) Zone = 1 or 3	88.9	27
36	Land cover 1982 = 1 Road = [2256,6488] River = [649,1054]	26.4	201

Table 4.1 (cont.)

Rule no.	Premise	% Occurrence of shrimp farms	No. of samples observed
37	Land cover 1982 = 2 Road = [0,186) Soil = 1,3,4,5,7,8,9or11	0	72
38	Land cover 1982 = 2 Road = [0,186) Soil = 6	50	2
39	Land cover 1982 = 2 Road = [186,6488] Zone = 0 River = [0,69)	14.3	7
40	Land cover 1982 = 2 Road = [186,6488] Zone = 0 River = [69,473)	100	35
41	Land cover 1982 = 2 Road = [186,6488] Zone = 0 River = [473,1054]	50.0	4
42	Land cover 1982 = 2 Road = [186,6488] Zone = 1,2 or 3	50.0	10
43	Land cover 1982 = 3 River = [0,473) Road = [0,186)	1.2	244
44	Land cover 1982 = 3 River = [0,473) Soil = 1,3,4,5,8 or 11 Zone = 0 Road = [186,1441)	0.8	382
45	Land cover 1982 = 3 River = [0,473) Soil = 1,3,4,5,8 or 11 Zone = 0 Road = [1441,2256)	5.5	91

Table 4.1 (cont.)

Rule no.	Premise	% Occurrence of shrimp farms	No. of samples observed
46	Land cover 1982 = 3 River = [0,473) Road = [186,2256) Soil = 1,3,4,5,8 or 11 Zone = 1 or 3	42.9	7
47	Land cover 1982 = 3 River = [0,473) Road = [186,2256) Soil = 1,3,4,5,8 or 11 Zone = 2	8.6	81
48	Land cover 1982 = 3 River = [0,473) Road = [186,2256) Soil = 6,7 or 9	11.6	198
49	Land cover 1982 = 3 River = [0,473) Road = [2256,6488] Zone = 0	0.0	12
50	Land cover 1982 = 3 River = [0,473) Road = [2256,6488] Zone = 1,2,or 3	71.4	7
51	Land cover 1982 = 3 River = [473,1054] Road = [0,186)	7.9	126
52	Land cover 1982 = 3 Road = [186,431) River = [473,649)	13.7	51
53	Land cover 1982 = 3 Road = [186,431) River = [649,1054]	38.2	34
54	Land cover 1982 = 3 River = [473,1054] Road = [431,1441)	11.4	88

Table 4.1 (cont.)

Rule no.	Premise	% Occurrence of shrimp farms	No. of samples observed
55	Land cover 1982 = 3 River = [473,1054] Road = [1441,6488] Zone = 0 or 2	0.0	50
56	Land cover 1982 = 3 River = [473,1054] Road = [1441,6488] Zone = 1 or 3	20.0	5
57	Land cover 1982 = 4 River = [0,225) Road = [0,776) Soil = 1,3,4,6,8,9 or11	5.9	152
58	Land cover 1982 = 4 Soil = 5 River = [0,139) Zone = 0,1 or 3 Road = [0,431)	0.0	11
59	Land cover 1982 = 4 Soil = 5 River = [0,139) Zone = 0,1 or 3 Road = [431,776)	66.7	6
60	Land cover 1982 = 4 Road = [0,776) Soil = 5 River = [0,139) Zone = 2	0	33
61	Land cover 1982 = 4 Road = [0,776) Soil = 5 River = [139,225)	66.7	6
62	Land cover 1982 = 4 River = [0,225) Road = [0,776) Soil = 7	40.0	20
63	Land cover 1982 = 4 River = [0,225) Road = [776,2256)	38.2	34

Table 4.1 (cont.)

Rule no.	Premise	% Occurrence of shrimp farms	No. of samples observed
64	Land cover 1982 = 4 River = [0,225) Road = [2256,3293)	100	6
65	Land cover 1982 = 4 River = [0,225) Road = [3293,6488]	50.0	4
66	Land cover 1982 = 4 River = [225,1054] Road = [0,186) Soil = 1 or 6	64.7	17
67	Land cover 1982 = 4 River = [225,1054] Road = [0,186) Soil = 3,4,5,7,8 or11	0.0	57
68	Land cover 1982 = 4 Road = [0,186) Soil = 9 River = [225,473)	50.0	4
69	Land cover 1982 = 4 Road = [0,186) Soil = 9 River = [473,1054]	3.8	26
70	Land cover 1982 = 4 River = [225,1054] Road = [186,2256)	44.1	118
71	Land cover 1982 = 4 River = [225,1054] Road = [2256,3293)	92.3	26
72	Land cover 1982 = 4 Road = [3293,6488] River = [225,649)	100	3
73	Land cover 1982 = 4 Road = [3293,6488] River = [649,1054]	11.1	9
74	Land cover 1982 = 5	43.4	339
75	Land cover 1982 = 6	1.2	84

Legend :

Land cover : 1 = Mangrove 4 = Swamp and/or mangrove clearing

2 = Standing tree 5 = shrimp farm

3 = Paddy field 6 = Grass land

Zone : Legal status of the area

0 = Area where exclude from zoning area

1 = Preservation zone

2 = Economic Zone A

3 = Economic Zone B

Soil : Soil type

Symbol	Soil series	Texture	Drainage	Permeability	Suitability
1	Rayong	sand	over fast	fast	grass land,coconut,scrub
2	Ban Thon	sandy loam	moderate	moderate	scrub
			good	fast	
3	1 & 2 association				
4	Tha Chin	clay	very bad	fast	nipa,mangrove
5	Samut Prakan	clay	bad	slow	rice
6	Cha-am	clay	bad	---	rice, peanut, watermelon
7	Ongkharak	clay/ loamy clay	bad	slow	rice
8	Kleang	clay/ sandy clay	bad	slow	rice
9	Chonburi	clayey loam, sandy phase	rather bad	moderate fast	rice, watermelon
10	Renu	---	---	---	---
11	Pak Chan	clay/ clayey loam	good	moderate	forest, corn, orchard

The accuracy when using rules obtained from the tree to classify data in the test data set were summarized in table 4.2.

Table 4.2 The accuracy of classification rules.

Land cover variable	% Accuracy			
	1986	1988	1990	1992
Fixed as landuse 1982	94.4	75.9	73	71.1
Changed by relating to following every other past year	94.4	77.9	78.8	78.6

Using the Knowledge Driven Model to Classify Data

This study used program KBCLASS which is called for 'knowledge based classifier' to classify area converted to shrimp farm. This program was developed by Supichai Tangjaitrong (Chulalongkorn University, private communication, 1994) to classify data according to a set of production rules in a knowledge base. As described by Supichai, the program use production rule stored in a knowledge base as its guide to perform the classification. The format of rules is a generic IF...THEN statement that is the same as a format of rules converted from decision tree by knowledge SEEKER. (Fig.4.3)

File Format

RULE_*i* IF

var_x1 operator operand
var_x2 operator operand
. . . . .
. . . . .
THEN

*var*_Y = *value belief*

Meaning of variable

Variable Name	Descriptions
*var*_X*i*	Name of the *i* th variable
operator	Logical operator: =, <= , >= , ≠
operand	Value of the condition: e.g., 5, fat, often
*var*_Y	Name of the classified result
value	Value of the classified result
belief	Grad of belief in the value

An example file

RULE_1 IF
Land cover = 1
River = [0,34)
Road = [0,186]
Soil = 1,3,6,8 or 11
THEN
Landuse 1988 = shrimp farm 9.1%
Landuse 1988 = others 90.9%
– – – –

Figure 4.3 Format of the knowledge base

The format of variables information in the knowledge base must be specified in another text file. (Fig.4.4)

File format

no_of_variables
var_name img_name img_channel vaugeness_factor weight missing value

Meaning of variables

Variable name	Descriptions
no_of_variables	Number of variables in the knowledge base
var_name	Variable name specifying data layer
img_name	Image file name of the associated variable
img_channel	Channel number of the image for *var_name* stored in the image file
vagueness_factor	Value of parameter *v* as use in equation 4.1
weight	Contribution weight
missing_value	A value to indicate that data element is missing or unknown

An example file

5					
river	riverdis.img	1	1	2	*
land cover82	82map.img	1	-1	1	*
road	roadid.img	1	1	1	*
soil	soilmap.img	1	-1	1	*
zone	zonemap.img	1	-1	1	*

Figure 4.4 Information used to specify data attribute

Five images were required for the five independent variables in the knowledge base. Every variable were assigned equal weight for classification except *distance from source of water*. From pre-classification, higher accuracy obtained when assigned weight of *distance from source of water* variable as twice of other variables. Furthermore, *distance from source of water and distance from road* were assigned as *fuzzy set* since they have loosely define boundaries. The others which boundaries are definite were assigned as *crisp set*. The equation for calculating fuzzy memberships is defined as:

$$\text{Membership} = \frac{1/(1+(x-c)^2)}{v} \qquad \text{....................................(4.1)}$$

where x = observed value

c = constant of the rule

v = vagueness factor : assign $v = -1$ for crisp variables

$v > 0$ for fuzzy variables

The program read and parsed the rules into classification lookup table (Fig. 4.5), then read each row of image classified. The supporting weight of each variable in the rules was considered from the contribution weight of every variable in the rules. The weight was computed by :

$$W_{(r,j)} = w_j \Big/ \sum_{i=1}^{N} w_i \qquad \text{.......................................(4.2)}$$

where $W_{(r,j)}$ is the supporting weight of variable j in rule r

w_i or w_j is the contribution weight of variable (i or j) given by users

N is the number of variables in rule r

RULE_1 IF
*Land cover*1982 = 1
River = [0,341)
Road = [0,186)
Soil = 1,3,6,8 or 11
THEN
*Landuse*1988 = Shrimp 9.1%
*Landuse*1988= Others 90.9%
RULE_2 IF
Land cover 1982 =1
River = [0,341)
Road = [186,431)
Soil = 6,8 or 9
Zone = 3
THEN
*Landuse*1988 = Shrimp 66.7%
*Landuse*1988= Others 33.3%
\- - -

parse rules into the classification table

Rule #	Variable	Operator	Lower	Upper	Weight
1	*Land cover82*	=	1	0	0.2
2	*Land cover82*	=	1	0	0.16

1	*River*	[...)	0	341	0.4
2	*River*	[...)	0	341	0.33

1	*Road*	[...)	0	186	0.2
2	*Road*	[...)	186	431	0.16

Remark :. The contribution weight of River is assigned as double of other variables.

Figure 4.5 Parsing classification rules into a classification table

On classifying a pixel, the program compare each variables with those list in the classification table, then add the supporting weight to the level of supporting evidence of the considering rule. When every record in the classification has been considered, KBCLASS sorts the level of supporting evidence and fires the rule that having the maximum level.

KBCLASS gives three type of classification result, (1) the rule number that data satisfy, (2) level of belief in the rule, and (3) level of supporting evidence for firing the rule. This study used only the rule numbers of the classification result which stored in output images with IDRISI. format. Plate 4.1 shows a sample of simulated image of shrimp farm in 1992 which has the possibility of landuse change into shrimp farm over than 50%.

Justifying the Model Result

Method and Result

From the result in previous section, each pixel in the output images contained the rule number that related to the possibility of landuse change into shrimp farm of each pixel. In order to verification of the knowledge driven model, the next step changed the rule number of each pixels to the level of shrimp farm occurrence according to setting level by module ASSIGN in program IDRISI (Table 4.3).

Table 4.3 Setting level for converting the rule numbers

% possibility of conversion	Level
$c = 0$	1
$0 < c \leq 25$	2
$25 < c \leq 50$	3
$50 < c \leq 75$	4
$75 < c \leq 100$	5

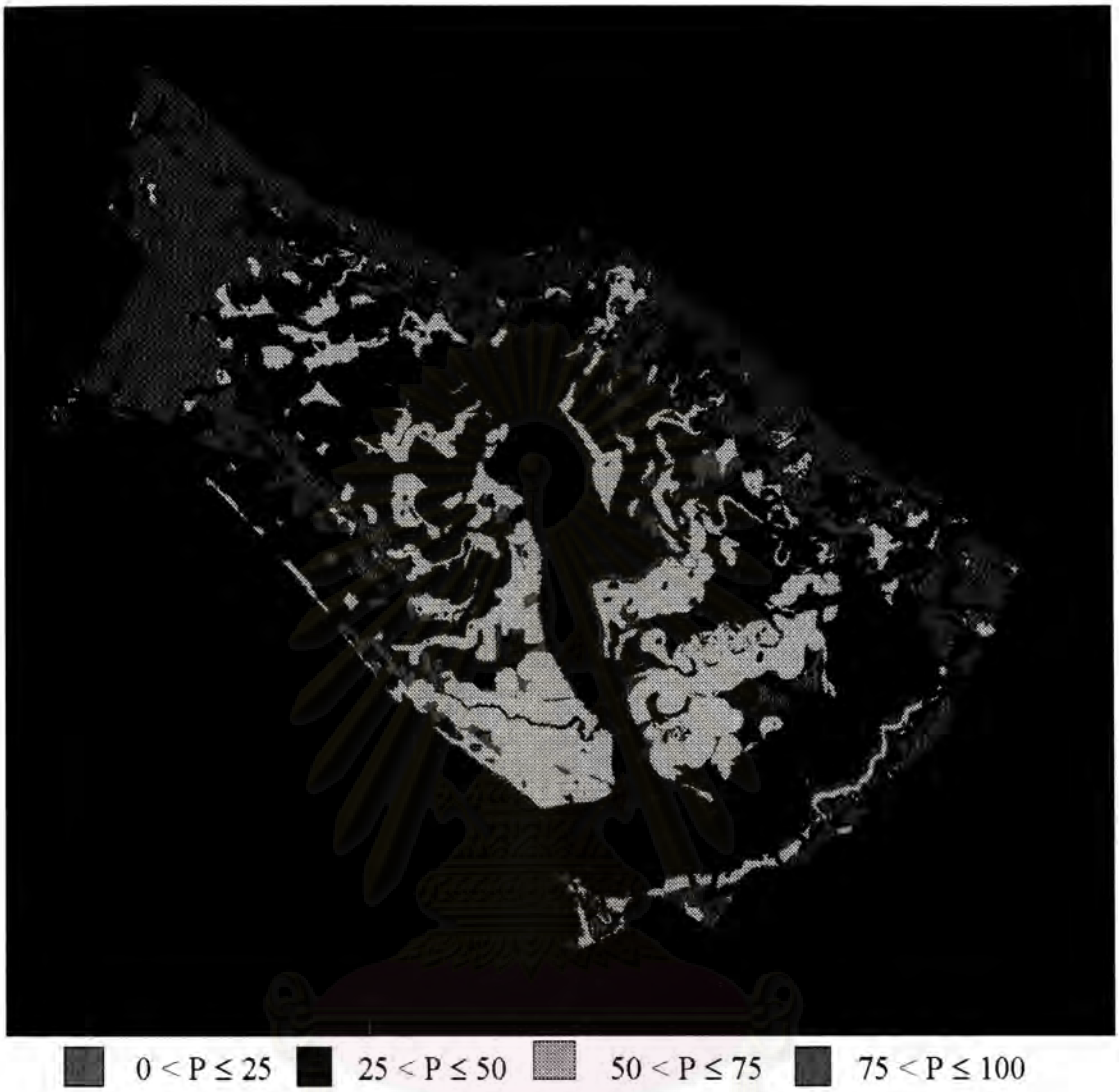

Plate 4.1 Possibility Image of landuse change into shrimp farm in 1992.

At this stage, each pixel of output images contained the level of shrimp farm occurrence. The output images were then compared with the image of shrimp farms of the same year by module CROSSTAB in IDRISI. This resulted in cross-tabulations table listing the tabulation totals (Table 4.4 and 4.5) and also shown in Figure 4.6 and 4.7. From comparison of determining shrimp farm area by classification with by simulation, the results also confirmed validation of the developing model (Table 4.6).

Table 4.4 Cross-tabulation of the images of shrimp farms against the simulated images (fixed land cover 1982 as independent variable).

Simulated image	Classified Landuse	% pixel of shrimp farms in each setting level						
		0	1	2	3	4	5	TOTAL
			p=0	$0<p\leq 25$	$25<p\leq 50$	$50<p\leq 75$	$75<p\leq 100$	
1986	Other	810,026	56,356	189,390	6,219	1,681	5,036	1,068,708
	Shrimp farm	736	2,230	15,853	2,812	3,305	6,356	31,292
		0.09%	**3.81%**	**7.72%**	**31.14%**	**66.29%**	_55.79%_*	**2.84%**
1988	Other	796,966	15,101	89,904	72,604	29,024	1,391	1,004,990
	Shrimp farm	389	1,055	15,095	33,261	41,391	3,819	95,010
		0.05%	**6.53%**	**14.38%**	**31.42%**	**58.78%**	_73.30%_*	**8.64%**
1990	Other	797,925	9,956	72,058	37,869	51,177	4,415	973,400
	Shrimp farm	2,010	1,608	10,298	24,359	71,912	16,413	126,600
		0.25%	**13.91%**	**12.50%**	**39.14%**	**58.42%**	**78.80%**	**11.51%**
1992	Other	805,036	22,890	101,814	18,958	38,841	2,381	989,920
	Shrimp farm	3,460	959	14,079	5,982	69,778	15,822	110,080
		0.43%	**4.02%**	**12.15%**	_23.99%_*	**64.24%**	**86.92%**	**10.01%**

Table 4.5 Cross-tabulation of the images of shrimp farms against the simulated images (changed land cover by relating to the following every other past year)

Simulated image	Land cover variable	Classified Landuse	% pixel of shrimp farms in each setting level						
			0	1	2	3	4	5	TOTAL
				p=0	$0<p\leq 25$	$25<p\leq 50$	$50<p\leq 75$	$75<p\leq 100$	
1988	**Land cover**	Other	802,979	9,617	101,574	82,065	4,642	4,131	1,004,990
	1986	Shrimp farm	195	645	16,951	46,559	7,553	23,107	95,010
			0.02%	**6.29%**	**14.30%**	**36.20%**	**61.94%**	**84.83%**	**8.64%**
1990	**Land cover**	Other	805,226	20,793	74,455	46,066	11,683	15,177	973,400
	1988	Shrimp farm	362	1,662	13,706	26,901	14,644	69,325	126,600
			0.04%	**7.40%**	**15.55%**	**36.87%**	**55.62%**	**82.04%**	**11.51%**
1992	**Land cover**	Other	797,774	10,531	69,064	75,348	34,695	2,508	989,920
	1990	Shrimp farm	509	749	11,075	50,678	42,026	5,043	110,080
			0.06%	**6.64%**	**13.82%**	**40.21%**	**54.78%**	_66.79%_*	**10.01%**

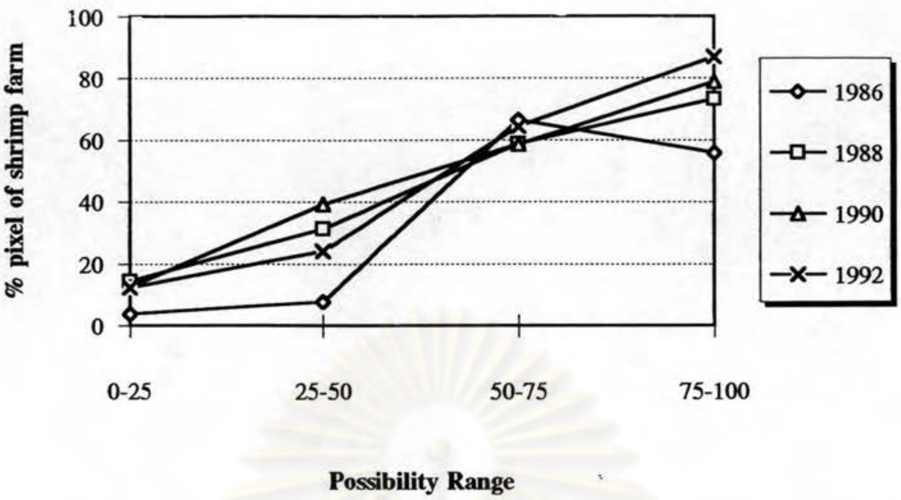

Figure 4.6 Result of cross-tabulation of shrimp farm images against the simulated images (fixed land cover 1982 as independent variable).

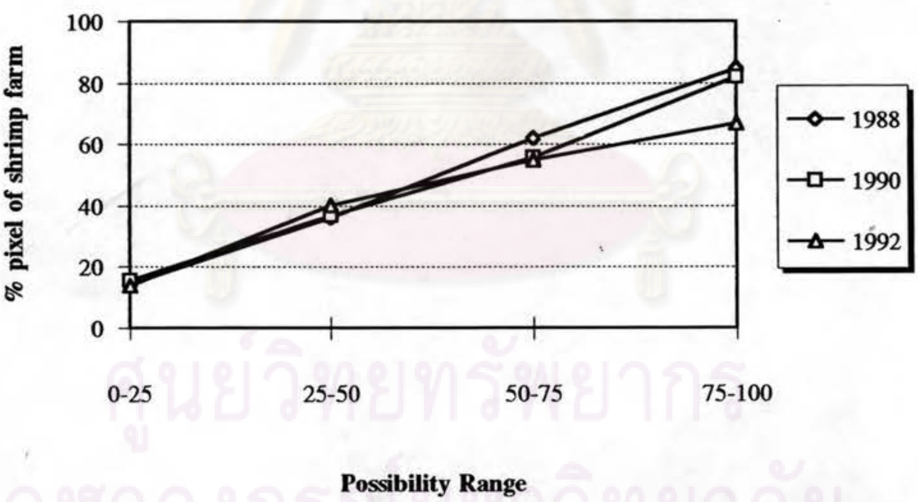

Figure 4.7 Result of cross-tabulation of shrimp farm images against the simulated images (changed land cover by relating to the following every other past year)

Table 4.6 Area of shrimp farm comparison between determining by classification with by simulation.

Year	Shrimp farm area(km^2)	
	Classification	Simulation
1982	16.57	-
1986	29.98	28.16
1988	91.04	85.51
1990	121.31	113.94
1992	105.48	99.07

Summary of the procedures about developing and justifying the model are shown in Table 4.7

Table 4.7 Procedures for developing and justifying the model

Step	Procedure	Program	Remark
1	Sampling train set data	SAMPLE (IDRISI)	Extract data of each variable from images
2	Develop a classification tree and convert to rule-based representation	Knowledge SEEKER	Use train set data to develop the tree and use the test set data to determine the tree accuracy
3	Classify possibility that a pixel change to shrimp farm	KBCLASS	Use rules (knowledge driven model) to generate simulated images
4	Verification of the model	ASSIGN, CROSSTAB (IDRISI)	Change the rule number of each pixel to the level of shrimp farm occurrence

Discussion

- From table 4.4, most of percentages of shrimp farms in each level from both ways of using land cover variable fall in their range of criteria of shrimp farm occurrence. These results ensuring that the developed knowledge driven model could simulated close to the real system being in the condition of conversion the land to shrimp farms. Comparing the results between two ways of using land cover variables, there were slightly difference which means that the land cover might not have as high influence on decision making process of the investors. The decision, therefor, rely on other physical factors, such as, source of water, distance from road, soil suitability. The findings resemble the result of investigation by Boonchana Klankumsorn *et al.*(1989).

- The production rules in table 4.1 generated base on change of landuse in a long time period (1982-1992) showed high validation result. Therefore, this set of rules could be represented as desirous conditions of the investors to convert the land to shrimp farm which will be advantage for coastal landuse planing.

From the rules which have percent occurrence of shrimp farm over than sixty and have number of samples observed more than fifteen (rule no. 29, 33, 35, 40, 66 and 71), it could be concluded the condition of area which have high possibility of change into shrimp farm. Type of landuse in all rules was mangrove or swamp except in the rule number 40 which was standing tree. Standing tree in that rule might be resulted from error in classification process. Really, it might be mangrove. Distance from source of water were in the range of 69-1,054 meter. For distance from road, the maximum distance was 6,488 meter. Distance from road will be shorter than in the rules if can add new roads in digitizing process. The topographical map used in the study was produced since 1981.

- In the attempt to define the interval boundaries in level of shrimp farms occurrence with lower value (i.e., $c=0$, $0<c<10$, $10<c<20$,...), the accuracy of the percent pixel of shrimp farm decreased which indicated that the model have their limited level of validation.

- The expansion of shrimp farms dependent to time, then the area with suitable factors might not be converted into shrimp farms. Hence, the percent of shrimp farms in setting level ' 6 ' in 1986 and 1988 which were earlier year of shrimp farm expansion did not fell in the range of the criteria.

- Validation of the knowledge driven model associate with *uncertainty*. Uncertainty could arise from several different sources, including *uncertainty knowledge*, *uncertainty data, incomplete information* and *randomness* (Rolston, 1988; Dym and Levitt, 1991). In this study, uncertainty might also due to all these four causes.

Uncertainty knowledge might result from doubt on the part of knowledge driven model. The rule-base representation might be imprecise since it was unable to sharply define the idea or the concept behind the information. This problem involved the classification process through a tree structure and the alternative selection of variable in the step of developing rules of the program KnowledgeSEEKER.

Each data layers might be inaccurate due to many reasons. In land cover or landuse, the error could be in the classification process of remotely sensed data. Also in other data layer which used digitizer as instrument in capture, the quality of digitizer (precision and accuracy), rate of sampling frequency (especially when digitized a curve) or even the reliability of the original map might lead to uncertain data.

Information might also be incomplete because of the limited resource in data acquisition. The topographical map had not been regularly updated, therefore new infra-structures (i.e., road, irrigation) were not included in the available maps. Furthermore, decision making process of the investors certainly could rely on other factors not included in this study, such as, government land reformation, quality of water etc.. In practice, these data might be rather difficult to reach.

Information gathered might not be reliable because it was unable to obtain it with sufficient precision. In this case, the study recognized the inherent randomness. Remotely sensed data often contains spatial noise that resulted from a number of reasons, such as atmospheric haze, specula reflection or uneven illumination (due to topography or uneven lighting). These noise might effect the accuracy in processing image classification.

- The knowledge driven model generated in this study had been developed based on the data layer collected only from the study area. The model need to be validated in other area which might have different factors of site selection for shrimp farms in order to reduce their characteristic on site specific limitation.

Conclusion

The production rules in table 4.1 might be useful to simulate the future landuse in other coastal areas at the initial stage of shrimp farm expansion. Data layers for entering as independent variables are commonly available. Information gained will be advantage for decision maker for landuse planning and management.

The concept of developing knowledge driven model for representing the behavior of environmental problems concerning the spatial information is not limited to shrimp farm expansion problem, therefore this concept would below the ability of decision makers to manipulate the information using production rules for various fields of environmental problems and various aspects of environmental management.