JIMMA UNIVERSITY

COLLEGE OF NATURAL SCIENCES

DEPARTMENT OF INFORMATION SCIENCE

DESIGNING A CASE BASED REASONING SYSTEM FOR IDENTIFYING SUITABLE LAND FOR CEREAL CROPS

BY: - TARIKU MOHAMMED

June, 2016

Jimma, Ethiopia

JIMMA UNIVERSITY
COLLEGE OF NATURAL SCIENCES
DEPARTMENT OF INFORMATION SCIENCE

DESIGNING A CASE BASED REASONING SYSTEM FOR IDENTIFYING SUITABLE LAND FOR CEREAL CROPS

A thesis submitted in partial fulfillment for the award of the degree of master of Information Science (Information and Knowledge Management)

Principal Advisor: -Million Meshesha (PhD)

Co-Advisor: - Chala Diriba (MSc)

June, 2016

Jimma, Ethiopia

JIMMA UNIVERSITY

COLLEGE OF NATURAL SCIENCE

DEPARTMENT OF INFORMATION SCIENCE

DESIGNING A CASE BASED REASONING SYSTEM FOR IDENTIFYING SUITABLE LAND FOR CEREAL CROPS

By: - TARIKU MOHAMMED

Name of Member of the Examiner Board

Name	Title	Signature	Date
___________	Chairperson	___________	____________
___________	Principal Advisor	___________	____________
___________	Co-Advisor	___________	____________
___________	External Examiner	___________	____________
___________	Internal Examiner	___________	____________

DEDICATION

This thesis is dedicated to my mother for her encouragements and hopes to see me more capable and competent in my endeavors.

DECLARATION / Approval Sheet

This thesis is my original work and has not been submitted as a partial requirement for a degree in any university. All the material sources used in this work are duly acknowledged.

____

TARIKU MOHAMMED
June, 2016

This thesis has been submitted to the department for examination with our approval as university advisors:

Principal Advisor: Million Meshesha (PhD)……………….
Co -Advisor: Chala Diriba (MSC) ……………

June, 2016

Acknowledgement

I am deeply indebted to my advisor, Dr. Million Meshesha, and Mr. Chala Diriba for their commitment, patience, encouragement and guidance throughout the research work.

I wish to thank my beloved friends and classmates who have one way or another contributed to the success of my research. I want to thank all the department staffs for their kindly guidance and support in each step of this research preparation. I also want to extend my gratitude to Ethiopian institute of agricultural research center office employees for their cooperation and hospitality.

Above all, I thank God who surrounded me with so many wonderful people in all my life.

Contents

List of Table

List of Figure

List of Acronyms and Abbreviations

ALES: Automated Land Evaluation System

AI: Artificial Intelligence,

CBR: Case based Reasoning

DA: development agents

EIAR: Ethiopia Institute of Agriculture

FAO: Food and Agriculture Organization

GIS: Geographic Information Systems

GUI: Graphical User Interface

ISLE: Intelligent System for Land Evaluation

JCOLIBRI: Java Case and Ontology Libraries Integration for Building Reasoning infrastructures

KBS: Knowledge Based Systems

LECS: Land Evaluation Computer System

RBS: Rule Based Reasoning

Abstract

Agriculture is the basis of the Ethiopian economy. The country's economic development depends, in large part on sustainable improvements in agriculture However; its productivity is not kept pace with population growth. In Ethiopia, there is scarcity of agriculture experts; the available once may not be accessible to every farmer. By having an agricultural knowledge based system, the problem of experts in Agriculture can be reduced. It is therefore the aim of this study to develop a Case-based system that enable to make proper decision in the process of land and cereal crop matching so as to select suitable cereal crops for the farm unit under cultivation. This research was conducted following design science approach. Purposive sampling technique was used to select 10 domain experts for knowledge acquisition. To develop land cereal crop matching case based reasoning system, important knowledge was acquired through interview and document analysis. The acquired knowledge was modeled using hierarchical decision tree. jCOLIBRI and ArcGIS was used for developing case-based system (CBS). The developed CBS provide a method in the process of land cereal crop matching proposing a solution to a new problem or providing relevant experiences to the decision maker. GIS tools were used for preparing, handling and generating spatial and non-spatial information as a tabular form for CBR tools. The prototype of CBRLCCM system utilizes multiple knowledge to determine suitable, optimal cereal crops for a farm unit. This knowledge consists of representative cases to reflect physical, economic, environmental and social factors that affect the choice of land use for cereal crops. Domain experts' evaluation by visual interaction with the prototype achieves 84% user acceptance. In addition, performance of the prototype system was evaluated using case testing method which scores f-Measure of 76%. This system is promising to develop an applicable system for improving the productivity of farmers by assisting agricultural expert and development agents who advise farmers on their daily needs. However, further study should be done to include inputs from climate prediction model so as to predict future land use choice.

Chapter one

Introduction

1.1. Background of the study

Agriculture is the basis of the Ethiopian economy. The country's economic development depends mainly on sustainable improvements in agriculture. In addition, Agricultural Development Led Industrialization strategy is expected to play an important role in generating surplus capital to accelerate the overall socio-economic development of the country. A high rate of agricultural growth has far reaching positive implications for economic development of low income countries in terms of enhancing employment and rushing poverty reduction (Mellor and Dorosh, 2010).

Agricultural productivity in Ethiopia has not kept pace with population growth, and the country is now in a worse position nutritionally than it was 30 years ago: food production has achieved a growth of about 2.5% per year, while population has increased at a rate of over 3% per year (Hailu, 2012). Even though 85 percent of the country's population lives in the rural areas, the performance of the agricultural sector in Ethiopia has remained weak and it is heavily influenced by weather condition (Mulat, 2009). Furthermore, the productivity of the sub-sector is decreasing as a result of poor management systems, shortage of skilled experts who provide advice for farmers at Woreda level (Ejigu, 2012). Despite the importance of agriculture in its economy, Ethiopia has been a food deficit country since the early 1970s (FDRE, 2002). Therefore, the performance of Ethiopian agriculture reveals that it has been unable to produce sufficient quantities to feed the country's rapidly growing human population (Kassa and Degnet, 2004).

Cereal production and marketing are the means of livelihood for millions of households in Ethiopia. It is the single largest sub-sector within Ethiopia's agriculture, far exceeding all others in terms of its share in rural employment, agricultural land use, calorie intake, and contribution to national income. The contribution of cereals to national income is also large (Rashid, 2010). According to available estimates, cereal crops contribution to agricultural value added is 65 percent (Diao & Alejandro, 2010) which translates to about 30 percent of GDP.

Land evaluation is the procedure of measuring the capability of land for specific land use purposes (FAO, 2007). The process of land assessment forms a vital part of land management and helps in preparing an effective land use plan. The purpose of land use planning is to utilize available resources in a land unit to maximize economic returns without degrading environmental values. Multi-disciplinary knowledge is required for a comprehensive evaluation of a farm unit's resources and to identify its true potentials and limitations. In case of agricultural land use, the results of land evaluation can help in choosing the most suitable land use and a profitable crop for a farm unit.

In any agricultural production system, accumulation and integration of related knowledge and information from many diverse sources play an important role. In practice, there is no substitute for knowledge and experience in identifying problems and choosing the most appropriate management technique for addressing them (Calu, 2009). Therefore, agriculture requiring information and application of knowledge from different interacting fields of science and engineering to make a suitable decision making that in turn depends on interplay of these data and knowledge. This needs agricultural specializations and technical awareness in farmer or a human expert to help the farmers in decision making (Abu-Naser *et al.*, 2008). Therefore, when there is a shortage of professionals, intelligent system can help the farmer for decision making (Ejigu, 2012).

In agriculture, applications of intelligent system are mainly found in the area of the crop production management, pest management, diagnostic systems, overall planning systems as well as economical decision making (Sarma et al., 2010). CBR is a family of AI techniques that simulates the human behavior in solving a new problem. Thus in CBR, reasoning is based on remembering. When confronted with a new problem, it is natural for a human problem-solver to look into their memory to find previous similar instances for help (Shi & Yeh, 2001). In CBR approach, new problems are handled by remembering old similar ones and moving forward from there. Referencing to previous cases is beneficial in dealing with situations that persist. However, CBR technique is based on two tenets about the structure of problem solving process. The first one is similar problems have similar solutions and the second one is future problems may have similar to current problems (Leake, 1996). When the two assumptions hold, CBR becomes an effective reasoning strategy. CBR has been used as a part of spatial reasoning system to assist spatial/non-spatial retrieval mechanism. Most of these approaches consist of CBR and GIS to form a hybrid system to solve spatial problems. In agricultural land evaluation spatial and non-spatial information are vital in representing the features of a case (Shi and Yeh, 2001). CBR as its name indicates, uses cases to reason about a given problem. In its problem solving process, it reuses old similar cases to understand the problem, suggest a solution, and/or to keep it from failure. Cases record the past, giving us and computer a way to make assumptions about the present. The CBR process (4RE cycle) seems feasible for problem solving situations. CBR technique is a promising way to build and apply more powerful land use planning support system in contrast to rule-based or model-based reasoning, CBR uses concrete knowledge directly and its inference is basically the processes of retrieval and adaptation (Shi & Yeh, 2001).

GIS has potentiality to provide a rational, objective and non-biased approach on making decisions in agriculture land suitability evaluation (Chuong, 2008). Moreover, GIS has the ability to integrate variety of geographic technologies such as Global Position System (GPS) and Remote Sensing. This function also provides an interface between GIS and other modeling software which can integrate non-spatial data. An example, suitability maps can be integrated with non-spatial data, such as socio-economic data to model the effect of these data on the land use. This function of GIS can save time and cost in the evaluation of land use options compared with conventional methods (Burrough, McDonnell, et al., 2008). Currently GIS techniques have been used in many land suitability studies in the process of land evaluation based on Guideline of land evaluation for agriculture (FAO, 2007). GIS can generate, store and display spatial information related to a case. Therefore, a system that combines CBR and GIS can very important to agricultural land use planning process and decision maker in such a combined system CBR will provide the method for decision support in proposing a solution to a new problem or providing relevant experiences to the decision maker *(*Shi and Yeh, 2001). In the handling of spatial data, GIS can be an analysis, data generator, a database management system and a visualization tool as a fundamental part of agriculture land use planning system. According to Holt & Benwell (1996), lack of analytical and modeling functionality is a major deficiency of current spatial information systems in land use planning. Hence, there is a perceived need to integrate spatial information systems with additional analytical approaches to overcome this deficiency. Therefore, the current study was focus on combine GIS results with CBR that can be used for agricultural land use decision support system for cereal crop suitability assessment.

1.2 Statement of the problem

Ethiopia has great agricultural potential because of its vast areas of fertile land, diverse climate, generally adequate rainfall, and large labor pool. About 73.6 million hectors' (66%) of the country's area is potentially suitable for agriculture however; low crop production in Ethiopia is noticeable due to the farmers does not know what kind of crop they should cultivate for their farm units instead, they are cultivating crops in traditional methods rather than following the scientific approach (Fasil, 2002).

Change is a continuous process. Resources, ecosystem, biophysical environment, and land use planning on the surface of the earth undergo changes over time. Climatic constraints are one of a main factor for agricultural land systems, either by restricting ecological processes such as plant growth rate, soil erosion, recycling of nutrients and land quality, or by limiting management activities, especially those related to the timing of specific practices, such as ploughing, sowing or harvesting (Brown *et al*., 2008).

According to Teklu (2005) land suitability analysis for Sorghum and Maize Crops Using GIS Approach in Dera Wereda, Ethiopia. Results showed that the proportion of the total land was marginally suitable for sorghum production (68.78%) and moderately suitable for maize (81.26%) production in the Wereda. On the same study area with different period of time Ebrahim (2014) conducted on land Suitability assessment for Sorghum and Maize Crops Using GIS Approach in Dera Wereda, Ethiopia. Ebrahim noted that the largest proportion of the land was only moderately suitable for maize (70.67%), but marginally suitable for sorghum production (59.75%) in the Wereda. According to Teklu and Ebrahim studies on the same area with different period of time indicated that there is different outcome for suitable analysis of the same type of crops based on the criteria of FAO Guideline.

Therefore, factors which are important in decision making of land suitability analysis can change due to different factors it needs further research to understand the dynamic pattern that have not been totally clarified and identified. The main restricting factors for good land suitability analysis in GIS don’t consider the dynamic condition of different factor that important for decision support system in agricultural land use planning. GIS can evaluate based on the existing or current condition of land quality and environment factor however this conditions are dynamic it can be change over time.

Scholar who studied on land suitability evaluation using GIS application such as Rabia, (2012); Pirbalouti et al. (2015); Gatheru & Maingi, (2010) determine physical land suitability for different crops based on the criteria for physical land evaluation was separately analyzed for their suitability for supporting the crops based on the FAO Guideline crop requirements specified and land evaluation. The major data sources were climatic, soil and topology data which have been considered to undertake suitability assessments.

However, all of these systems relied heavily upon the physical land evaluation, neglecting social and economic factors. Socioeconomic conditions of farmers such as farmer’s skills, size of land holding, current market price for a crop, its market demand, neighboring land uses and government policies and regulations have a considerable effect on land use decisions. None of these systems took current knowledge (market price, demand for a crop produce, weather forecast, government policies) into consideration. Instead, most of these systems have a static knowledge base which does not take this dynamic information. It is however agreed that it is impossible to preconceive all the factors necessary for a decision-making problem. So, provisions need to be made to incorporate new, additional factors into the decision support system.

To alleviate these shortcomings, FAO worked together to prepare a common framework so as to provides a description of land qualities and guidelines for physical and economic land evaluation. FAO (2007) framework are the inclusion of socio-economic and local political factors in the land evaluation process, assessment of land with respect to specific land use types, a multi-disciplinary approach to land evaluation, setting priorities for environmentally sustainable land uses and providing alternative land use options for users.

In this respect, in the process of agricultural land use planning support system, CBR can help in making consistent decisions through handling previous experiences or the memory of how similar cases had been solved in the past. In addition, CBR can take account for dynamic as well as static knowledge which cannot include in the GIS features. This is the type of problem solving cycle which is most suitable for CBR. Where as in the handling of spatial data, GIS can be an analysis, data generator, a database management system and a visualization tool as a fundamental part in the process of agricultural land use planning system. Therefore, a system that combines GIS result with CBR can very helpful in the agricultural land use planning. In such a combined system CBR will provide the method for consistent decision support in proposing a solution to a new problem.

Most geo-information tools do not readily fit the changing needs of the planning profession. Land use planning support systems are generally regarded as systems in which technologies dedicated to the agricultural experts are brought together. Land use planning support system specifically support the whole of or some part of a unique planning task (Geertman & Stillwell, 2004). Currently Land use planning support system has arrived in concept and in application. It began in response to planners' fascination with GIS as a reminder that geographically referenced information and spatial analysis techniques alone cannot adequately support all of the planning (Harris, 1989). Its models include expert familiar tools (such as GIS, land allocation models) for conducting analysis, projecting future conditions, modeling spatial interaction and newly emerging artificial intelligence techniques such as artificial neural networks, fuzzy-logic, case based reasoning, data mining, analytic hierarchy process and other reasoning procedures.

In Ethiopia, agriculture experts are not always available, may not be accessible to every farmer. Due to a shortage of extension agents, each agent has to serve on average 1090 farmers (Kassa and Degnet, 2012). By having an agricultural expert system, the problem of scarcity or shortage of experts, especially in Agriculture, can be reduced (Rafea et al., 1998). Therefore, fewer experts in agriculture can be replaced with expert system. Furthermore, it can be store much of the information that an expert needs to make decisions and can make them on hand for others; Therefore, the notion of knowledge based agriculture has an adequate prospective to improve the agricultural production. Moreover, Abrham (2009) noted that, farmers worry about a sustainability of partnership, limited option of available technologies and lack of appropriate and timely decision on tested technologies that improve farmer's productivity. According to Abrham, the lack of trained agents (experts), staff turnover for training purpose, weak exchange of information has been affecting farmer's income and their economy.

Agricultural land use planning domain has been experiencing various problems. If GIS are to be used to solve these problems, then systems with a diverse range of analytical functions are needed. Hence there have been attempts and approaches to integrate additional analytical approaches with GIS. Therefore, CBR acting as an important role in spatial reasoning system to assist spatial and non-spatial retrieval mechanism.

The aim of this research is therefore to elicit tacit knowledge of land evaluator (crop expert) and explore ways to develop a CBR system that can guide farmers to choose a suitable cereal crops for their land unit.

To this end, the study explores and answers the following research questions to come up with solution for the above mentioned problems.

- What kinds of knowledge are required for matching land cereal crop decision making?
- What are the most suitable techniques and ways for knowledge acquisition, knowledge modeling and knowledge representation?
- To what extent the proposed prototype CBR performs and get user acceptance?

1.3 Objective of the study

1.3.1 General objective

The general objective of the study is to design and develop CBR that enable to make proper decision in the process of matching land with cereal crops so as to select suitable farm unit under cultivation for a given cereal crops.

1.3.2. Specific Objectives

To achieve the general objective of this study, the following lists of specific objectives are attempted.

- To acquire the domain knowledge needed for developing the proposed CBR system.
- To model and represent the cases from different data (spatial and non-spatial) knowledge involved in identifying suitable land for cereal crop.
- To develop a prototype case based reasoning system.
- To evaluate system performance and user acceptance of the prototype CBR.

1.4 . Scope and limitation of the study

There are five major cereals crops, (maize, tef, wheat, barley and Sorghum) which are produced in Ethiopia (Rashid, 2010). The scope of this study was focus on developing a prototype CBR that enable to make proper decision in the process of land cereal crop matching so as to select suitable major cereal crops for the farm unit under cultivation at Ethiopia institute of agricultural research center which are found in Addis Ababa. Furthermore, in an attempt to develop a CBR the study was performing the task of knowledge engineering such as knowledge acquisition, knowledge modeling, knowledge representation and construction towards CBS development that addresses the land cereal crop matching. The quality of decisions depends on the quality and the variety of cases used in the system. Due to time constraints, the case collected from the crop expert reflects only a few factors, which serve the purpose of validating a case-based approach to agricultural land evaluation. There were also problems to get localized attributes for all type of cereal crops requirement for agricultural land use planning from experts. Due to this few cereal crops were not including on these study.

1.5. Significance of the study

The findings of this study resolve agricultural researchers or experts in such a way that experts and researchers can get easily and timely access to identification suitable land for cereal crops from knowledge base which stores the cases that experts used to solve problem.

By replicating this system to different cereal crop growing Woredas, the system can increase production of cereal crops in agricultural sectors through provision of adequate knowledge from stored cases in the process of identifying suitable land for cereal crop approach. Therefore, those development agents who assist farmers can consult and access the system to identifying suitable land for cereal crops and make appropriate decisions on farmers‟‟ daily needs. The system will benefit farmers through indirect means. This is through development agents can be trained and consulted from identifying suitable land for cereal crop in the case based system which can be implemented at Woreda level without limitation of distance and scarcity of human experts.

1.6. Research Methodology

In this study design science research method was employed. Design science research is a set of analytical techniques and perspectives for performing research in Information Systems (Simon 1996). Design science research involves the design of novel or innovative artifacts and the analysis of the use and/or performance of such artifacts to improve and understand the behavior of aspects of Information Systems. It seeks to create innovations that define the ideas, practices, technical capabilities, and products through which the analysis, design, implementation, and use of information systems can be effectively and efficiently accomplished (Brooks, 1987). Generally, for prototype system development the following steps were used in the study. These were data collection and preparation, modeling and cases representation, system design and development, and finally evaluation of the system performance and user acceptance.

1.6.1. Study site, population and Sampling Techniques

The study site is Ethiopia institute of agricultural research center where land suitability analysis cases collected for land cereal crop matching. Purposive sampling techniques was employ to select sample which can help to acquire the required knowledge from the domain experts who have deep knowledge of about cereal crops, which are concerning how to cultivate and identify the suitable land for cereal crops. In which domain experts were selected based on their knowledge on cereal crops or specialization from the selected institution. For this study the researcher selected 161 cases among a total of 298 cases. Physical Land area with suitability level highly, marginally and moderate is considered for incorporation component of these study for further analysis. Which are non-suitable land for cereal crops and the repeated cases omitted from suitability evaluation.

Also the study area was selected purposively depending on where more land suitability analysis information found collectively. For this reason, ten research experts from Ethiopia institute of Agricultural Research Center cased selected purposively for interview. Thus, a total of ten experts who are crop specialists and land evaluator was participated in the interview for this study.

1.6.2. Method of Data Collection

For the purpose of this study, both primary and secondary data was needed to acquire the required domain knowledge. Primary sources domain expert on crop specialized and land evaluator at Ethiopia institute of agricultural research center Addis Ababa. Secondary sources were used as sources of knowledge in this research are: FAO guideline, different article, Internet resources, databases, books, and manuals. To acquire relevant knowledge for proposed CBS, Interview and document analysis techniques was employ. The research was needed data on climate, topography and soil characteristics of the specific area. In addition to these, data are collected to assess indicators of land utilization in various altitudinal ranges.

Questionnaire was used for the user acceptance of the system being developed. The data was collected from professionals. For testing and evaluation of the proposed system the cases are collected from Ethiopia institute of agricultural research center. Visual interaction method was used to collect data for user's acceptance of the system was developed.

Collection of information to create the case base was done by conducting a study of the domain and by having meetings with experts on land suitability analysis. So the first step in creating a good case structure is to acquire those features that are important to describe situations in the problem domain. The cases were generated within GIS environment. The proposed system was tried to retrieve tabular data and link the inference to spatial and exterior data. Hence, after retrieving the similar cases from the case library, to enable an extended connection to further data (about the similar cases) is considered to provide knowledge acquisition. Therefore, Knowledge acquisition is an essential part of creating the case structure.

In this research, to develop CBR for land cereal crop matching, the relevant knowledge was acquired through a process called knowledge acquisition. This is the process whereby knowledge of group of experts or other sources such as GIS, manuals are gathered, verified, validated and put into CBS.

1.6.3. Case Modeling

After the knowledge is acquired from different sources, the next step was organizing and structuring of knowledge. Knowledge modeling is the representation of information in the form of logic for the purpose of processing knowledge to simulate intelligence (Makfi, 2011).

In this study the acquired knowledge was modeled using hierarchical tree that links agricultural land use of cereal crops to make appropriate decisions making. Knowledge modeling contains input, knowledge model and output.

The main inputs of the systems are spatial and non-spatial data which include water supply (either rain fed or irrigated), soil type (Nutrient Availability, Nutrient Retention, rooting Conditions, Soil Workability and Oxygen Soil Drainage class), topology (degrees of slop), erosion, the socio-economic factors such as farmer's skills, profitability and crop rotation also can be considered factors which are important to produce land suitability for the farm unit. Finally, the output will be recommendations for the farmer that enables to choose suitable cereal crops.

FAO Guidelines for Land Evaluation (FAO 1976) are the basis of present research. It will be used further for the analyzing agricultural structure in order to lay the foundation for collecting, evaluating and analyzing information. Further modification into local context on the basis of parameter prevails in study area will made according to FAO (2007) instruction. Identification of suitable land class on the basis of parameter is presented as follows;

a) Land suitability orders that reflects kinds of suitability: S (suitable) and N (Non suitable).
b) Land suitability classes that reflects the degrees of suitability within orders: S1 (high suitable), S2 (medium suitable), S3 (low suitable), N (Non suitable).
c) Land suitability subclasses that reflect kinds of limitations required within classes, for example: n = (soil nutrient), t° = (temperature), e = (erosion hazard), etc.

1.6.4. Case Representation

Cases are the fundamental units of CBR. They are the essence of CBR and their structure in effect determines how CBR operates. A case is a contextualized piece of knowledge, representing an experience that teaches a lesson fundamental to achieving the goals of a reasoned (Kolodner, 1993).

Case representation is the main part of CBR system, so case library was correctly representing the experience and knowledge (about the domain). For this research feature-value case based representation method was used. The reason for representing cases using feature value case based representation is that this approach uses old experiences to understand and solve new problem. Main steps for building a case library are defining what a case is (contents/general information), deciding which features should be used to describe the case and input the feature values for each case. Therefore, to acquire a data is not enough for developing a case library; this data should be identified, classified, indexed and reformatted if required. In this study Database and case base are used for different purposes. Database systems are designed to do exact matching between queries with stored information, while the goal of CBR is to retrieve most similar cases and knowledge inference. In addition, by the support of relational database management systems (RDBMS) that stores information in tables (rows and columns of data) and conducts searches by using data in specified columns of one table to find additional data in another table. GIS software facilitates to establish linkages between inner and outer components, so RDBMS could benefit the CBR process. Afterwards, we can store geographical data within GIS environment so as to build effective case library.

1.6.5. Implementation Tools

jCOLIBRI 1.1 software (CBR tool) and ArcGIS 10.1 (GIS tool) software was used for developing this study. In such an integrated model, CBS could provide a method for agricultural land use planning in proposing a solution to a new problem or providing relevant experiences, tips or lessons to the decision maker. ArcGIS software was used for preparing and handling spatial and non-spatial data jCOLIBRI software was used to utilize the GIS database as a case library, indexing of these cases and retrieval processes. Firstly, spatial data and non-spatial attributes are generated by using ArcGIS as a tabular input for the CBR software jCOLIBRI. Then the researcher was calibrating this database to develop a case library structure for jCOLIBRI.

According to the selection principles given above, jCOLIBRI was chosen to use as a CBR software tool. This software tries to help application designers to develop and quickly prototype CBR systems interactively. jCOLIBRI is a Lesser General Public License (LGPL) software which means; it is free, available for researchers and some source codes are open, however you can modify or redistribute it under the terms of the LGPL rules. jCOLIBRI is an object-oriented framework in Java programming language that aims to formalize CBR and to provide a design and implementation assistance with software engineering tools. jCOLIBRI formalizes the CBR knowledge using a domain-independent +CBR ontology (CBR onto) which is mapped into the classes of the framework, a knowledge level description of the CBR tasks and a library of reusable Problem Solving Methods (PSMs) (Recio-Garcia et al., 2005).

1.6.6. Evaluation procedure

After a proposed system is developed, evaluation procedures were conducted to check the performance of the prototype system and acceptability by the users. The evaluation process was more concerned with system user acceptance and system performance. User acceptance efforts was concerned with figuring out how well the system addresses the user needs, whereas validation efforts determine if the system performs the intended task effectively.

To assess issues in user and system interaction, questionnaire method was used to gather feedback. The questionnaire was including both closed ended and open ended Questions that will be constructed using clear and objective questions. The results of the system were also compared with the manual test results of agricultural experts. The system testing was evaluated through precision, and recall consideration.

1.7. Organization of the Thesis

This thesis is organized into six chapters: the first Chapter briefly discusses about the introduction part of the study, background of the problem area, the general and specific objectives of the study, the research methodology, the scope and limitation, programming tools used to develop the prototype, evaluation procedure and significance and application of the results of the research. Chapter two gives an overview of related research work and background introduction about case based reasoning system. It introduces the available related and relevant documents to give an overview about case based reasoning, CBR techniques, case based reasoning life cycle, comparison of CBR with other techniques, advantages of using CBR, overview of land evaluation, application of CBR in land development control and related works within the domain area. Chapter three focuses on the case representation methods and creation of case modeling to develop land cereal crops system. Chapter four focuses on designing and implementation of the prototype which discuss the detail of developed system using selected programming tool. Testing and evaluation of the system by domain area is discussed in the fifth chapter and finally chapter six gives final conclusions of the research done and forward recommendations for future studies.

1.8. Operational Definitions of terms

ArcGIS: ArcGIS is an integrated collection of GIS software products that provides a standards-based platform for spatial analysis, data management, and mapping. ArcGIS can be integrated with other enterprise systems such as work order management, business intelligence, and executive dashboards.

Artificial Intelligence (AI): AI is the area of science which focuses on creating systems that can engage on behaviors that humans consider intelligent. Additionally, intelligence is the computational part of the ability to achieve goals in the world.

Attribute: Information stored about a graphic element. Usually, a specific and non-spatial characteristic of an object or entity.

Case Based Reasoning (CBR): To use previous similar cases to solve, evaluate or interpret a current new problem.

Case Based System (CBS): A system (model) approach that based on case based reasoning.

Case Library: It is the core of the CBR system. The cases should contain case specific (also domain specific) details as much as possible.

Decision Support System (DSS): Can be described as an interactive, computer-based system designed to help decision makers solve poorly structured problems. Using a combination of models, analytical techniques and information retrieval, such systems help develop and evaluate appropriate alternatives. Decision support systems should focus on strategic decisions, not operational ones. More specifically, they should contribute to reducing the uncertainty faced by managers when they need to make decisions regarding future options.

Features: Features and components are used interchangeably in GIS. Both terms refer to elements of a land use scenario, such as buildings, roads, etc.

Geodatabase: The geo-database is the common data storage and management framework for ArcGIS and can be utilized wherever it is needed on desktops, in servers (including the Web), or in mobile devices. It supports all the different types of data that can be used by ArcGIS such as attribute tables, geographic features, satellite and aerial imagery, surface modeling, data survey measurements.

Geographic (Data) Model: A data model in geographic information systems is a mathematical construct for representing geographic objects or surfaces as data. For example, the vector data model represents geography as collections of points, lines, and polygons; the raster data model represent geography as cell matrixes that store numeric values; and the Triangulated irregular network (TIN) data model represents geography as sets of contiguous, non-overlapping triangles.

Geographic Feature: An object, real or imaginary, with geographic position classed by point, line or area; for example, a valve, road centerline, or an in-place utility pole.

Geo-processing: Geo-processing is based on a framework of data transformation. A typical geo-processing tool performs an operation on a dataset (such as a feature class, raster, or table) and produces a new dataset as the result of the tool. Geo-processing also supports the automation of workflows by providing a rich set of tools and a mechanism to combine a series of tools in a sequence of operations using models and scripts.

Geo-referencing: To geo-reference something means to define its existence in physical space. That is, establishing its location in terms of map projections or coordinate systems. The term is used both when establishing the relation between raster or vector images and coordinates but also when determining the spatial location of other geographical features

GIS: Geographic Information System(s). Describes any automated system for spatially managing and analyzing geographic information.

Global Positioning System (GPS): A technology that enables an individual to identify the location of an object by triangulating his/her coordinates from a network of satellites.

Graphical User Interface (GUI): A human-machine interaction that relies on graphic symbols and a pointing device to control a computer rather than entry of text from a keyboard.

jCOLIBRI: " 'Java Cases and Ontology Libraries Integration for Building Reasoning Infrastructures" is an object-oriented framework in Java for building Case-Based Reasoning (CBR) systems. It includes mechanisms to Retrieve, Reuse, Revise and Retain cases and is designed to be easily extended with new components.

KBS: (Knowledge Based Systems) A system that uses stored knowledge to solve problems in a specific domain. KBS is a program for extending and/or querying a knowledge base. The Computer User High-Tech Dictionary defines a knowledge-based system as a computer system that is programmed to imitate human problem-solving by means of artificial intelligence and reference to a database of knowledge on a particular subject. KBS are systems based on the methods and techniques of AI. Their core components are the knowledge base and the inference mechanisms.

Layer: A logical collection of geographic entities among which a compulsory physical relationship exists.

Model Builder: Model Builder is an application in which you create, edit, and manage models within ArcGIS software. It is also useful for automated geo-processing.

Model: Model is a pattern, plan, representation (especially in miniature), or description designed to show the main object or workings of an object, system, or concept. Basically a model is a simplified abstract view of the complex reality. Model may also refer to abstractions of concepts and theories.

Point: A graphic element identifiable by a single co-ordinate pair.

Polygon: A multisided figure representing an area on a map. A polygon is a spatial feature defined by the series of arcs comprising its boundary and a label-point establishing its centroid. A closed plane figure usually with more than four sides but in GIS, any closed plane figure, such as parcels, boundaries, etc.

Raster: Data displayed as discrete picture elements (pixels). A cellular data structure composed of rows and columns. Groups of cells represent features. The value of each cell represents the value of the feature. Image data is stored using this structure.

Remote Sensing (RS): RS is defined as the acquisition of information about an object without being in physical contact with it RS is any of the technical disciplines for observing and measuring the earth from a distance, including satellite imaging, global positioning systems, radar, sonar, aerial photography, etc.

Spatial Data: Information about the location, shape and relationships among geographic features.

Spatial Decision Support System (SDSS): Spatial Decision Support System is a framework for integrating database management systems with analytical models, graphical display and tabular reporting capabilities and the expert knowledge of the decision-makers. Such systems can be viewed as spatial analogues of decision support systems developed in operational research and management science to address business problems.

Topography: Shape or configuration of the land surface; represented in map form by contour lines.

Topology: Descriptions of geographical relationships of features; especially which features are adjacent to or connected to other features. The explicit representation of the position of a feature, relative to features it defines, or is defined by.

XML: Extensible Markup Language (XML) is a method for putting structured data (such as that in a worksheet) in a text file that follows standard guidelines and can be read by a variety of applications. Designers can create their own customized tags, enabling the definition, transmission, validation, and interpretation of data between applications and between organizations.

CHAPTER TWO

LITERATURE REVIEW

Conceptualizing the basic ideas of knowledge based system is essential to understand sufficiently the notion of KBS. So this chapter attempts to discuss review of literatures on land

Evaluation concept and common cereal crops production in Ethiopia. It deeply elaborates case based reasoning concepts, architecture, and knowledge acquisition and representation methods. It also discusses knowledge based systems techniques and the application particularly in Agriculture.

2.1. Land Evaluation Concepts

The FAO (2007) defined land evaluation as the process of measurement of land performance when used for particular purpose. In this way land evaluation can be useful for predicting the potential use of land based on its attributes (Rossiter, 1996). Land evaluation has developed from soil survey understanding and land classification. Soil survey interpretation is predictions of performance, not recommendations for the use of soils (Beek, 1980). Agricultural land use requires not only that good soil, but also there are other factors limit the productivity of the land such as erosion, climate, hazard and topography. currently, these factors are included in the most of land evaluation systems.

The basic feature of land evaluation is the assessment of the prerequisite of land use types with the character of the available land resources, and involves the explanation of surveys and studies of soils, vegetation, climate and landforms. Fundamental to the evaluation process so is the fact that different kinds of land uses have different requirements (Dent and Young, 1993). Land evaluation presents information and recommendations which can assist planners and decision makers to decide which crops to cultivate where and the limitation of land use.

Land evaluation is the selection of suitable land and suitable cropping. The main production of land evaluation exploration is a land classification that indicates the suitability of different types of land for specific land uses, mostly described on maps with accompanying reports (FAO, 1981).

According to FAO (2007), land evaluation should provide answers to such questions as:

- What other uses of land are physically possible, economically and socially relevant?
- What inputs are essential to achieve a required level of production and minimize the unfavorable effects?
- What are the current land uses and what are the consequences if current management practices stay the same?

2.2. Land Evaluation Approach

The FAO framework for land evaluation is just a set of guidelines and they are not strict instruction manuals. However, evaluators have to select land characteristics and qualities, which fit their requirements, which are different from one environment to another. Therefore, computer systems used in different environments and different sets of data may not be used for other sets of data and conditions (FAO, 2007). As a result, a number of computer systems have been used to develop land evaluation methods in several regions of the world for example ALES, LECS, ISLE, LEIGIS and LUSET.

The basic requirements of applying the FAO framework are the selection and definition of land utilization types for which the land is to be evaluated. The requirements of the land utilization types are compared with the land resources. In this process, land resources are described as land qualities and land characteristics. In 1976, the FAO provided a general framework for land suitability classification. The framework in itself, does not propose a specific method for doing this classification (Keshavarzi, Sarmadian, et al., 2010); rather it is a set of methodological guidelines for the determination of land suitability. It was basically designed to address any kind of environment and at any scale, and to be utilized especially in regions with limited data (FAO, 1976). The FAO framework has three different guidelines. These guidelines are: 1) land evaluation for rain fed agriculture (FAO, 2007), 2) land evaluation for irrigated agriculture (FAO, 1984), and 3) land evaluation for natural forests (FAO, 1984). These guidelines are designed to assess crop, management, environmental and conservation requirements. The guidelines for rain fed agriculture may be considered the norm for land evaluation. The main different between the guidelines for land evaluation for rain fed and the guidelines for irrigated agriculture is that the latter takes into the account quantity and quality of water resources and economic factors. Special features of guidelines for land evaluation for natural forests are therefore that the land-use types may be related to conservation rather than production, that the land use is commonly multiple uses (including wood production, conservation, recreation, grazing etc.). A checklist of land qualities for assessing land suitability that suggested from the guidelines for land evaluation for rain fed agriculture. In later years, the set of methods in land evaluation were emerged based on the FAO framework (FAO, 2007). The important definitions that are used in the framework (FAO, 1976) are presented in the glossary The FAO framework describes a methodology for land suitability classification and the term *suitability* is used rather than *capability*. The FAO identified land suitability as "the fitness of a given tract of land for a defined use" (FAO, 1976). According to the FAO, the term "land suitability evaluation" could be interpreted as the process of assessment of land performance when the land is used for specified purposes.

The FAO assessed and classified land suitability with respect to particular uses since what is suitable for one kind of cultivation may not be suitable for another. The process of land suitability classification is assessment and grouping of specific areas of land in terms of their suitability for defined uses (FAO, 1976). For instance, an alluvial flood plain with impeded drainage might be highly suitable for rice cultivation but not suitable for many kinds of agriculture or for forestry (FAO, 1981). The concept of land suitability is only meaningful in terms of specific kinds of land use, each with their own requirements, e.g. for soil moisture, rooting depth etc. The qualities of each type of land, such as moisture availability or liability to flooding, are compared with the requirements of each use.

The framework classifies the suitability of land into four categories: land suitability orders, classes, subclasses and units. Suitability orders indicate whether land is assessed as suitable (S) or not (N) for the use under consideration. Classes indicate the degree of suitability (up to five), for example, highly suitable (S1), moderately suitable (S2), marginally suitable (S3), currently not suitable (N1) and permanently not suitable (N2).

2.3. Significance of Cereal Crops in Ethiopia

In terms of caloric intake, cereals dominate the diets of Ethiopian households. Rashid (2010**)** suggested that an average Ethiopian consumes 1858 kilocalories. Out of the total calorie consumption, five major cereals (maize, tef, wheat, barley and sorghum) account for more than 60 percent, with maize and wheat representing 20 percent each. The low share of tef in calorie consumption often come as surprise to urban Ethiopians as tef is the predominant staple food in the middle and high income households. Wheat is a staple crop in the highlands of Ethiopia. Most of the wheat grain produced by small-scale farmers was consumed or retained as seed on the farm and little surplus (19.4%) went to the market.

Barley (Hordeum vulgare) is one of the most important cereal crops, mainly grown by smallholder farmers at mid- and high altitudes (Minale et al., 2011). Tef (Eragrostis tef (Zucc.) is locally named as tef (Amharic), taff (Oromigna) and taf (Tigrigna) (Hailu,2001). It is one the major indigenous staple food crops with finest grain originated and grown in Ethiopia. Ethiopia has 350 different varieties of tef and very few of the generally known locally mentioned by the farmers as “Neck tef" the white colour grain, “Key tef" the red colour grain and “Sergegna which is the mixed of the two. This crop is used for a range of purposes including pan cake (injera), porridge (genfo) and traditional beverages (Taddese, 2005). Maize (Zea mays L.) is locally named as Bekollo or Yebahir mashilla (Amharic), Bokollo (Oromigna) and meshella_bah’ri or e'ffun (Tigrigna) (Hailu, 2001). It is one of the most important cereal crops used for both food and local alcohol preparation. In most developing countries (Latin America, Africa and Asia), maize is grown as a staple food crop (grain maize); while in other countries like USA and Brazil, it is an important animal feed or is a basic compound for bio-fuel (ethanol) production (Verheye, 2010).

2.4. GIS Applications in Land Evaluation

GIS-based techniques for land use suitability analysis developed from the practice of manually overlaid maps which were developed in the USA in the last century. GIS capabilities for spatial analysis overcome the drawbacks of the paper map overlay approach (Malczewski, 2004). However, GIS has now become a dominant tool for land use planning due to its ability to deal with different functions which is very useful for land use planning. These functions, the most important are database management, cartographic analysis and modeling function. The ability to integrate data in GIS is one of the most important advantages of the system, involving collection of data from different sources, formats, and scales and making them compatible with each other (Flowerdew, 1991). The main feature of integrated data management is the ability to present the information of different layers at the same time, which can help planners and decision makers by showing together distinct factors that affect land use (FAO, 2007). Moreover, GIS has the ability to integrate variety of geographic technologies such as Global Position System (GPS) and Remote Sensing.

The modeling function provided by GIS can benefit land evaluation by providing the ability to analyze and model data layers by automatic approach. Once a model has been built and validated, the repetition of the analysis, as assumptions and /or conditions change, is a quick and easy task. This function also provides an interface between GIS and other modelling software which can integrate non spatial data. An example, suitability maps can be integrated with non-spatial data, such as socio-economic data to model the effect of these data on the land use. This function of GIS can save time and cost in the evaluation of land use options compared with conventional methods (Burrough, McDonnell, et al., 1998).

2.5. Concepts of Artificial Intelligence

The term Artificial Intelligence (AI) refers to the activity of building intelligent systems. It is a technology of making computers to simulate human being’s intelligence (Raza, 2009). An intelligent system is a system that exhibits and possesses some basic attributes such as performing some actions, reasoning about a particular domain, making decision and goal oriented problem-solving capability. A system or an agent can be said to be intelligent when the agent's performance cannot be distinguished from that of a human performing the same task (Honavar, 2006).

The prime goal of AI research is to raise human beings understanding in all aspects like in human being’s perceptual, reasoning, learning, and creative processes (Honavar, 2006). The first major and successful AI research application technologies are expert systems or knowledge based systems (Pomykalski et al., 1999). Early efforts in building AI programs were intended to create general-purpose problem solvers. Later on with the advancement of the technology domain specific Knowledge Based System (KBS) had been developed for the various domains, including agriculture.

2.6. Knowledge based reasoning techniques

Rule-based and case-based reasoning are the two popular approaches used for problem solving in intelligent systems (Prentzas and Hatzilygeroudis, 2002). Rule based reasoning is a system whose knowledge representation in a set of rules and facts. Symbolic rules as one of the most popular knowledge representation and reasoning methods, their popularity is mainly due their naturalness, which facilitates comprehension of the represented knowledge. The basic forms of a rule, if<condition> then <conclusion> where <condition> represents premises and <conclusion>represents associated action for the premises. The condition of rules is connected between each other with logical connectives such as, AND, OR and NOT., thus forming a logical function. When sufficient conditions of a rule are satisfied, then the conclusion is derived and the rule is said to be fired. Rules based reasoning was dominantly applied to represent general knowledge. Rule based expert systems have a significant role in many different domain areas such as medical diagnosis, electronic troubleshooting and data interpretations. A typical rule based system consists of a list of rules, a cluster of facts and an interpreter (Prentzas, 2007).

Case-Based Reasoning (CBR) is a general artificial intelligence paradigm for reasoning from experience. CBR methodology has been investigated in improving human decision-making and has received much attention in developing knowledge based systems in medicine (Quaglini, et al., 2001).

2.7. Case-Based Representation

Case-Based Representation is a computer technique which combines the knowledge-based support philosophy with a imitation of human reasoning when past practice is used that is mentally searching for alike conditions happened in the past and reusing the knowledge gain in those circumstances (Leake, 1996). The idea of case based reasoning is founded on the idea of using explicit recognized knowledge's to solve new problems. The decision maker uses previous, explicit knowledge called cases, to help him solve a present problem. He retrieves the suitable cases from a larger set of cases. The similarity between a present situation and the retrieved case are the basis for the latter's selection (Gonzalez &Daniel, 1993).

2.7.1. **Cases**

A case can be said to be the record of a previous experience or problem. The information obtained about this past knowledge will by necessity depend on the domain of the reasoned and the purpose to which the case will be place In the case of a problem solving CBR system, the details will usually include the requirement of the problem and the relevant attributes of the situation that are the conditions of the problem. The other fundamental part of the case is the solution that was applied in the previous situation. Depending on how the CBR system reason with cases these solution may contain only the facts of the solution or processes concerned in gaining the solution. It is also important to include the achieved measure of success in the case description if the cases in the case base have achieve different degrees of achievement or breakdown (Leake, 1996).

The information contained in a case base leftovers exact to the case in which it is stored (Kolodner, 1993). Because of the specific knowledge of a case base we find that related information and knowledge applicable in a specific situation is stored in close proximity. Thus rather than drawing knowledge from a wide net the knowledge needed to solve a specific problem case can be found grouped together in a few, or even one location.

2.7.2. **Case based reasoning system**

At the highest level, a case based reasoning system incorporates the reasoning mechanism and the external facets: the input specification the output suggested solution the memory of past cases that are referenced by the reasoning mechanism (Prentzas, 2007).

According to Sankar, et al. (2001) the most CBR systems, the case-based reasoning mechanism (problem solver) has an internal structure divided into two major parts; the case retriever and the case reasoned as shown in Figure 2.8.2 The case retriever's task is to find the appropriate cases in the case base while the case reasoned uses the retrieved cases to discover a solution to the specified problem explanation. This reasoning generally involves both formative the differences among the retrieved cases and the current query case; and modifying the retrieved solution appropriately, reflecting these differences.

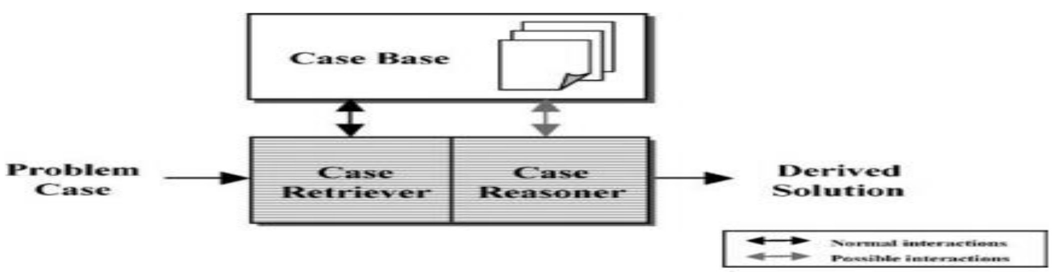

Figure 2.7.2.1: Two major components of CBR system (Source: Sankar, et al., 2001)

2.7.3. **Case based reasoning life cycle**

Solving a problem by CBR involves getting a problem description, measuring the similarity of the current problem to previous problems stored in a case base with their known solutions retrieving one or more similar cases and attempting to reuse the solution of one of the retrieved cases, possibly after adapting it to account for differences in problem descriptions. The solution proposed by the system is then evaluated. Following revision of the proposed solution if required in light of its evaluation, the problem description and its solution can then be retained as a new case, and the system has learned to solve a new problem (Prentzas, 2007).

At the heart of the system, case- based reasoning has been formalized for purpose of computer reasoning as a four-step procedure (Aamodt and Plaza, 1994): retrieve, reuse, revise and retain. As depicted in figure 2.8.3, the first step is retrieve. Given a target problem; retrieve cases from memory that are relevant to solving it. A case consists of a problem, its solution, and, typically, annotations about how the solution was derived. Retrieve is followed by reuse where the solution from the previous case is mapped to the target problem. This may engage adapting the solution as needed to fit the new circumstances. Having mapped the previous solution to the target situation, test the new solution in the real world (or a simulation) and, if necessary, revise. After the solution has been successfully revised by domain experts, retain/store the resulting experience as a new case in memory for future use.

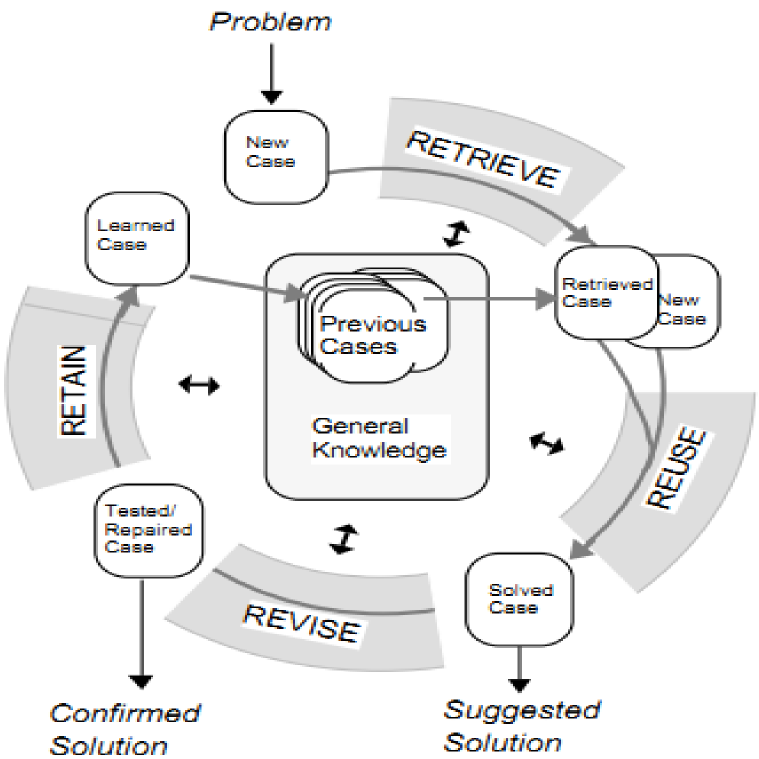

Figure 2.7.3.1: The CBR cycle (Source: Aamodt and Plaza, 1994)

2.7.4. **Advantages of case based reasoning**

A case based reasoning approach has tremendous advantages in the development of knowledge based system. According to Pal and Shiu (2004), the following are merits of using CBR approach.

- **Reduce the knowledge acquisition task:** the knowledge acquisition tasks of CBR contain most of the collection of the relevant prevailing experiences/cases and their representation.
- **Learn over time:** CBR systems are used, they meeting problem circumstances and make extra solutions. If solution cases are tested in the real world, and a level of success is determined for those solutions, then these cases can be added in the case base and used to help solving future problems. As cases are added to CBR system

would be able to reason in a wider variety of situations, and with a higher degree of refinement and success.

- **Keep away from repeating entire the steps that need to be taken to arrive at a solution**: In problem areas that need important processes to create a solution from scrape the alternate approach of adapting a preceding solution can importantly decrease this processing prerequisite. Moreover, reusing a previous solution also permits the real phases taken to get that solution to be reused for solving other problems. **Can be used in many different ways and broad range of domains:** The number of ways a CBR system can be implemented is almost unlimited. It can be used for many purposes. A few examples of CBR systems implementation areas are: creating a plan, making a diagnosis and arguing a point of view. So, the data dealt by a CBR system is likewise able to take many forms and the retrieval and adaptation appoaches will also vary. Whenever stored past cases are being retrieved and adapted, case-based reasoning is said to be taking place. CBR can be applied to extremely diverse application domains due to the seemingly limitless number of ways of representing indexing, retrieving and adapting cases.
- **Reflect human reasoning**: as there are many situations where humans use a form of CBR, it is not difficult to convince implementers, users and managers of the validity of the paradigm. Likewise, humans can understand a CBR system's reasoning and explanations, and are able to be convinced of the validity of the solutions they receive from a system.

Though CBR is helpful for different types of problems and domains there are situations when it is not the most fitting methodology to employ. According to Prentzas (2007), some of the limitation issues in CBR are inability to express general knowledge, knowledge acquisition problems, inference efficiency problems and Provision of explanations.

2.8. CBR Techniques

There are different techniques in case based reasoning. Among them the common knowledge engineering tasks and issues which are crucial in developing the CBR systems are: case representation, case indexing, case storage, case retrieval, case adaptation, learning and generalization (Salem *et al*., 2007).

2.8.1. **Case Representation**

Case is a specific piece of knowledge representing an experience (Aamodt and Plaza, 1994). It contains the information which is content of case and situation where that information or knowledge can be used. Different type of data can be kept in a case. CBR community has a lack of consensus what information should be stored in a case. A Case is a combination of two components. These components are description of a problem and its solution. Problem description consists of a set of attributes and values and based on the description attributes values solutions are predicted. Case representation selected also be contingent on a number of factors like the representation used in the case base, number of features that are used to match cases during searches, the purpose to which the system is being put, and he number and complication of cases have been kept.

2.8.2. **Case Indexing**

Case indexing refers to conveying directories to cases for future retrieval and contrasts. This choice of indices is important to being able to retrieve the right case at the right time. This is because the indices of a case determine in which context it would be retrieved in future. Indexes should reflect the important aspects of the case, the attributes that influenced the outcome of the case and also those which describe the circumstances in which it is expected that they should be retrieved in the future (Salem *et al*., 2007).

Index is a computational data structure can be held in memory and also can search quickly (Aamodt and Plaza, 1994). In medical systems, where patient's age, sex, height and weight can be used as index features, that information is helpful for future retrieval. The patient's photograph can be included as an unneeded feature which cannot be used in the retrieval. Picture should be helpful for doctor for remind patient. Index should have the following features (Aamodt and Plaza, 1994; Pal and Shiu, 2004): be predictive, should show the purpose for which case will be used, it should be easy to recognize it in the future and it must address the future use of the case base. Indices can be selected in both ways: manually and automatically. If you choose index manually then you have to decide a purpose of a case.

Although passing on indexes is still mainly a manual process relies on human experts, various tries of using automated approaches were proposed in the literature. According to Bonzano (1997), inductive techniques are used for learning local weights of features by comparing similar cases in a case base. Other methods include indexing cases by features and dimensions that are predictive across the entire problem domain (Acorn and Walden, 1992); by computing the differences between cases; adaptation guided indexing and retrieval (Smyth and Keane, 1998) and explanation-based techniques.

2.8.3. **Storage**

One important aspect of the efficient CBR system is case storage (Aamodt and Plaza, 1994). It represents a logical view of what is stored in case. For efficient retrieval, case base should be organized in a manageable way. These methods referred as a case memory model.

Retrieval

Case retrieval is a process of finding cases which are closest to current case (Pal and Shiu, 2004). Case retrieval is the process of finding within the case base those cases that are the closest to the current case. For efficient case retrieval, there should be selection criteria which could judge a case. To carry out case retrieval there must be criteria that determine how a case is judged to be appropriate for retrieval. The selection criteria are necessary to decide which case is the best one to retrieve, that is, to determine how close the current and stored cases on the basis of similarity values.

The actual processes involved in retrieving a case from the case base depend very much on the memory model and indexing procedures used. There are a number of factors to consider when determining the method of retrieval. Among them: the number of cases to be searched, the amount of domain knowledge available, the ease of determining weightings for individual features, whether cases should be indexed by the same features or whether each case may have varying important features are a few to be considered (Salem *et al*., 2007).

There are different retrieval methods that range from a simple nearest neighbor search to the use of intelligent agents. According to (Pal *et al*., 2001) the most common, traditional retrieval methods are nearest neighbor retrieval, inductive, knowledge guided approaches and validated retrieval.

Nearest Neighbor Retrieval: In nearest neighbor retrieval, the case retrieved is chosen when the weighted sum of its features that match that query is greater than the other cases in the case base. Some features that are considered more important in a problem solving situation may have their importance denoted by weighting these features more heavily in the matching (Watson and Marir, 1994).

Nearest-neighbor retrieval (NNR) is a technique to measure how similar the target case is to a source case (Watson, 1999). It processes retrieval of cases by comparison of a collection of weighted attributes in the target case to source cases in the CBR library. If there is no matched case in the CBR library, CBR system will return the nearest matched source case. The return of the nearest case match can be represented by the following equation (Watson and Marir, 1994):

$$\text{NN}(\text{I},\text{R}) = \frac{\sum_{i=0}^{n} wi \text{ x } \text{sim}(f_i^I, f_i^R)}{\sum_{i=0}^{n} wi}$$

Where: W is the importance weighing of an attribute;

I is the target case;

R is the source case;

i is individual attributes from 1 to n;

sim is the local similarity function

f_i^I and f_i^R are the values for attributes i in the input case (I) and case in the case base (R) respectively; and

n is the number of attributes in each case;

The weights allocated to each feature/attribute provide them a range of importance. But determining the weight for a feature value is a problem and the easy way is to calibrate this weight by an expert or user in terms of the domain knowledge.

This calculation is frequent for every case in the case library to rank cases by similarity to the target. Similarity measures are functions used to compare two case entries. A case entry may have symbol or data values which attribute of the entry that should be compared is given as a parameter together with the two case entries. The similarity function will return a value between 0 and 1, ranging from no similarity to completely equal (exact match).

Algorithms similar to this are used by the majority CBR tools to carry out nearest neighbor retrieval. Similarities are usually normalized to fall within a range of zero to one (where zero is totally dissimilar and one is an exact match) or as a percentage similarity where 100% is an exact match. In other words, the weight ranges from 0 to 1, where 0 means irrelevant, and 1is very important (essential). The scale is linear, meaning an entry with weight 0.2 is twice as important as an entry with 0.1.

Inductive approaches: When inductive approaches are used to determine the case base structure, that is to determine the relative importance of features for discriminating between similar cases, the resulting hierarchical structure of the case base provide a reduced search space for the case retriever. This may in turn reduce the search time for queries (Watson and Marir, 1994).

Inductive retrieval: is a technique to extract rules or construct decision trees from the past cases. This technique processes a target case based on indexed source cases. The source cases are normally indexed by keywords and stored into a set of cases. The set of cases are divided into a decision tree structure. When inductive approaches are used to determine the case base structure the resulting hierarchical structure of the case base provides a reduced search space for the case retrieval. This may in turn reduce the search time for queries in retrieval. If target case is not found in the decision tree at runtime, the CBR system may not retrieve a source case (Watson, 1997).

Knowledge Guided Approaches: define retrieval use domain knowledge to determine the features of a case which are important for that case in particular to be retrieved in future. In some situations, different features of each case will have been important for the success level of that case (Watson and Marir, 1994).

Validated Retrieval: Validated Retrieval consists of two phases, firstly the retrieval of all cases that appear to be relevant to a problem, based on the main features of the query case. And the second phase involves deriving more discriminating features from the group of retrieved cases to determine whether the cases are valid in the current situation. The advantage of this method is that inexpensive methods can be used to make the initial retrieval from the case base, while more expensive methods can be used in the second phase as they are applied to only a subset of the case base (Watson, 1997).

2.8.4. **Case Adaptation**

Case adaptation is the process of translating the retrieved solution into the solution appropriate for the current problem. It is a technique to alter retrieved case for reproducing new solution for new problem (Lang and Lau, 2002). It may be the most important step which adds intelligence. Case adaptation improves overall problem solving ability of CBR.

2.8.5. **Learning and maintenance in CBR systems**

Once an appropriate solution has been generated and outputted, there is some expectation that the solution will be tested in reality. Using real world assessment, the CBR system can be updated to take into account any new information uncovered in the processing of the new solution. Information can be added to the system for two purposes: first, the more information that is stored in the case base, the closer the match found in the case base is likely to be and second, the purpose of adding information to the case base is for is to improve the solution the CBR is able to create (Salem *et al*., 2007).

Learning

A very important feature of case-based reasoning is its coupling to learning. The driving force behind case based methods has to a large extent come from the machine learning community; and case-based reasoning is also regarded a subfield of machine learning (Aamodt, 1993). Thus, the notion of case-based reasoning does not only denote a particular reasoning method, irrespective of how the cases are acquired, it also denotes a machine learning paradigm that enables sustained learning by updating the case base subsequent to a problem has been solved. Learning in CBR occurs as a natural by-product of problem solving. When a problem is successfully solved, the experience is retained in order to solve similar problems in the future. When an attempt to solve a problem fails, the reason for the failure is identified and remembered in order to avoid the same mistake in the future.

CBR favors learning from experience, since it is usually easier to learn by retaining a concrete problem solving experience than to generalize from it. Effective learning in CBR requires a well worked out set of methods in order to extract relevant knowledge from the experience, integrate a case into an existing knowledge arrangement, and index the case for later corresponding with similar cases.

Maintenance

Leake and Wilson (1994) defined case-base maintenance as the implementation of policies for revising the organization or contents (representation, domain content, accounting information, or implementation) of the case-base in order to facilitate future reasoning for a particular set of performance objectives. When applying CBR systems for problem solving, there is always a tradeoff between the number of cases to be stored in the case library and the retrieval efficiency. The larger the case library, the more the problem space covered but, it would also downgrade the system performance if the number of cases grows to an unacceptable high level. Therefore, removing the redundant cases or less useful cases under an acceptable error-level is one of the most important tasks to maintain CBR systems.

2.9. CBR Tools

There are different commercial and non-commercial tools which are used in developing CBR systems. Some of CBR tools which are indicated in the work of Iqbal and Hassan (2006) are described as follows.

CBR-Express and ReCall

CBR-Express is developed by Inference Corporation and is among most successful application of CBR. The interface of CBR Express is build using Asymetrix Tool Book. It is basically a help desk system and has simple case structure and use nearest neighbor matching to retrieve cases. Recall is a CBR based tool made by ISoft which is an AI company. It is coded in C++ and offers a combination of nearest neighbor and inductive case retrieval.

ReMind and Kate

ReMind is produced by Cognitive Systems Inc., and developed initially for Macintosh. But now it is also developed for Windows and UNIX. ReMindoffer template, nearest neighbor, inductive and knowledge based retrieval. Kate is a CBR tool which is developed by AcknoSoft (Watson &Marir, 1994) that can run on MS Windows, Mac, or SUN. Kate is made up of a set of tools: Kate-Induction, Kate-CBR, Kate Editor and Kate-Runtime. This tool supports both kinds of nearest neighbor and induction algorithms. Kate-Induction is an ID3-base induction system that supports object-oriented representation of cases. Cases can be imported from many databases and spread sheets. Kate-CBR uses nearest neighbor approach.

CBR-Works and my CBR

CBR-Works can be seen as a CBR-Shell providing all necessary tools to model, maintain, and consult a case base (Schulz, 1999). CBR-Works comes from the German company TECINNO, running on MS Windows, Mac, OS/2, and various UNIX platforms. CBR-Works can import case-bases from Microsoft Excel and in the CASUEL case format. CBR-Works is developed for e-commerce applications but can be used for other purposes also. It contains elements from all knowledge containers and can perform full CBR cycle. On the other hand; myCBR is open source, developed under the GPL license. It can be viewed as a successor of CBR Works and contains many useful features. Key motivation for implementing myCBR was the need for a compact and easy-to-use tool for building prototypical CBR applications in teaching, research, and small industrial projects with low effort.

jCOLIBRI

Belen Diaz-Agudo developed domain independent architecture called COLIBRI in 2002 (Recio and Diaz-Adugo, 2002). It stands for Cases and Ontology Libraries Integration for Building Reasoning Infrastructures. COLIBRI helps in the design of knowledge intensive CBR systems. As an object oriented framework, jCOLIBRI's aim was to support COLIBRI architecture. It is basically technical evaluation of COLIBRI which support in distributed architecture a Digital Logic (DLs) engine, Graphical User Interface (GUI) clients for producing CBR system from reusable components and object-oriented framework in java.

In the study jCLIBRI was used. It is non-commercial compatible tool which supports the full CBR cycle (Retrieval, Reuse, Revise and Retain). By using jCOLIBRI, it is possible to work with external databases using different connectors; like text connectors in our case. Using jCOLIBRI is also suitable for developing large scale applications.

2.10. Using CBR Technique for Geographic and Spatial Analysis

CBR is driven by two motivations. The first one is the desire to model human behavior (from cognitive science). The second one is the pragmatic desire to develop technology/technique to make AI systems more effective (Leake, 1996). In this respect, the use of land use planning, such as CBR, can reduce the time to process the applications and help in making consistent decisions for similar applications. When dealing with such problems, experts usually draw upon previous experiences or the memory of how similar cases had been solved in the past. This is the type of problem solving cycle which is most suitable for CBR. Also it can overcome the problems with black box inference process of some AI techniques and rule-based reasoning.

According to Holt & Benwell (1996), lack of analytical and modeling functionality is a major deficiency of current spatial information systems. Therefore, there is a perceived need to integrate spatial information systems with additional analytical approaches to overcome this deficiency. CBR is proposing a methodology for building and applying more practical PSS for planning. The results derived directly from real cases in CBR process are more convincing and acceptable to planners. Urban planning often requires the experiences and expertise of planners and assistance of planning models and analytical methods in its complicated decision-making process. However, the capability of people in handling large amount of information and the availability of experienced planners are often limited (Shi & Yeh, 2001). By the usage of CBR technique, the traditional working style of a planner (which are based on his/her knowledge of past records) in dealing with applications could be simulated. The following section tries to highlight CBR issues/methods relate to GIS and city planning process.

This is an example of “search to find a similar spatial pattern and associated non-spatial attributes” from the study of Holt & Benwell (1996). They state that users may need to know about a spatial phenomenon before they make a decision. Knowing what has been done before can aid the decision-maker. All process is simplified for easy comprehension.

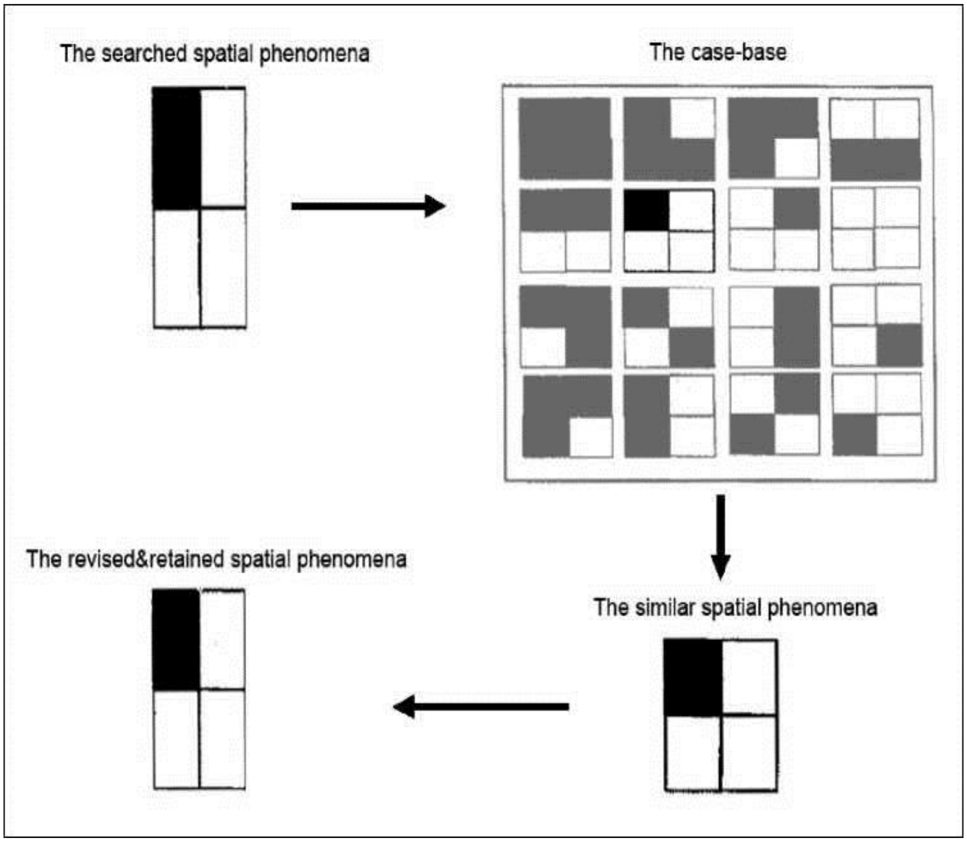

*Figure 2.10.1: **Using CBR for spatial phenomena (Source: Holt & Benwell, 1996)***

2.11. Knowledge based system in agriculture

In agriculture expert systems, unite the accumulated expertise of individual disciplines e.g., plant pathology, entomology, horticulture and agricultural meteorology into a framework that most excellent addresses the spacific on-site needs of farmers (Prasad & Vinaya, 2006).

In the nineties, several expert systems have also been developed in agricultural field. An agro forestry expert system was developed to help land use research scientists, officials, farmers, and other individuals interested in increasing the benefits gained from applying agro forestry management techniques in developing countries (Prasad & Vinaya, 2006).

According to Robinson (1996), Expert systems are used in a wide range of areas in agriculture. Its major usage areas are: Crop Management Advisors: These kinds of systems help the farmers by giving decision helping on the process of growing a certain type of crop. For instance, to generate fertilizer recommendations; This system deals with specific task fertilization in growing crops.

Livestock Management Advisors: Similar to crop growing advisors, this system gives combined advices for animal breeders. Example; Application of conditional causality in an combined KBS for daily farms. The system has health production and financial modules. It gives decision support for daily farm management tasks.

Planning system: Production planning systems with ES in agriculture, deals with identifying and suggesting projections, plans for future cropping activities. Example implementation is: CROPS: A whole farm crop rotation planning system to implement sustainable agriculture. Main purpose of the system is to obtain sustainable and profitable cropping plans which meet given production needs and various constraints.

Pest Management Systems: These systems help farmers to deal with harmful creatures with optimal management solutions. Such as an expert system for integrated pest management of apple orchards. The system helps the farmer first identifying the problem and then gives advice for taking actions.

Diagnostic Systems: Different from the pest management, diagnosis is concerned with any kind of disease in plants and crops. These tools are like the well-known MYCIN. It works the same way. An addictive reasoning expert system shell for plant disorder diagnosis: It is a domain- specific generic tool for diagnosing plant diseases.

Conservation/Engineering Systems: Problems dealing with engineering solutions to conservation problems. A typical example is: Development and validation of an expert system for soil erosion control planning in Prince Edward Island: It is used for conserving soil by recommending the appropriate engineering solution to control soil erosion within typical cropping systems.

Process Control Systems: This system monitors some sensors and takes corrective actions using some instruments. Determination of greenhouse climate set points by SERRISTE: its function is to maintain certain conditions in a glass house for winter production of tomatoes.

Marketing Advisory Systems: Gives advices to farmers on marketing different products.

Cattle and GRAIN Marketing: Helps farmers to select different marketing alternatives for their cattle, and grain.

The problems in agriculture are often multidisciplinary and very complex because of affecting complex events. Expert systems approaches will succeed with this kind of problems. It has many methods for uncertainty and reasoning using whatever on the hand. ES in Agriculture will help farmers and animal breeders make their decisions more efficiently and timely. Currently many people are forced to make decisions about agricultural activities without enough knowledge. Many of them have inadequate training about agriculture and needs to be managed.

2.12. Related works

Shi & Yeh (2001) developed a system that integrates a CBR shell (ESTEEM) and GIS package (ArcView) to build Case Based System (CBS) in Hong Kong, China. They tried to show how CBR can be used to handle planning applications in development control. Authors state that CBR can simulate the present working style of a planner in dealing with development applications which is based on the knowledge of past application records. They used the previous planning application cases to support the suggestions to the decision makers rather than generalizing rules and then performing a logic inference to get the conclusion.

Elaleem (2010) carried out a study whose aim was to determine the physical land suitability areas for barley, wheat and maize crops in the north western region of Libya. The FAO framework for land evaluation with Fuzzy Analytical Hierarchy Process (AHP) and Ideal Point methods were employed to determine land suitability classes for the selected crops. Pairwise comparisons method was applied for determining the weights of criteria for land characteristics. The findings emphasized that soil factors represented the most sensitive criteria affecting all the crops considered. In contrast, erosion and slope were found to be less important in the study area. The study applied manual Fuzzy logic method based on some membership functions developed by some researchers. However, the membership functions that have been successfully developed in a different environment may not be appropriate for other environment.

Van Huynh and Michael (2005) carried out a study whose aim was to determine the physical land suitability areas for grapefruit crop production in Vietnam and sustainable agriculture development of a representative village Thuy Bang, Hue, Vietnam. The methodology used for the physical land suitability analysis for “Thanh Tra” pomelo is a multi-criteria evaluation approach within GIS context, based on FAO land evaluation framework, modified for Vietnamese conditions. The methodology consists in matching land qualities against crop requirements of “Thanh Tra” grapefruit. The important parameters were categorized into six maps namely; soil unit’s map, slope map, texture map, soil effective depth map, organic material map, soil fertility map. Land Evaluation Units (LEUs) map and physical land suitability classification were obtained by overlapping the above mentioned maps within a GIS system. The study concluded that lack of irrigation; erratic rainfall and poor soil fertility are the most serious problems influencing yield and quality of “Thanh Tra” pomelo

Messing et al. (2003) developed Land suitability classification in China based on the FAO Framework (1976). Fifteen Land characteristics were selected to classify Land qualities into six classes namely: available water, slope aspect, erosion hazard, soil workability, available nutrients and flooding hazard. Then GIS was used for the comparison between the current land use and the land suitability for agriculture. The result was four scenarios for planning suitable land use in the study area.

All of the above related works are instances of effort to automate the FAO land evaluation framework taking advantage of the pervasiveness of the computer and the veracity offered by GIS in the land mapping and manipulation of spatial data indeed, since the FAO land evaluation framework and the emergence of GIS as an effective tool in land evaluation and geo-spatial information analysis, a number of computer systems have been developed for land evaluation and spatial analysis based on the framework. In addition, GIS can be an analysis, data generator for spatial data of specific place of lands which can change over time due to different mechanism. So the major constraints seen in the above reviewed work are mainly incapability to consider the dynamic nature of a land since the growing cereal crop potential of a particular land will be changed over time which means it considers only static factors which are neglecting dynamic factors. Also most of land evaluation systems developed by different researcher are location dependent in that the system only works for the study area where the data are collected but may not be functional for other locations. Furthermore, inability to collectively consider physical land suitability with socio-economic factor was also another main limitation identified during review of related literature. Instead, most of the reviewed study gives emphasis only towards one of the above factors either physical land suitability or socio-economic factor which ignores application of both physical and socio-economic factors jointly in land crop matching.

However; this research adopted a knowledge base approach to farm level agricultural planning that combines case-based reasoning systems methodology with GIS tools. The present approach is different from the other approaches to land evaluation as this can represent different types of factors such as qualitative, quantitative and heuristic factors that are associated with land evaluation for choosing an optimal cereal crop for a farm unit. Mainly the developed CBRLCCM system distinctively consider both physical land evaluation with socio-economic factor during land cereal crop matching which provides better result on suitability analysis. Alongside, the system also works for different geographical location with various land character since the case library is filled by real cases with raw cereal crop requirement so it is not location dependent. Generally, in these approach GIS can store, display and generate spatial information related to suitability analysis depending on the physical and socio-economic factor for each cereal crops. Whereas CBR has been used as a part of spatial reasoning system to assist spatial/non-spatial retrieval mechanism. In such an integrated system CBR will provide the method for decision support in proposing a solution to a new problem or providing relevant experiences in the process of land cereal crops to the land evaluator.

CHAPTER THREE

CASE ACQUISITION, MODELING AND REPRESENTATION

3.1 Case Acquisition

Case acquisition is the most important process and vital stage in case based reasoning system development. How knowledge is extracted and represented from the available cases to determine the usefulness of the system. Knowledge acquisition process has many steps. Some of them are: selecting a problem to be solved by the system, interviewing an expert, questionnaires, observation, record reviews, codifying the knowledge in some representation language, and refining the case base by testing it and extending its capability (Clancey, 2004). This process is generally called knowledge elicitation.

The process of case acquisition in this research includes some basic activities such as gathering the required cases, analyze them and then identify important factor for matching land cereal crop. In order to acquire required cases for this study both primary and secondary sources of knowledge are used. The techniques used to extract relevant knowledge from these sources are reviewing related documents and manuals, and interviewing domain experts.

3.1.1 Reviewing Related Documents and Manuals

Document analysis involves gathering knowledge from existing documentations. Hence, document analysis has been carried out to acquire explicit knowledge which is found in various secondary sources of knowledge. Accordingly, different source books and article, Internet resources and FAO Guidelines for Land Evaluation (FAO 1976, 1993, and 2007) are the basis of present research and these documents are reviewed and analyzed. Soil database and different maps (thematic) are also the basis of the suitability analysis. Which are administrative base map, topographic maps and aerial photographic of the area (scale 1: 50,000) were used. Those thematic maps are created and edited, overlaid and visualized on ArcGIS software. As the result, relevant and technical knowledge were extracted and structured in a manner that suitable for knowledge modeling and finally knowledge representation.

Therefore, for the land evaluation purpose, as criteria set out in the guideline of FAO for this research, following general foundations are adopted.

- The natural conditions including soil characteristics and agro-ecological factors,
- Land use requirements and ecological requirements of cereal crops,
- Socio-economic conditions, spatial data source and attributive data,

3.1.2 Interviewing Domain Experts

Both structured and unstructured interviews were employed to elicit tacit knowledge from domain experts. Since one of the important step on this research is extracting tacit knowledge which is embedded and personalized in expert's mind. For this reason, ten experts from Ethiopian Institute of Agricultural Research were selected purposively for interview. Thus, a total of ten (10) crop experts who are specialized in land evaluation and land use planning as well as experts who has professional backgrounds specifically on the selected cereal crops have been participated in the interview for this study. These experts were interviewed about major cereal crops requirements so as to matching with farm unit under cultivation, how crop advisory experts interact with different farmers to assist them on crop selection during land cultivation, what are the basic criteria's to be considered during land cereal crop matching, the significance of socio economic factor on crop selection in addition to the physical land suitability. During face to face communication, the information obtained from experts has been recorded manually.

Profiles of domain experts participated in the interview process are presented in the Table 3.1 below.

Table 3.1.2.1: Profile of experts who participated in the knowledge acquisition

No	Educational level of experts Work	Number of experts	Occupational role
1	MSC	2	Crop advisor
2	MSc	3	Plant pathology
3	PHD	4	Researcher and Land evaluator
4	PHD	1	Researcher & national coordinator of crop advisor

These experts indicated in the Table 3.2.1 above are crop professional conducting research and working on land evaluator and crop production. Based on their profession and experience, these experts identified that the common limiting factor that evaluating land cereal crops matching in Ethiopia are soil characteristics, land topography, erosion and climate factor. The experts responded that land evaluation for crop matching mostly carried out using visual identification through observation and laboratory examination. In case of visual identification through observation, the land evaluator is identified by experienced crop experts by applying both tacit knowledge accumulated through experience and explicit knowledge such as using colored and structure of land quality and FAO guidelines books for land evaluation. The experts responded that visual identification method has great advantages during survey time so as to identify soil texture and depth of a land unit. Laboratory examination has a great importance so as to identify soil PH, salinity, organic matter, excess of salts and nutrient retention value for land unit.

Hence, experienced and knowledgeable experts can identify the limiting factors for crops through observation and laboratory examination for evaluating land resource and choosing suitable cereal crop in farm unit. In general, by reviewing different literature as well as by interviewing experts in the domain the researcher identified fourteen (14) major description attributes to develop CBRLCM system which includes Slope, Soil erosion, Average temperature, Rainfall, Drainage, Soil depth, pH value, Organic matter, Texture, Farmer skills, Equipment, Crop rotation, Profitability, Government policy. Depending on the above provided query description, the system able to recommend four (4) classes of suitability analysis for each five cereal crops those are Barley, Maize, Wheat, Sorghum and Tef.

In deed the detail of this knowledge required for easy representation and case based system development that focuses on land evaluation so as to select suitable cereal crop under cultivation is discussed, structured and modeled in the following sections.

3.1.3 Land Evaluation Approach

The FAO framework is an approach for land suitability evaluation, which classifies land in terms of suitability ratings from highly suitable to not suitable. The assessment of land performance is based on its physical suitability for the proposed land utilization types. This will provide estimates of the maximum available suitable area for each type. The approach involves the implementation and interpretation of basic surveys of soils, climate and terrain properties (Wadaey, et al., 2011).

The basic requirements of applying the FAO framework are the selection and definition of land utilization types for which the land is to be evaluated. The requirements of the land utilization types are compared with the land resources. In this process, land resources are described as land qualities and land characteristics.

Moreover, based on the discussion with domain experts during the interview and analyzing various secondary sources (books, articles, guidelines, manuals, Internet), the following knowledge has been elicited for the prototype development of this research work. This knowledge focuses on land evaluation criteria so as to select suitable cereal crop in farm unit under cultivation.

3.1.4. Selection of Land Characteristics (LC) and Land Qualities (LQ)

FAO, 2007 suggested a list of land qualities which should be considered for land suitability assessment, as shown in (Table 3-1.4.1). The selection of land qualities for land suitability classification is based on agronomic experience at research stations and existing farms.

Table 3. 1.4.1: land qualities considered for land suitability assessment

Group	Land Qualities	Land Characteristics	Unit
Soil	Rooting Conditions	Root able Depth	***cm***
	Texture	Soil Texture	Class
	Nutrient Availability	Soil Reaction	PH
	Organic matter	Organic matter	***%***
	Drainage	Available Water Holding Capacity	***class***
Climate	Length of Growing Period	Evapotranspiration	***mm/month***
		Rainfall	mm/month
Erosion	Erosion Hazard	Soil Erosion model (USLE)	***tha~1tyr~***
Topography	Potential for Mechanization	Slope Steepness	%

3.1.5. Crop Requirements

For each land utilization type it is very important to generate the best conditions for its cultivation which ranged between optimal conditions and the conditions that are unsatisfactory. FAO stated that there is no easy solution to the problem of collecting land use requirements data. Therefore, the evaluator has to collect local and regional experiences and compare them in order to evolve knowledge and worldwide experience in this field to identify the best prediction of the land use requirements.

The land characteristics are derived from the standardized crop environmental requirements and were arranged in tabulated format. In addition, notes on land qualities are given for each crop that are later used for evaluating the suitability of each crop. the land suitability class expresses the degree of fitness of a given type of land for a specified land utilization type. The distinction between different classes of suitability depends on the land suitability criteria that control the limits between suitable and unsuitable, between highly suitable, moderately suitable and marginally suitable land. Therefore, land suitability criteria depend again on the criteria for optimal land use.

It is not common to find handbooks on the cultivation of crops giving the perfect local land conditions. Such knowledge must be gathered from domain expert and literature review of optimal crop requirements and used to build the land use requirements. This information and knowledge may then be used to generate the critical limits of land characteristics and qualities. These critical limits are matched with data which are collected from Ethiopia institute agricultural research center to find the land suitability. Therefore, the researcher reviewed different literature regarding to the land use requirements and matched it with the local data and experience which is collected from the EIAR. In this study, similar land characteristics were used for the all type cereal crops. Hereunder the requirements of the five major cereals (maize, tef, wheat, barley and sorghum) crops based on the literature and domain expert where it is summarized below in tables 3.1.5.1-3.1.5.5.

Table 3.1.5.1: Land suitability rating for Land characteristics for barley

Land characteristics	Highly Suitable	Moderately suitable	Marginally Suitable	Not Suitable
Root able Depth (cm)	>100	>100-70	>70-30	<30
Soil Texture (classes)	SL, SL CL L, CL,	S CL, S CL L	L S, S L	S

Available Water-holding Capacity (mm)	Well	Moderate	Poor	No
Soil Reaction (PH)	6.2-8	>6.2-5.3	>5.3-5	<5
Organic matter	>1.2	1.2-0.8	<0.8	>1.2
Soil Erosion (ton-i ha-i yr-i)	0-2	2-5	5-7	>7
Slope (%)	0-2	>2-4	>4-8	>8
Mean temp. of the growing cycle	10-6; 10-18	6-4 18-24	4-2 24-28	<2 >28
Mean RF of the growing cycle	650-300 650-1100	300-200 1100-1300	200-150 1300-1500	<150 >1500

Clearly describes that barley is produced in areas with an average temperature in the growth cycle of less than 20 ^{o}C and in a well distributed precipitation of rainfall 300-1100mm in the growing cycle. The general pH requirement of barley crop is between 5.5 -8.5 with an optimum pH of 6.2 - 8.0. Moreover, barley prefers well-drained, warm soils with a medium texture at a depth between 1.5-2.0m. The main requirements for barley are summarized in Table *3.1.5.1*.

Table 3.1.5.2: Land suitability rating for Land characteristics for wheat

Land characteristics	Highly Suitable S1	Moderately suitable S2	Marginally Suitable S3	Not Suitable N
Root able Depth (cm)	>120	>80-120	>50-80	<50
Soil Texture (classes)	SL, SL CL L,	S CL, S CL	LS, SL	S

	CL, L,	L		
Available Water-holding Capacity (mm	Well drained	Moderate	Poor	No drainage
Soil Reaction (PH)	6.5 -8	>5.5-6.5	>5-5.5	<5
Organic Matter (%)	>1.5	<1.5-1	<0.5-1	<0.5
Slope Steepness (%)	0-2	>2-4	>4-8	>8
Soil Erosion (ton-i ha-i yr-i)	0-2	2-5	5-7	>7
Mean temp. of the growing cycle	10-16; 10-18	6-4 18-24	4-2	-<2 >28
Mean RF of the growing cycle	650-300 650-1100	300-200 1100-1300	200-150 1300-1500	<150 >1500

In this study, like barley similar land characteristics were used for wheat crops. Also barley grows in most highland parts of Ethiopia Wheat needs at least 200 mm of well-distributed rainfall. Soils best suited to wheat are sandy loam to clay loam texture with good internal drainage. The optimal soil depth is more than 0.9 m. The optimum soil pH ranges between 6.2 and 8. High air humidity combined with high temperature causes wheat rust disease. The main requirements for wheat are summarized in Table *3.1.5.2*.

Table 3.1.5.3: Land suitability rating for Land characteristics for maize

Land characteristics	Highly Suitable S1	Moderately suitable S2	Marginally Suitable S3	Not Suitable N
Root able Depth (cm)	>2m	1.5-1.9	1.0-1.4	<0.9
Soil Texture (classes)	Loamy 8s Silty loam	Clay loam	Silt-clay	Clay

Available Water-holding Capacity (mm	Well drained	Moderate	Poor	No drainage
Soil Salinity (EC) ds/m	0-6	>6-7.4	>7.4-9.5	>9.5
Soil Reaction (PH)	5.8-7.8	7.8-8.5	5.2-5.8	>8.5/ <5.2
Organic Matter (%)	>1.5	<1.5-1	<0.5-1	<0.5
Mean Growing Temp ^{o}C	22-32	20-16/ 32-35	16-14/ 35-40	22-32
Total Growing Season RF/ MARF	400-900	300-400/ 900-1,200	300-150/ 1,200 -1,400	400-900
Soil Erosion (ton-i ha-i yr-i)	0-2	2-5	5-7	>7
Slope Steepness (%)	1500-2200	1000-1500/ 2200-2400	2400-3000	<1000/ >3000

As depicted from the table 3.1.5.3. Maize grows in the temperature range of 14-40^{0}C and its optimal growth is at temperatures between 18^{0}C and 32^{0}C. Like other crops, maize has a tolerance to a wide range of environmental conditions. Maize favors a well-drained, well aerated with deep loam and silt loam soils with an adequate organic matter presence. Maize has a maximum potential rooting depth to a distance of 2 meter with suitable pH ranges from 5.2-8.5, but with the optimal suitability of 5.8-7.8. Hence understanding physical and chemical parameters of the soil for a given farm land is mandatory to obtain optimal maize production.

Table 3. 1.5.4: Land suitability rating for Land characteristics for sorghum

Land characteristics	Highly	Moderately	Marginally	Not

	Suitable S1	suitable S2	Suitable S3	Suitable N
Rootable Depth (cm)	>90	90-50	50-20	<20
Soil Texture (classes)	Sic, si, sc, 1 si, sil,	SI, c>60v	S, Is, s, lcs	cm, sicm, cs
Available Water-holding Capacity (mm	Moderate	Poor	Well	No
Soil Reaction (PH)	5.5-8.2	5.3-5.5/ 8.2-8.3	5.2-5.3/ 8.3-8.5	>8.5/ <5.2
Organic Matter (%)	>1.5	<1.5-1	<0.5-1	<0.5
Mean Growing Temp ^{O}C	24-32	23-18 / >32	18-15	-
Total Growing Season RF/ MARF	400-900	300-400/ 900-1,200	300-150/ 1,200 -1,400	400-900
Soil Erosion (ton-i ha-i yr-i)	0-2	2-5	5-7	>7
Slope Steepness (%)	500- 1,500/1,800	1,500-2,400		>2,500/ <500

Sorghum shows a marked preference for loamy, deep, and well drained. The temperature range for the growth of sorghum is between 0 and 38°C. However, the growth is optimal at temperature between 15 and 32°C. Otherwise flowering will be poor or will not happen at all. The optimum soil pH ranges between 5.5 and 8.2. Summary of the land requirements for sorghum is shown in (*Table 3.1.5.4*).

Table 3. 1.5.5: Land suitability rating for Land characteristics for Tef

Land characteristics	Highly Suitable	Moderately suitable	Marginally Suitable	Not Suitable
Root able Depth (cm)	>50	25-50	10-25	<10
Soil Texture (classes)	SiL, SiC, SC, CL, , Si, C	L, SCL, SL,HC,SiC L	LS	s
Available Water-holding Capacity (mm	Well drained	Moderate	Poor	No drainage
Soil Reaction (PH)	5.2-7.5	5-5.2;7.5-8.0	4.5-5.0; 8.0-8.5	<4.5 >8.5
Organic matter	>1.0	0.8-1.0	0.5-0.8	<0.5
Mean Growing Temp oc	15-20	12.5-15; 20-25	25-30	<12.5 >30
Total Growing Season RF/ MARF	400-550	300-400; 550-800	200-300; 800-1200	<200 >1200
Soil Erosion (ton-i ha-i yr-i)	0-2	2-5	5-7	>7
Slope Steepness (%)	0-5	5-15	15-30	>30

Tefs shallow rooted crop that grows better on soils that have high clay content which retain soil moisture well and is tolerant of waterlogged conditions in the early vegetative stage. Because of the shallow rooting system, t*ef* crop is easily exposed to water erosion and is more suitable to grow in a flat or gentle slope and if the altitude is above 1900m, the maximum yield of *tef* can be obtained with the pH >7.5 and poorest yield below pH <5.5.

For this study the researcher extract cases and the case library is build based on the physical land suitability analysis and FAO Framework which provides details about the socio-economic factor on suitable crop selection so as to matching land characteristics with crop requirements to produce five layer namely: Soil, climate, erosion hazard, slope and socio-economic which are important for land suitability for the selected crops. These layers were integrated into the GIS environment and then the overall land suitability analysis result presented as information for each kind of selected crops.

3.1.5.1. Soil layer

The available spatial information to this study were collected from Ethiopian institute of agricultural research. The thresholds values for crop requirements are deduced from domain expert with discussion and studies in Agriculture Resource Centre in Ethiopia and other sources. The details of requirements for each crop were obtained from literature and data provided by the Agriculture Research Centre in Ethiopia. Physical and chemical soil characteristics were stored in spreadsheet model in excel (see Figure 3.1.5.1.1). "if" function was used to set the suitability class; each type of soil takes a degree of suitability class for each soil characteristics by matching soil characteristics with crop requirements for each crop (See Figure 3.1.5.2).

D	E	F	G	H
Drainage	soil depth(cm)	pH	organic matter	texture
well	65	6.1	1.1	loam sandy
well	90	7.9	1.2	loam
well	200	7.6	1.1	Sandy loam
Moderate	85	8.1	1	clay loam
Poor	50	1.9	2	clay loam
Moderate	96	8	2	sandy loam
Poor	22	8.5	1.1	clay
well	160	6.2	1.3	silt loam
well	156	6.4	1.4	silt loam
Moderate	142	6.5	1.5	loam
Moderate	122	8.4	2.5	sand
Poor	41	5.1	1.3	loamy sand
Poor	15	5	2.3	loamy sand
Moderate	55	9.2	1.1	clay
well	251	6.6	0.8	sandy clay loam
well	301	8	1.5	sandy loam
poor	95	5	1.9	clay
Moderate	150	7.9	2	loam
well	105	6.1	2.3	Sandy clay loamy
well	190	7.1	1.8	silt loam
well	200	7.6	1.7	silt loam
Moderate	85	8.1	1.1	sandy loam
Poor	88	1.9	1.2	clay loam
Moderate	96	8	1.1	sandy loam
Poor	25	8.5	1	clay loam
well	168	6.2	2	silt loam
well	189	6.4	2	silt loam
Moderate	142	8	1.1	Sandy loam
Moderate	122	8.7	1.3	sand
No drainage	41	5.1	1.4	loamy sand

barley | wheat | maize | sorgem | Teff | socio-economic

Figure 3.1.5.1.1: Database for soil characteristics

The overall soil suitability classes for each crop was determined using Mode function in excel and exported from the spreadsheet model to the soil classification layer in GIS and then the soil layer was created based on the requirements for each crops.

crops		barley				maize					wheat				
		suitability classes					suitability classes					suitability classes			
land use requirement	unit	S1	S2	S3	N	unit	S1	S2	S3	N	unit	S1	S2	S3	N
land characteristic(LC)															
Drainage	class	Well drained	Moderate	Poor	No draina	class	Well drair	Moderate	Poor	No draina	class	Well drair	Moderate	Poor	No drainage
soil depth	cm	>100	>100-70	>70-30	<30	m	>2	1.5-1.9	1.0-1.4	<0.9	cm	>120	>80-120	>50-80	<50
soil reaction(PH)	PH	6.2-8	>6.2-5.3	>5.3-5	<5	PH	5.8-7.8	7.8-8.5	5.2-5.8	>8.5/ <5.2	PH	6.5 -8	>5.5-6.5	>5-5.5	<5
organic matter	%	>1.2	1.2-0.8	<0.8	>1.2	%	>1.5	<1.5-1	<0.5-1	<0.5	%	>1.5	<1.5-1	<0.5-1	<0.5
texture	class	SL, SL CL L,	S CL, S CL	LS, SL	s	class	Loamy 8s	Clay loam	Silt-clay	Clay	class	SL, SL CL L	S CL, S CL	LS, SL	s

		barley													
land use requirement	unit														
land characteristic(LC)		case1	case2	case3	case4	case5	case6	case7	case8	case9	case10	case11	case12	case13	case14
Drainage	class	1	1	1	2	3	2	3	1	1	2	2	3	3	2
soil depth	cm	2	2	1	2	3	2	4	1	1	1	1	1	3	4
soil reaction(PH)	PH	1	1	1	2	4	2	2	1	1	1	2	4	4	4
organic matter	%	1	1	1	2	1	1	2	2	2	1	1	2	2	2
texture	class	1	1	2	1	1	2	2	1	1	1	1	2	2	2
	overall suitab	1	1	1	2	1	2	2	1	1	1	1	2	2	4

suitability model / collected data / Sheet1 / barley / maize / wheat / sorgem / Teff / socio-economic / 1

Figure 3.15.1.2: Suitability Model for the study area

caseid	suitability class
case1	1
case2	1
case3	1
case4	1
case5	2
case6	1
case7	2
case8	3
case9	1
case10	1
case11	1
case12	3
case13	3
case14	2
case15	1
case16	1
case17	2
case18	2
case19	1
case20	1
case21	1
case22	1

... | collected data | **maize** | barley | wheat | sorgem | Teff | socio-economic

Figure 3.1.5.3: Land suitability for maize

1 is high suitability (S1), 2 is moderate suitability (S2), 3 is margin suitability and 4 means not suitable (N)

3.1.5.2. Slope Layer

Slope has been considered as one of the evaluation parameters in cereal crop suitability analysis. Based on the four slope classes (S1, S2, S3 and N). The slope layer was produced from the contour map of the specific area map which was collected from Ethiopian institute of agricultural research center. Surface function of ArcGIS was applied to convert contour map to slope grid information. The classes of Slope layer were produced according to slope suitability categories required for each crops. It is clear that the slope is most limiting factor so as to determine land suitable evaluation is shown in Figure 3.4.

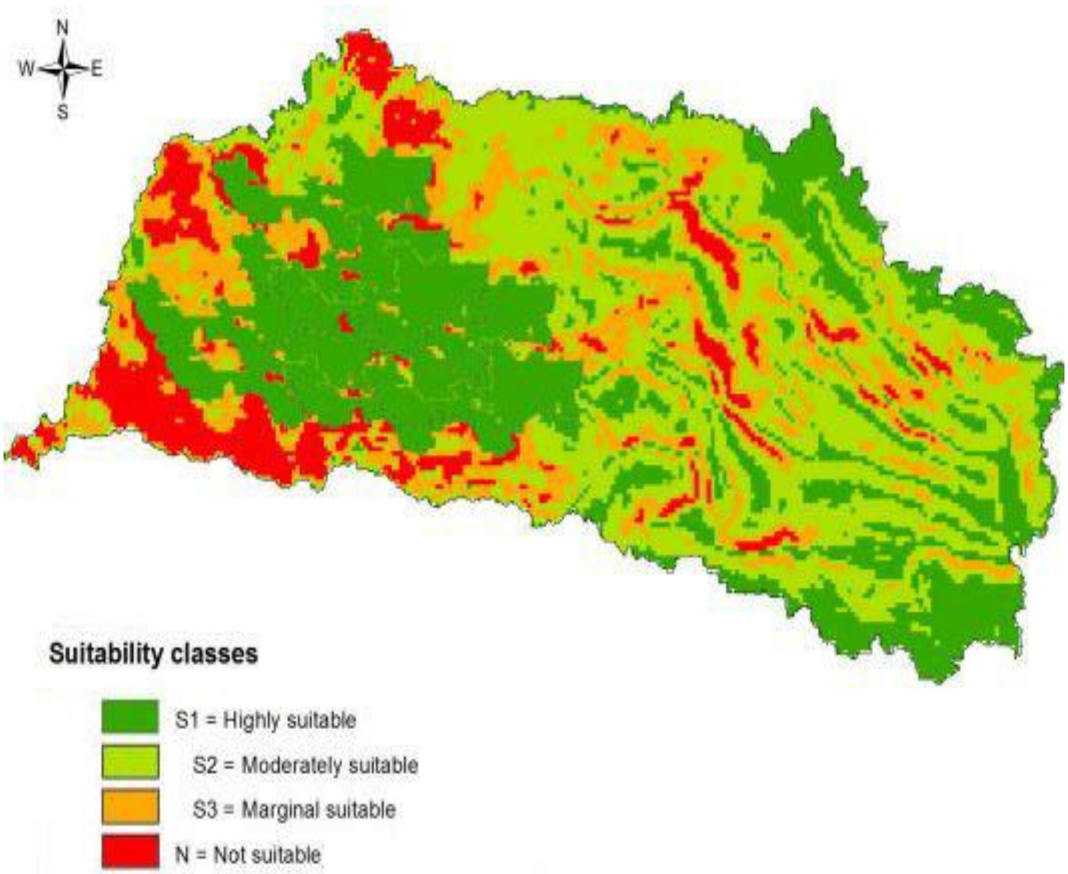

Figure 3.1.5.2.1: Slope suitability map

3.1.5.3. Climate Layer

Temperature and rainfall are the two main climatic factors that can affect land suitability. Rainfall is used to determine the Length of Growing Period. This term refers to the period of the year in which agricultural production is possible from the viewpoint of moisture availability and absence of temperature limitations. In addition, the amount of soil moisture stored in the soil profile can be taken into account. Temperature determines the distribution of crops, soil and farming system of a certain area. To produce climate layer, the suitability classes for the average rainfall and temperature were exported to GIS. Climate layer was created in ArcGIS for each crops.

3.1.5.4. Erosion Layer

Erosion layer gives information or indicators about predicted hazard erosion that could be occur in the area. The value of erosion hazard was exported to ArcGIS to produce suitability class for each crops. The erosion hazard layer was produced by reclassify the value into four classes namely: high suitability (S1), moderate suitability (S2), margin suitability (S3) and non-suitable (N).

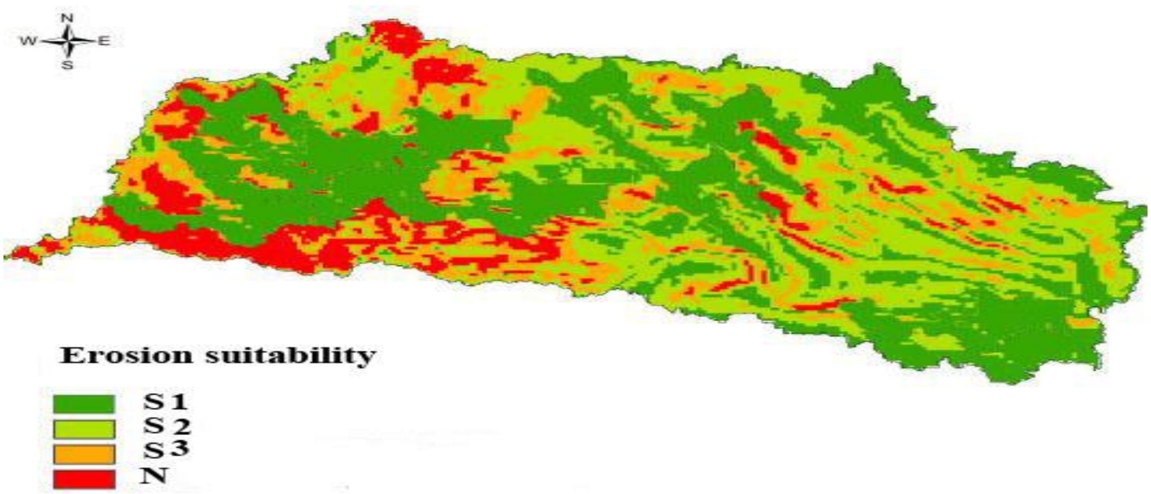

Figure 3.1.5.4.1: Soil Erosion layer

3.1.5.5. Socio-economic layer

The FAO framework is simple and broadly aids at developing a land evaluation model. Based on FAO guideline and domain expert interview the researcher making analysis for socio-economic factor prepared different alternative cases in Excel which yields final weight of the sub-criteria according to its influence. Weight of the alternatives corresponds with importance of the criteria in the cereal cultivation in farm unit. Furthermore, the alternative descriptive cases for the solution have close relationship with each other and the solution part must be addressed at the same time for more logical feedback. These attributive characteristics are incorporated into GIS based data of land suitability evaluation as a system. Use of social and economic parameters provides reflection of real situation. GIS technique used for the land suitability evaluation improves efficacy of the outputs of the evaluation.

Hence, it has been adapted in the current research that economic, social and political factors which influence the crop selection process. Social and cultural characteristics influence cultivation pattern land evaluation is effected by social attributes. Importance of social parameter for suitability classification should be raised from the opinion of individual and corresponding weight can be given for analysis purpose. These parameters are then classified under socio-economic class as sub criteria for evaluation purpose. The weight of criterion and sub criteria give with the discussion of expert and reviewing FAO framework for socio-economic factor. Hence, it is also possible to provide remedial measures or explanations in order to switch to a second choice in case the user is not willing to settle for the first crop suggested by the system.

Land suitability analysis for cereal crops are the cumulative use of all the sectors including physical land characteristics, economic parameters and social traits. Physical parameters are prerequisite; so multi-criteria evaluation will be proceeding only if land area is physically suitable. Land area with suitability level S1 S2 and S3 are considered for incorporation of social and economic component for further analysis. Non suitable land areas (N) are omitted from suitability evaluation.

A	B	C	D	E	F	G	H	I	J
farmer skills	**equipment**	**crop rotatic**	**profitibility**	**government p**	**barley suitability**	**maize suitabilit**	**wheat suitabi**	**sorghum sui**	**Teff suitability**
wheat	yes	maize	wheat	wheat	moderately	moderately	highly	marginally	moderately
barley	yes	sorghem	barley	barley	highly	highly	moderately	highly	moderately
maize	no	wheat	maize	maize	moderately	highly	marginally	moderately	marginally
sorghum	no	barley	sorghum	sorghum	marginally	moderately	marginally	highly	moderately
teff	yes	maize	teff	teff	marginally	moderately	marginally	marginally	highly
wheat	yes	sorghem	wheat	wheat	moderately	marginally	marginally	moderately	marginally
barley	yes	teff	barley	barley	highly	highly	moderately	moderate	moderately
maize	no	barley	maize	maize	moderately	highly	marginally	marginally	moderately
sorghum	no	wheat	sorghum	sorghum	highly	moderately	highly	highly	marginally
teff	yes	maize	teff	teff	highly	highly	highly	highly	highly
wheat	no	barley	wheat	barley	highly	marginally	highly	marginally	marginally
wheat	no	sorghem	barley	maize	highly	highly	highly	highly	moderately
barley	yes	maize	maize	sorghum	moderately	highly	marginally	moderately	marginally
maize	yes	sorghem	sorghum	teff	marginally	moderately	marginally	highly	moderately
sorghum	no	wheat	teff	wheat	marginally	moderately	highly	moderately	highly
wheat	yes	maize	wheat	wheat	moderately	moderately	highly	marginally	moderately
barley	yes	sorghem	barley	barley	highly	highly	moderately	highly	moderately
maize	no	wheat	maize	maize	moderately	highly	marginally	moderately	marginally
sorghum	no	barley	sorghum	sorghum	marginally	moderately	marginally	highly	moderately
teff	yes	maize	teff	teff	marginally	moderately	marginally	marginally	highly

... collected data | barley | wheat | maize | sorgem | Teff | socio-economic

Figure 3.1.1.5.1 - socio-economic database suitability

3.1.6. Overall suitability evaluation in ArcGIS

A land suitability model was constructed using GIS capabilities and modeling functions. The GIS Model Builder was used to organize and integrate spatial processes to model the land suitability. The spatial and non-spatial factors (soil, climate, slope, socio-economic and erosion) were integrated into the GIS environment as information layers and overlaid to produce overall land suitability assessment for a particular land utilization type.

The overall land suitability evaluation was produced from the spatial overlay of five layers namely; Soil, climate, slope, socio-economic and erosion layer ArcGIS software. The overall suitability evaluation for barley, wheat, maize, sorghum and tef were produced by using the weighted overlay technique in ArcGIS. The weighted overlay technique allows different weights to be applied for different thematic layers. The weighting values of each layer are given depending on the importance of each layer. In this study the weighted values were 25% for soil, and 20 % for climate, slope and socio-economic, 15% erosion layers. These values were supported by the reviewing FAO guideline and discussion with domain expert's s in Ethiopia institute of Agricultural research center. To find suitable class for each crop, a suitability model was created using model builder in Arc tools box and tools from spatial analysis tool sets. Then, after their suitability was assessed, the land suitability factors which were considered in this study are Soil, climate, slope, socio-economic and erosion was used as the input for each crops suitability model to find the most suitable class as shown in Figure 3.2. And in general land suitability analysis for agricultural crops was done using model builder and Flow chart showing the processes involved.

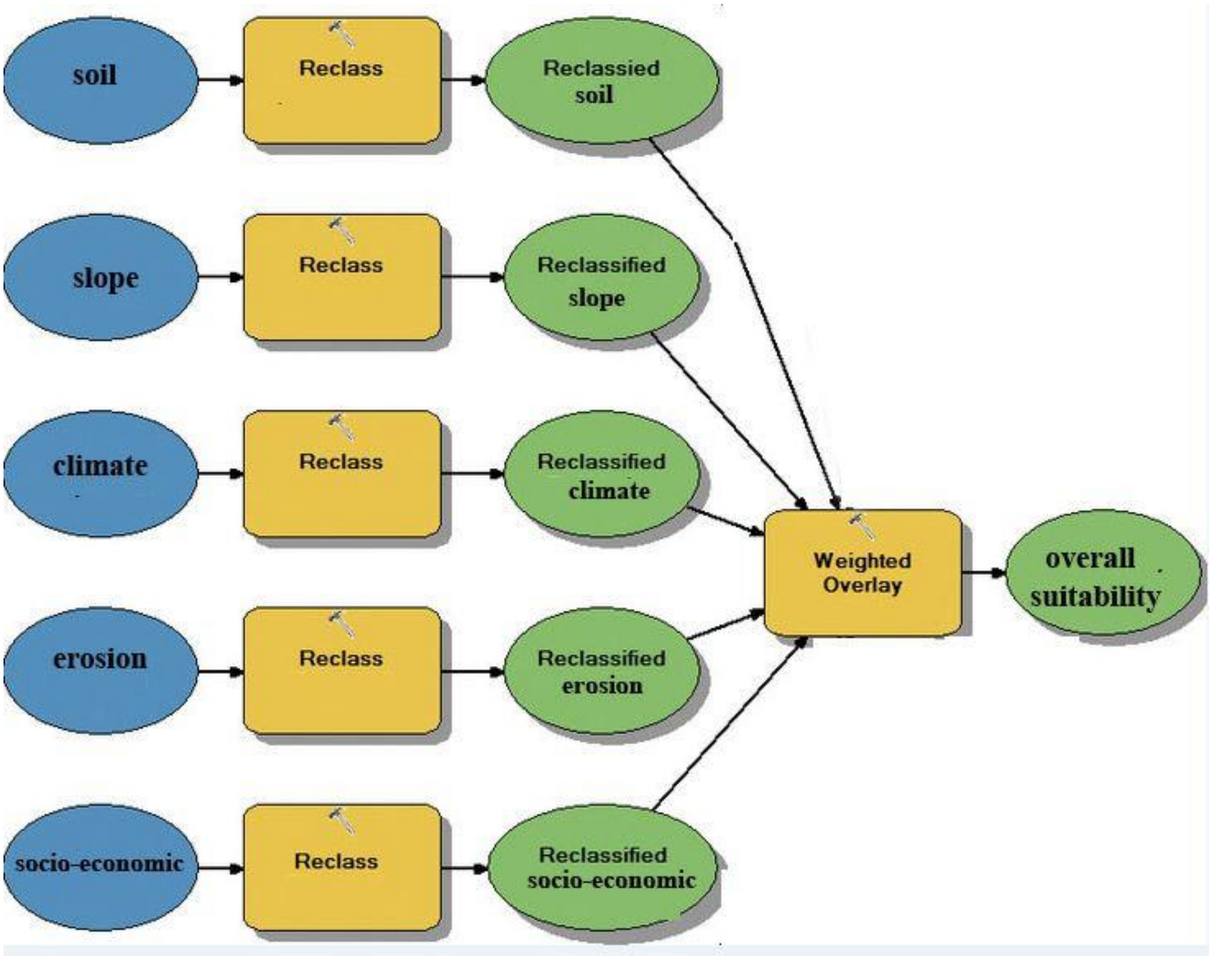

Figure 3.1.2.1: Suitability model

The data derive from the process of generalizing and analyzing information on land assessment, including land unit map, land use requirement and adapting classification analysis. Map of land suitability is generated with reasonable overlay of necessary thematic layers. It has basic spatial information layer. The data consist of attribute data and spatial distribution of the layer. The information layer in the thematic layer map of consist of the main information source is land unit data presented on land unit map to establish connection with land use requirements of cereal crops. To the degree the agronomical requirement of the cereal crop satisfied by land quality gives the measure of suitability level. The detailed will be described and analyzed in the preceding sections. Other non-spatial information for land suitability assessment including socio-economic factors are also a system of data organized in the form of database in excel. Upon rational analysis of the available data, suitability ratings according to domain expert and FAO guideline system of land evaluation was developed.

Regular grid of cells that is placed over specific location for each cell, thematic information is recorded about the underlying location, e.g. land use at the location. Each theme can be represented by a grid layer. In vector-based models, points, lines and polygons represent the geographical objects. Present study uses vector based model for each of the land mapping unit as geographical objects for every thematic maps. Land areas of single land mapping units are not homogenous. Thematic information about these geographical objects is stored in a separate attribute table. Every row in the table represents one geographical object and every column represents an attribute describing the geographical objects. A unique identification ID is attached to each object in order to link the attribute table to the geographical objects. Creating this link is only possible when every row in the attribute table starts with the attribute ID of the corresponding geographical object. Thematic map consists of information distributed spatially in the units of the map itself. Such units in the map are names as Land Mapping Unit (LMU). They are the basic constituent of thematic map. Therefore, LMU is defined as a land plot which is specifically identified in the map with distinct characteristics. The characteristics within each LMU homogeneously distributed differences or fluctuations in the land attributes will lead to the further fragmentation of the land mapping units into smaller pieces. From the agricultural point of view, each LMU is more suitable for specific types of cultivated cereal crops in the existing land condition. Each LMU is the individual entity of an area which bears characteristics of its own for example soil characteristics like soil texture, root able soil depth, organic matter.

In this manner, a geo-processing model (see Figure 3.7) - intersection of grid layer with other GIS data layers- is applied to compute analysis information, which drops to each cell unit. A sample geo-processing result is presented in Figure 3.8. By the finishing of this step, cellular representation of GIS based archive (main part of the case library) will be completed. Cellular representation is used only to generate “GIS based spatial unique cases” for land evaluation spatial structures with cellular automata technique.

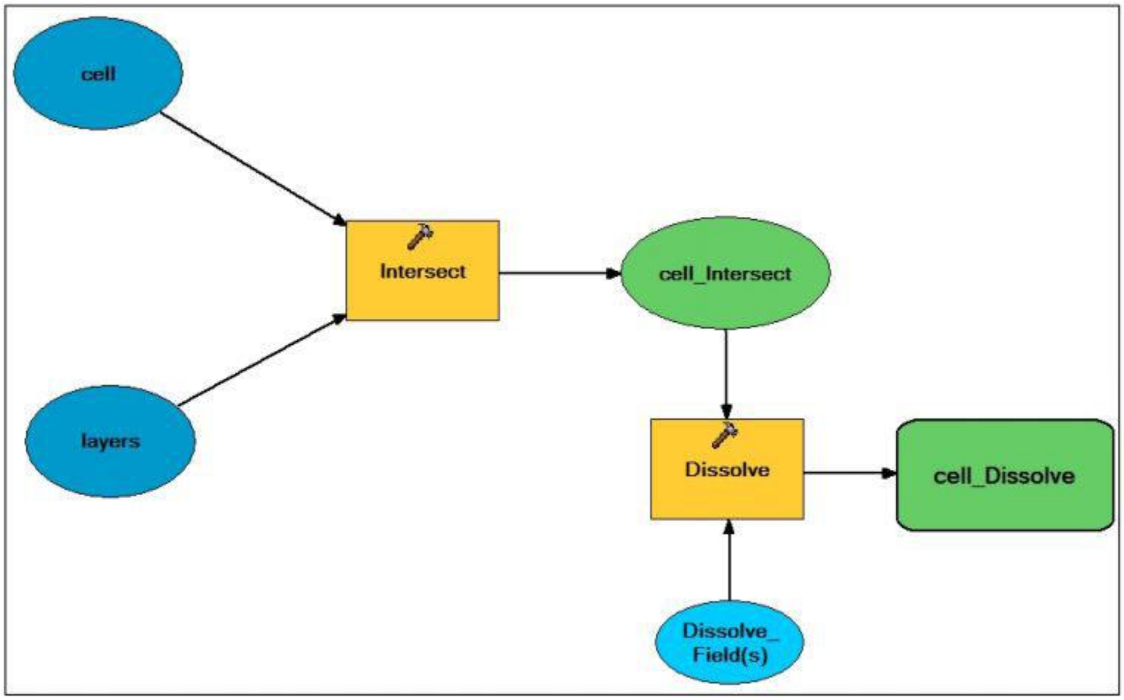

Figure 3.1.6.2: Intersection of layers by ArcGIS model builder

This lattice space and other data layers are used to build a geodatabase which constitutes the spatial core of the case library. Another advantage of using this geo-processing function is structural flexibility of the case library. For instance, if an additional data layer is added (or removed); case library could be updated by the execution of the model.

After the intersection process completed, final database (non-spatial) is produced by joining the necessary columns of all intersected layers (see Table3 .12). Another simple geo-processing model is utilized to perform this joining operation to create the database.

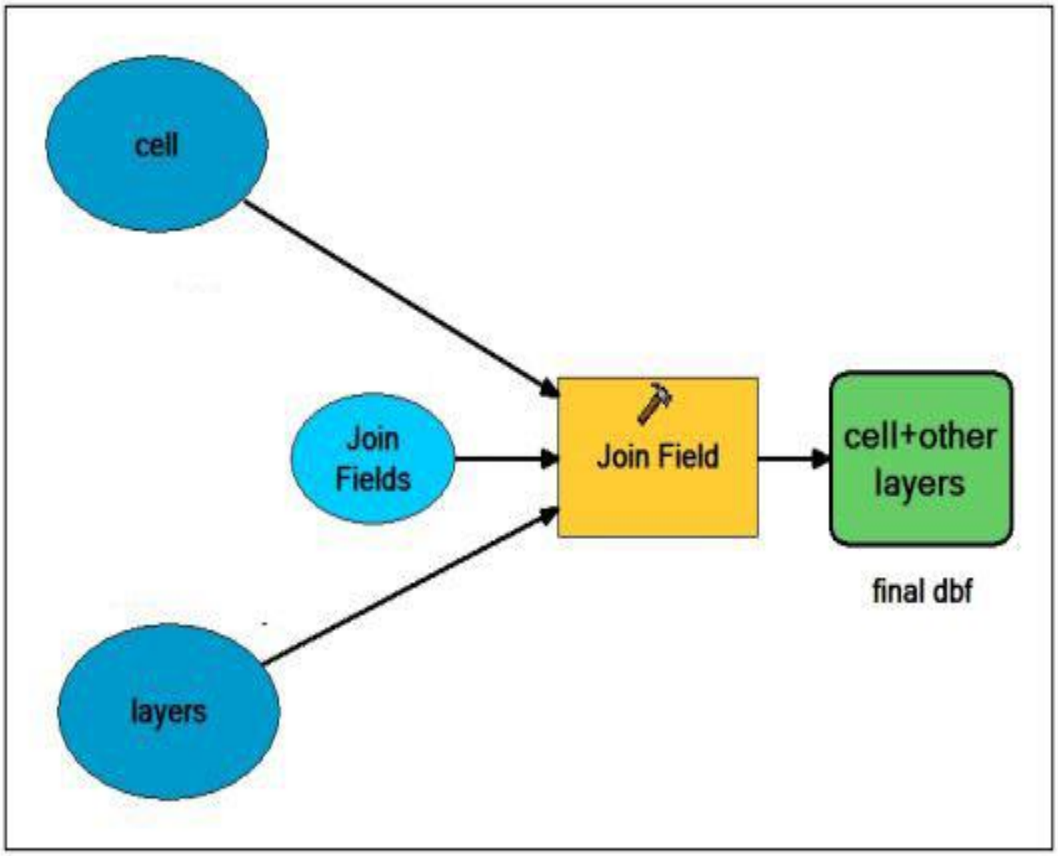

Figure 3. 1.6.3: Joining the required fields of layers by ArcGIS Model builder

The database can be defined as an organized collection of non-redundant data in a computer so that it can be expanded, updated, retrieved and shared by various applications (Malczewski, 2004). Feature values are stored in a table file "cell.dbf" (dbase data file format) in ArcGIS software.

caseID	slope(%)	soil erosion	Drainage	soil depth(c	pH	organic matte	texture	average tem	Rainfall	farmer skills	equipment	crop rotatio	profitibility	governmen	barley suit	maize suital	wheat suital	sorghum s	Teff sui
case1	2	very low	well	65	6.1	1.1	loam sandy	22	700	wheat	yes	maize	wheat	wheat	highly	moderately	highly	marginally	moderate
case2	3	very low	well	90	7.9	1.2	loam	22	700	barley	yes	sorghem	barley	barley	highly	highly	highly	highly	moderate
case3	2	very low	well	200	7.6	1.1	Sandy loam	21	650	maize	no	wheat	maize	maize	highly	highly	marginally	moderately	marginall
case4	4	very low	Moderate	85	8.1	1	clay loam	21	650	sorghum	no	barley	sorghum	sorghum	highly	highly	moderately	highly	moderate
case5	11	moderate	Poor	50	1.9	2	clay loam	28	850	teff	yes	maize	teff	teff	moderately	highly	moderately	marginally	highly
case6	10	moderate	Moderate	96	8	2	sandy loam	28	850	wheat	yes	sorghem	wheat	wheat	moderately	highly	moderately	moderately	marginall
case7	17	moderate	Poor	22	8.5	1.1	clay	26	350	barley	yes	teff	barley	barley	highly	highly	moderately	moderate	moderate
case8	2	very low	well	160	6.2	1.3	silt loam	26	350	maize	no	barley	maize	maize	moderately	highly	marginally	marginally	moderate
case9	4	very low	well	156	6.4	1.4	silt loam	28	450	sorghum	no	wheat	sorghum	sorghum	highly	highly	highly	highly	marginall
case10	3	very low	Moderate	142	6.5	1.5	loam	28	450	teff	yes	maize	teff	teff	highly	highly	highly	highly	highly
case11	5	low low	Moderate	122	8.4	2.5	sand	18	450	wheat	yes	maize	wheat	wheat	highly	marginally	highly	marginally	marginall
case12	12	moderate	Poor	41	5.1	1.3	loamy sand	18	450	barley	yes	sorghem	barley	barley	highly	highly	highly	highly	moderate
case13	13	moderate	Poor	15	5	2.3	loamy sand	18	450	maize	no	wheat	maize	maize	highly	highly	marginally	highly	marginall
case14	8	low	Moderate	55	9.2	1.1	clay	28	410	sorghum	no	barley	sorghum	sorghum	marginally	highly	marginally	highly	moderate
case15	1	very low	well	251	6.6	0.8	sandy clay loam	28	410	teff	yes	maize	teff	teff	marginally	highly	highly	highly	highly
case16	3	very low	well	301	8	1.5	sandy loam	28	410	wheat	yes	sorghem	wheat	wheat	highly	highly	highly	marginally	moderate
case17	8	low	poor	95	5	1.9	clay	28	390	barley	yes	teff	barley	barley	highly	highly	moderately	highly	moderate
case18	10	low	Moderate	150	7.9	2	loam	25	390	maize	no	barley	maize	maize	moderately	highly	marginally	moderately	moderate
case19	5	very low	well	105	6.1	2.3	Sandy clay loamy	25	440	sorghum	no	wheat	sorghum	sorghum	marginally	highly	moderately	highly	moderate
case20	4	very low	well	190	7.1	1.8	silt loam	25	440	teff	yes	maize	teff	teff	moderately	highly	marginally	marginally	highly
case21	2	very low	well	200	7.6	1.7	silt loam	23	540	wheat	no	barley	wheat	barley	highly	marginally	moderately	highly	marginall
case22	7	low	Moderate	85	8.1	1.1	sandy loam	23	540	wheat	no	sorghem	barley	maize	highly	highly	moderately	moderate	moderate
case23	9	low	Poor	88	1.9	1.2	clay loam	23	540	barley	yes	maize	maize	sorghum	highly	highly	moderately	marginally	moderate
case24	3	very low	Moderate	96	8	1.1	sandy loam	30	560	maize	yes	sorghem	sorghum	teff	highly	moderately	highly	highly	marginall
case25	12	low	Poor	25	8.5	1	clay loam	30	560	sorghum	no	wheat	teff	wheat	moderately	highly	marginally	highly	moderate
case26	3	very low	well	168	6.2	2	silt loam	30	570	wheat	yes	maize	wheat	wheat	highly	marginally	highly	marginally	marginall
case27	2	very low	well	189	6.4	2	silt loam	30	570	barley	yes	sorghem	barley	barley	highly	highly	highly	highly	moderate
case28	8	low	Moderate	142	8	1.1	Sandy loam	31	580	maize	no	wheat	maize	maize	moderately	highly	marginally	moderately	marginall
case29	10	low	Moderate	122	8.7	1.3	sand	31	580	wheat	yes	maize	wheat	wheat	marginally	moderately	marginally	highly	moderate
case30	19	high	No drainage	41	5.1	1.4	loamy sand	31	580	barley	yes	sorghem	barley	barley	moderately	moderately	highly	moderately	highly
case31	22	high	No drainage	15	5	1.5	clay loam	33	630	maize	no	wheat	maize	maize	highly	moderately	highly	marginally	highly
case32	10	low	Poor	55	9.2	2.5	clay	33	630	sorghum	no	barley	sorghum	sorghum	highly	highly	highly	highly	moderate
case33	2	very low	well	251	6	1.3	sandy clay loam	23	670	teff	yes	maize	teff	teff	highly	highly	marginally	moderately	marginall
case34	3	very low	well	301	8	2.3	sandy loam	23	670	wheat	yes	sorghem	wheat	wheat	highly	highly	moderately	highly	moderate

Figure 3.1.6.4: Final database (dbf file) after joining all intersected layers

The CaseID is a number that uniquely identifies the case; also all cells have a unique name/number, for instance “case2”. The descriptive features (columns) are critical parts of the case and they represent the key feature of a case, such as “drainage”, “soil depth”, “slope”. This final data base used to a case library for a jCOLIBRI which is converted into plain text format so as to compatible with jcolibri.

3.2. Conceptual Case Modeling

After the cases are is acquired from different sources, the next step was organizing and structuring of the knowledge applied for each case. During the case modeling stage the domain knowledge that is elicited by various techniques is represented in a knowledge model. A knowledge model is a structured representation of knowledge using symbols to represent pieces of knowledge and the relationships between them (Calu, 2009). Knowledge models include representational character based on languages such as diagrammatic and logic representations in the form of networks and ladders, tabular representation such as matrices and structured text such as hypertext. The generation and modification cycle of a knowledge model is an essential part of the knowledge modeling phase. The model helps to ensure that all stakeholders in a scheme understand the language and terminology being used and quickly takes information for authentication and adjustment where necessary. The knowledge models are also of great value during cross-validation with other knowledge domain (Emberey et al., 2007).

During the case acquisition stage, most of the case is unstructured and often in tacit form. The knowledge engineer will try to understand both the tacit and the explicit part of the knowledge and then use simple visual diagrams to stimulate discussion amongst users and knowledge experts. The knowledge engineer then has to construct the conceptual model from what has been discussed during the knowledge acquisition stage. This communicates the knowledge to the information specialist who will transform the model into workable computer programs or codes. This approach is similar to that of software engineering where models are used to represent user requirements.

Therefore, knowledge acquired through different knowledge acquisition techniques and it can be modeled with decision tree and hierarchical tree structure. Decision trees are produced by algorithms that identify several ways of splitting a data set into branch like segments. These segments form an inverted decision tree that originates with a root node at the top of the tree. The hierarchical tree diagram provides the analyst with an effective visual condensation of the clustering results. It is also useful in spotting outliers, as these will appear as one member clusters that are joined later in the clustering process. The numbers at the top and bottom of the hierarchical tree diagram represent equally spaced values of the criterion function. It gives a pictorial representation of the criterion function information (Chen et al., 2003).

For this study, hierarchical tree structure was used to represent knowledge modeling. Hierarchical tree structure can easily model concepts and clearly explains the concepts in the problem area. It models the knowledge in the hierarchical manner. This model starts from the main concept at the highest level of the hierarchy and other sub concepts that can affect or affected by the highest level concept put next to down ward in the hierarchy. The context of this hierarchical structure is used to demonstrate clearly the decision making process of land suitable analysis which are implemented by using jCOLIBRI programming tool.

The general structure of creating knowledge modeling contains input, knowledge model and output as shown below (Makfi, 2011). The main inputs for the system are spatial and non-spatial data which include water supply (either rain fed or irrigated), soil type (t Availability, Nutrient Retention, rooting Conditions, Soil Workability and Oxygen Soil Drainage class), topology (degrees of slop), The environmental(erosion), climate (mean temperature, rainfall and length of growing period), socio-economic(profitability, crop rotation, equipment and farmer skills) can be considered factors in which are important to produce land suitability for the farm unit. Finally, the output will be recommendations for the farmer that enables alternative crop choices based on the criteria.

Hierarchical structure was used to model the knowledge. The hierarchical structure as shown in figure 3.2.1 is derived from the knowledge acquired from the consultations of experts and secondary sources. These hierarchical structures are the base for the prototype case based system development and present below.

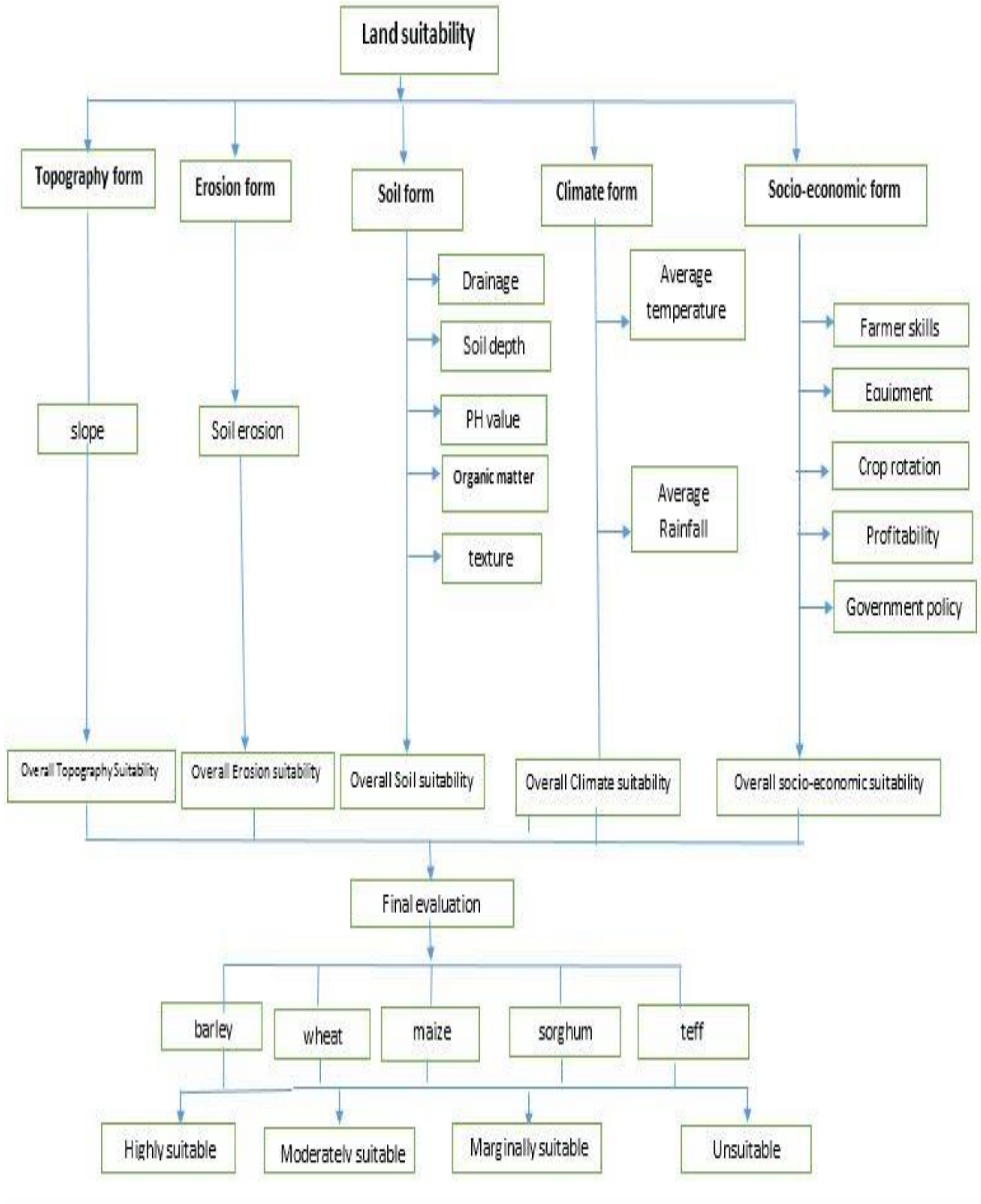

Figure 3. 2.1: hierarchical structure of land evaluation for cereal crops

3.2.1. Procedure for crop suitability analysis

The process of land suitability analysis is the assessment and grouping of specific areas of land in terms of their suitability for a defined use. Suitability of land is assessed considering rational cropping system, for optimizing the use of a piece of land for a specific use in Ethiopia. Suitability is assessed by comparison of the land use requirements with the land qualities. Land suitability could be assessed for current situation (actual land suitability) or after improvement (potential land suitability). There are four categories recognized for classification of land suitability. Land Suitability Orders indicating in the simplest of whether land is suitable (S) or not suitable (N) for specified use. Land Suitability Classes showing the degree of suitability within an order; Land Suitability Sub classes reflection the kinds of limitation or required improvements measures within classes; and Land Suitability Units indicating differences in required management within sub classes.

Ethiopia strategic agricultural land use plans are developed after a preliminary assessment of soil and land resources and they provide a good indication of possible broad land use categories that the region can support. Government, through Land Resource Assessment agencies, usually carries out land surveys techniques to determine suitable land uses for sustainable production in an area and to identify preferred land use types. Such preferred land use types and crops suitability results for a region are published in the form of maps, often the mapping scale range is 1: 50,000. These maps serve the purpose of guiding farmers and landowners to adapt and devise their own land use plans for their farm units.

Crop advisors perceive the land evaluation process as a problem-solving activity. They use their past experience and lessons from past mistakes to suggest crops based on the available resource conditions. Crop advisors make use of soil surveys, land classification results and suitability maps released by Government agencies. The current case-based reasoning for land cereal crops matching system is being developed with the help of the EIAR. The case based system outlined in the current research imitates the problem-solving approach adopted by a crop advisor and it is designed based on the knowledge provided by experts at EIAR. The

CBRLCCM system adopts a combined land evaluation and crop suitability analysis in decision-making. The initial involves identification of a suitable set of crops considering physical spatial attributes. In addition, it involves the user's interaction to identify suitable crops based on various social, economic and political inputs.

The assessment criteria used in the physical evaluation has been adapted from the EIAR Guidelines and Methodology for Assessing Crop Potential and other factors used in economic, socio-political evaluation are gathered from the agricultural experts at EIAR and FAO guideline.

Currently, EIAR uses soil and physiographic data to generate crop-wise suitability maps for a region on a mapping scale of 1: 50,000. These maps play an important role in strategic land use planning in land use decisions. Such plans help farmers to develop their farm-level land use plans. EIAR has developed an assessment methodology for physical land evaluation which evaluates soil and landscape parameters of a region to assess its suitability for some crops. This methodology uses soil and landscape mapping data bases compiled for the agricultural districts of Ethiopia. Each mapped soil landscape unit has been classified with respect to a range of attributes which affect agricultural land use and thereby crop suitability. These attributes include in the model. Each crop has a specific requirement of these attributes and a criteria table is developed to define these crop requirements in terms of these attributes. By matching the values of these land attributes with the requirements of specific crops, the relative particular crop can be predicted.

Each farm unit is evaluated based on a number of physical factors such as topography, waterlogging, drainage conditions, soil depth, water storage, chemical barriers to root growth, soil fertility, and erosion potential to classify the land for its suitability to grow a particular crop. The land is classified into one of the four classes of suitability, from High suitability to Low suitability, as shown in below for example, it produces a potential lands suitable for wheat map (see Figure 3.11), showing land units under four categories (from highly suitable to not suitable) in ArcGIS.

- S1 (highly suitable which is land having no significant limitations to sustained application of a given use.
- S2 (moderately suitable) which is land having limitations which in aggregate are moderately severe for a sustained application of a given use.
- S3 (marginally suitable) is land having limitations which in aggregate are severe for sustained application of a given use and will reduce productivity or benefits.
- N1 (currently not suitable) is land having limitations which may be surmountable in time but which cannot be corrected with existing knowledge at currently acceptable cost.

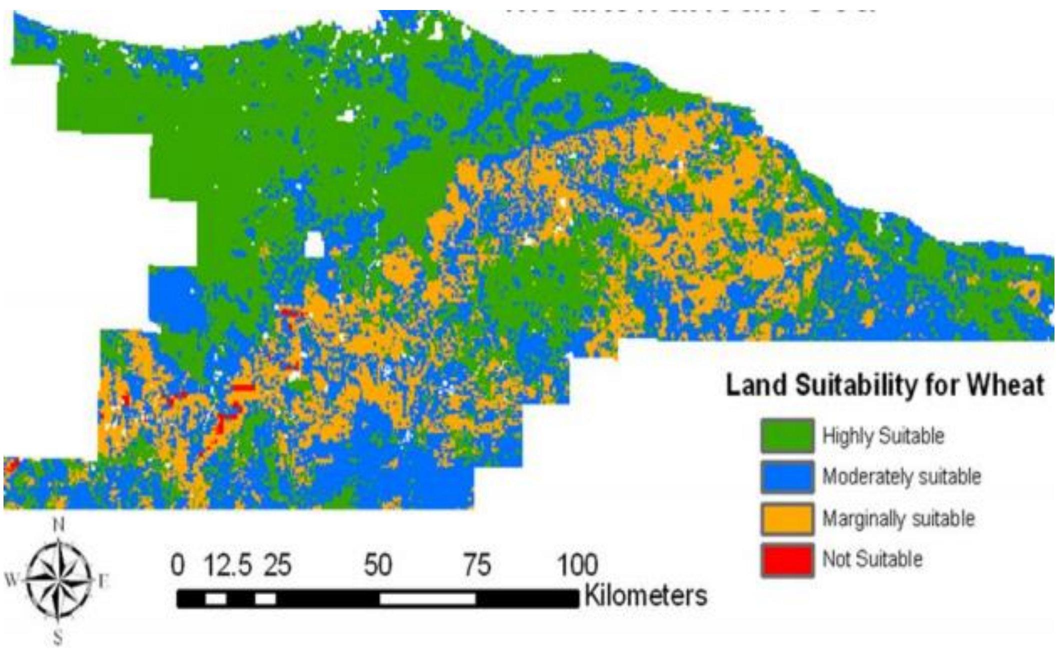

Figure 3.2.3: Land suitability map for wheat in arsi zone, Oromia region, Ethiopia (Kebede, 2014)

ArcGIS was used for preparing and handling special and non-spatial data as a source for case library construction so as to utilize in the process of case-based reasoning system development. The study to be executed for common cereal crop to identify whether a land unit is suitable for each crop based on the same cases.

The main purpose of the land suitability assessment guidelines and the row special data prepared by EIAR is that they would be useful in land use planning and strategic decision-making. The present case-based reasoning system builds on the existing special data to extend its capability to suggest suitable crops that can be grown in a farm unit. Unlike other crop advisory systems which are based on the physical land evaluation factors which only consider soil and physiographic factors, the present case-based reasoning considers additional factors for land evaluation and crop suitability and uses a non-spatial of socio-economic analysis based on FAO guideline and domain expert.

The main procedures for evaluating and classifying land are comprised:

- The study of relevant existing information and, wherever practicable, field appraisals of land conditions and experiences in a fully developed area having physical, and socio-economic conditions similar to the area under investigation.
- The selection of cropping and management alternatives and the description of prospective land utilization types (LUTs) for evaluation.
- The selection of types of data required for the evaluation and the preparation of all resource inventory.
- The selection of class-determining factors having significance from a physical and socio-economic standpoint, and the specification of critical limits to designate factor ratings and land suitability categories.
- The classification and mapping of provisionally evaluated land.
- Modification of the provisionally evaluated land classification based on additional pertinent physical, erosion and social-economic information by updated class determining factors and critical limits.
- The classification and mapping of the evaluated land delineating the location of the specific lands found to be suitable for crop cultivation.

The current research model is classified as spatial and non-spatial based on a dynamic resource base to provide dynamic land suitability. The model considers geographical location of a land unit for evaluation, the rainfall distribution occurring in a farm unit. The model evaluates land on a dynamic basis i.e., the suitability of a land unit for a particular land use varies with time. Whenever land is evaluated, current weather, socio-economic and government regulation factors will be taken into account. In order to build a successful decision, support system, it is essential to know how people make agriculture decisions in the real-world. Also, it is necessary to identify the specific decision criteria used by them in the whole process of decision-making. Only after a thorough understanding of the current decision-making process, will it be possible to suggest improvements. So the researcher followed the same procedure as it is mentioned above in order to develop the prototype case based system.

3.2.2. Suitability evaluation

This research extends the EIAR and FAO guideline approach to take the land evaluation and assessment for determining crop suitability to the farm level. It also overcomes some of the limitations of the land evaluation assessment made by EIAR which only consider the physical land suitability analysis. The current research assumes that a farmer is interested in knowing the best crop that can be grown in a farm unit and doesn't want to go through the individual evaluation processes for each crop. The best crop here refers to the one that is most profitable by economic standards and the one that fit within the strategic land use plan developed. This can be fulfilled by evaluating a farm unit for its suitability for cereal crops and then showing the crops suitable class status simultaneously to the farm unit. The developed prototype system evaluates a land unit for all possible cereal crop options at a single attempt.

Later, with inputs from the existing land attributes, the CBRLCCM system shows simultaneously the suitable (alternative) crop that can be grown in a land unit for the common cereal crops which cultivate in Ethiopia. The inputs from land owners include spatial and non-spatial data which include soil type (t Availability, Nutrient Retention, rooting Conditions, Soil Workability and Oxygen Soil Drainage class), topology (degrees of slop), The environmental(erosion), climate (mean temperature and rainfall), socio-economic (profitability, crop rotation, equipment and farmer skills) for the farm unit which is under consideration. The land owner has an option of including assessment factors such as government regulations, crop rotation, profitability and skills in the assessment process. The suitability evaluation process begins after the user enters necessary information on climate, soil and topography of the land so as to evaluate the cereal crops.

3.2.3. Soil form

Land suitability evaluation, on basis of soil conditions requires criterion mostly from the soil attributes. Which represents the main soil parameters used for generation of case based system which uses five attributes process for generating the final suitability for each crop. The important soil parameters are discussed here under. The framework developed by FAO comprises key terms and their description such as land characteristic, land quality, land utilization type and land use requirement. Land characteristic is a direct measure of simple attributes of the land that can be directly observed in a routine survey in EIAR. For example, soil texture or pH of the soil. Land quality is a complex attribute that can be derived by combining one or more land characteristics. It could be a qualitative or quantitative derivative of land characteristics. For example, a land quality named soil acidity is derived from pH measurements in the surface and deep soil. Land utilization type is a type of broad land use category. For example, cereal crops, dry land farming or intensive (irrigated) farming represent different land utilization types. Combinations of certain land qualities form the ability of the land to fulfill specific land use requirements. Land use requirements are the criteria required by certain land use (or a crop in particular) for successful implementation of a land utilization type. For example, the criteria (conditions) that are required by intensive farming may be different from dry land farming. Therefore, soil form has its sub criteria which are an important factor in crop cultivation. Texture is an important sub criterion which is provides significant information regarding water holding capacity, permeability, irrigation requirement and erodibility. Growth and development of the plant primarily based on the soil texture. Root penetration, nutrition absorption through soil particles, water holding capacity, water infiltration and percolation are affected by texture type. Soil depth for the cultivation is another important factor controlling cultivation methods as well as the selection of the crop type.

Moreover, Soil reaction is the degree of acidity or alkalinity of the soil and pH is the negative logarithm of the H ion activity. This refers to the relative activity of the H ion in the soil solution. In present investigation pH value ranges set for each kind of crops which presents an indication of the degree of availability of many soil nutrients and the favorability of soil condition to microbial activity which contributes to the fertility in turn. Water holding capacity (drainage) Water holding capacity is the amount of water taken by unit weight of dry soil when immersed in water. Water holding capacity gives an indication of the ability of the soil to provide moisture over a non-irrigated drought period. This capacity related to soil texture and soil organic matter. Sand possesses low WHC, while silt, clay and soil rich in organic matter have high values. Therefore, these the entire attribute of soil has an important factor so as to decide the overall soil suitability.

Soil suitability in the soil form consists of five main qualities: The overall suitability assessment process for soil form is presented in Fig. 3.11. Each quality in the form has its evaluation results. The final result is the overall suitability for soil, computed by the maximum limitation method for current suitability evaluation.

3.2.4. Slope form

Slope is considering an important factor in land suitability classification. It influences the irrigation system, irrigation efficiency, soil drainage, soil erosion, labor requirements and mechanization use. The slope layer was produced from the contour map of study area. Then slope maps of the study were derived using the “Spatial Analysis Slope” tool in ArcGIS 10.1 Slope using spatial analyst was calculated and a continuous elevation surface digital map, was grouped into categories. The data was further processed to yield important derivative products such as slope. Surface function of ArcGIS was applied to convert contour map to slope grid map. The four suitability ranges or slope suitability criteria classified for each cereal crops. The classified raster data layers were then converted to feature (vector) data layers for the overlaying analysis. Using data management tools in Arc Tool box, generalization of the feature (vector) data layers was performed to make a clearer slope suitability map.

3.2.5. Climate form

Temperature and rainfall are the main climatic factors that can affect land suitability. Climate change could affect agriculture in several ways such as the availability of water in rain fed agriculture areas land, degradation risks, amount of soil moisture and erosion. In addition, the amount of soil moisture stored in the soil profile can be taken into. The developed system considers the climate input as a numeric value to the assessment process of land suitability. Climate information in the system is described by two characteristics: The average rainfall (mm) and average temperature(c) per month.

The overall evaluation of the climate suitability to a specific crop will be shown in an evaluation box at the bottom of the climate form The system stores the selected input information and computes the suitability level of the average rainfall (mm) and the temperature per month. The final evaluation of climate suitability is calculated based on crop requirement.

3.2.6. Soil Erosion Hazard

Five major factors are used to calculate the soil loss for a given site. Each factor is the numerical estimate of a specific condition that affects the severity of soil erosion at a particular location. The erosion values reflected by these factors can vary considerably due to varying weather conditions. Therefore, the values obtained from the universal soil loss equation more accurately represent long-term averages (USLE,2006).

$$A \text{ (ton/ha/year)} = R * K * L * S * C * P \quad \text{--- Eq. 1}$$

Where A is the mean annual soil loss, R is the rainfall erodibility factor, K is the soil erodibility factor, L is the slope length factor, S is the slope steepness factor, C is the crop management factor and P is the erosion control practice or land management factor. A tolerable soil loss is the maximum annual amount of soil, which can be removed before the long-term natural soil productivity is adversely affected. The impact of erosion on a given soil type, and hence the tolerance level, varies, depending on the type and depth of soil. Generally, soils with deep, uniform, stone-free topsoil materials and/or not previously eroded have been assumed to have a higher tolerance limit than soils that are shallow or previously eroded.

Soil loss tolerance rates which have been classified into four class based on its rates included in table: 3.8 below.

Table 3. 2.6.1: soil loss tolerance rates class

Soil Erosion Class	Potential Soil Loss tones/hectare/year (tons/acre/year)
Very low (tolerable)	<6.7
Low	6.7 (3)–11.2
Moderate	11.2 (5)–22.4 (10)
High	>22.4

3.2.7. Socio-economic

The FAO framework is simple and broadly aids at developing a land evaluation model. Hence, it has been adapted in the current research that economic, social and political factors which influence the crop selection process. Social and cultural characteristics influence cultivation pattern land evaluation is effected by social attributes. Importance of social parameter for suitability classification should be raised from the opinion of individual and corresponding weight can be given for analysis purpose. These parameters are then classified under socio-economic class as sub criteria for evaluation purpose. The weight of criterion and sub criteria give with the discussion of expert and reviewing FAO framework for socio-economic factor. Hence, it is also possible to provide remedial measures or explanations in order to switch to a second choice in case the user is not willing to settle for the first crop suggested by the system.

The social, economic and political factors, (for more details refer to table 3.10), are generally non-spatial. Economic factors such as market price and market demand for a crop produce is important factor so as to determine profitability (revenue) of a specific farm unit for a crop produce. Market demand is crucial in determining the supply- demand economics of a particular crop. By accurately measuring the market demand or predicting seasonal demand for a crop product, its optimal selling price can be determined so as to determine the profitability. If the market demand is high, the price for that crop will be high and vice-versa. The market demands for the crops were randomly assumed at present as they could not be obtained from real market places. However, seasonal demands of various crops and their variations can be estimated by studying the market over a period of time.

Therefore, profitability (revenue/cost) ratio is one of the important factors which draw most attention of the commercial farmers. In present research profitability parameter influencing development of the crop production Ethiopia. Larger the cultivation parcel, higher would be the revenue/cost ratio. For instant land fragmentation is in continue state, this causes higher intensive labor input, this situation certainly causes to recede revenue/cost ration and decrease net profit. By accurately measuring the profitability or predicting based seasonal production of a specific crop, its result of revenue/cost ratio must be more than expenditure cost. Machinery availability and political factors are also represented in the expert system's knowledge bases Political factors such as government regulations or restrictions, subsidies on certain crops are essential to choose a crop and therefore are a part of the CBRLCCM. The political factor can be considered as a heuristic factor which depends on the human judgment of an expert who will input this information into the case-base system. These types of information are highly dynamic and need to be periodically updated by the expert. Much of the crop knowledge base data used in the Crop Advisor, especially the economic, and social data, have been collected from a crop expert through interviews and FAO Guideline.

Farmer skills

Farmers' skills are an important factor in land evaluation due to crop cultivation need skills, technique and motivation so as to well cultivating crops.

Profitability (revenue/cost) ratio

The cost from the crop yield must exceed the revenue of cultivating the crop, if not it would be a loss for a cultivator. The price of crop produce at markets is an essential element that drives the choice of crop. Also, the seasonal demand for the particular produce should be considered before selecting a crop. The selection of crops, which prices of the crop produces are highly fluctuating in the market place, tend to be risky.

Crop Rotation

The previous year's crop affects the productivity of a farm unit. Growing same crop year-after-year not only robs the productivity of the soil, but also it does not break the pests' life cycle. Crop rotation is successful alteration of crops preferable from two dissimilar families of crops.

Equipment

Some crops can be seeded, harvested with the help of farm machineries but others don't permit. On the other hand, with the help of machineries, large fields of wheat or barley can be managed with ease.

Government policies and restrictions

The government indirectly controls agricultural crops grown in a certain region by providing favorable subsidies, offering fare prices and establishing procurement prices for certain crops. On the other hand, it also restricts the growth of certain crops by the method of licensing, water restrictions and imposing zoning regulations.

3.3. Case Representation

After knowledge is acquired and modeled the next step is representing the knowledge with the appropriate case representation technique. CBR is a type of knowledge representation which uses previous experiences in form of cases to understand and solve new problems. For the prototype CBR for matching land use with crop requirements of system knowledge is represented in the form of cases. According to Gebhardt (1997), cases in many practical CBR applications are usually represented as two unstructured sets of attribute value pairs; as the problem and solution features. As it is easy to represent the cases (knowledge) collected using attributes with their characteristic values; the researcher selected feature-value pair format representation for the cases in the study. For an attribute-value representation it is relevant to find out which attributes with which semantics relevant to be chosen. So for representing the knowledge with feature-value pair format, case structure was constructed for matching land use with cereal crop requirements. The case structure has two important parts: problem descriptions (land quality) and solution (crop suitability class). Problem description, as part of the case structure consisted of attributes (characteristic and quality of land) which described the problems to be solved. The solution part provides the recommended suitable evaluation class provide to cereal crops based on the land descriptions. The Case Structure for land evaluation for cereal crops is descried.

Table 3.3.1: The Case Structure for land evaluation for cereal crops

Attribute name	Parameter
Slope	Description
Soil erosion	Description

Average temperature	Description
Rainfall	Description
Drainage	Description
Soil depth	Description
PH value	Description
Organic matter	Description
Texture	Description
Farmer skills	Description
Equipment	Description
Crop rotation	Description
Profitability	Description
Government policy	Description
Barley	Solution
Maize	Solution
Wheat	Solution
Sorghum	Solution
Tef	Solution

Therefore, the researcher identified the different problem description attributes and solution with the help of domain experts, FAO guidelines, manuals and different materials. The attributes which were used for land evaluation for cereal crops were: slope, erosion, climate, drainage, soil depth, PH value, organic matter, mean temperature, rainfall, farmer skills, profitability, crop rotation, equipment and farmer skills and recommended crop suitability. Most attributes selected have string and integer data types and some others have double and Boolean value. There were different challenges during identification and representation of case structure. There were different challenges during identification and representation of case structure. There were too many attributes that was found on land evaluation during the pre-selection stage of attributes. So that, identifying the most significant attribute was one of the challenges that took long time to come up with the final list of important attributes. So using the selected attributes, the researcher collected 298 sample spatial data for land evaluation from sample survey made by the Ethiopian institute of agricultural research center which only considers the physical land suitability. Then this sample data inserted into ArcGIS 10 with inclusion of socio-economic layer which are formed totally five layer into ArcGIS tools. As a result, 298 fully working cases with their alternative land suitability class solutions that comprise both physical land suitability as well as socio economic factor which is represented using feature-value format has been prepared and inserted in to the case library. For this study the researcher selected 161 cases among a total of 298 cases. Physical Land area with suitability level highly, marginally and moderate are considered for incorporation component of these study for further analysis. Which are Non suitable land for cereal crops and the repeated cases omitted from suitability evaluation.

Therefore, the researcher selected 161 correctly filled cases which are generated in ArcGIS successfully and totally 19 attributes selected to develop the prototype system.

CHAPTER FOUR

DESIGNING AND IMPLEMENTATION OF THE PROTOTYPE

The design and implementation part of this section involves the actual development of

a workable CBR system for land evaluation process in the Ethiopia institute of agricultural research. Relevant cases are collected from domain experts and FAO guideline. Therefore, having all the necessary cases and the knowledge from the domain expert and different relevant documents, the next task is coding the knowledge into computer using appropriate and efficient knowledge representation methods. For this research, jCOLIBR 1.1 CBR frame work is used to develop the prototype. The retrieval algorithm used in this research is nearest neighbor retrieval algorithm. This is because jCOLIBRI uses this algorithm for retrieval task.

4.1. Designing the Architecture of Case Based Recommender System

Figure 4.1.1 shows the framework of CBR system developed in this study for identifying suitable land for cereal crops. To develop CBR system, the researcher collects the important knowledge from relevant documents and domain experts, which are land suitability assessment for cereal crop production involves the interpretation of data relating to soils, climate, topography and socio-economic into a suitable format to allow land suitability analysis to take place. Land qualities and their associated land characteristics are arranged in five categories. The process of assessing the land suitability is based on matching land characteristics with crop requirements to produce thematic map layers for each category in ArcGIS which are used as case library for CBS.

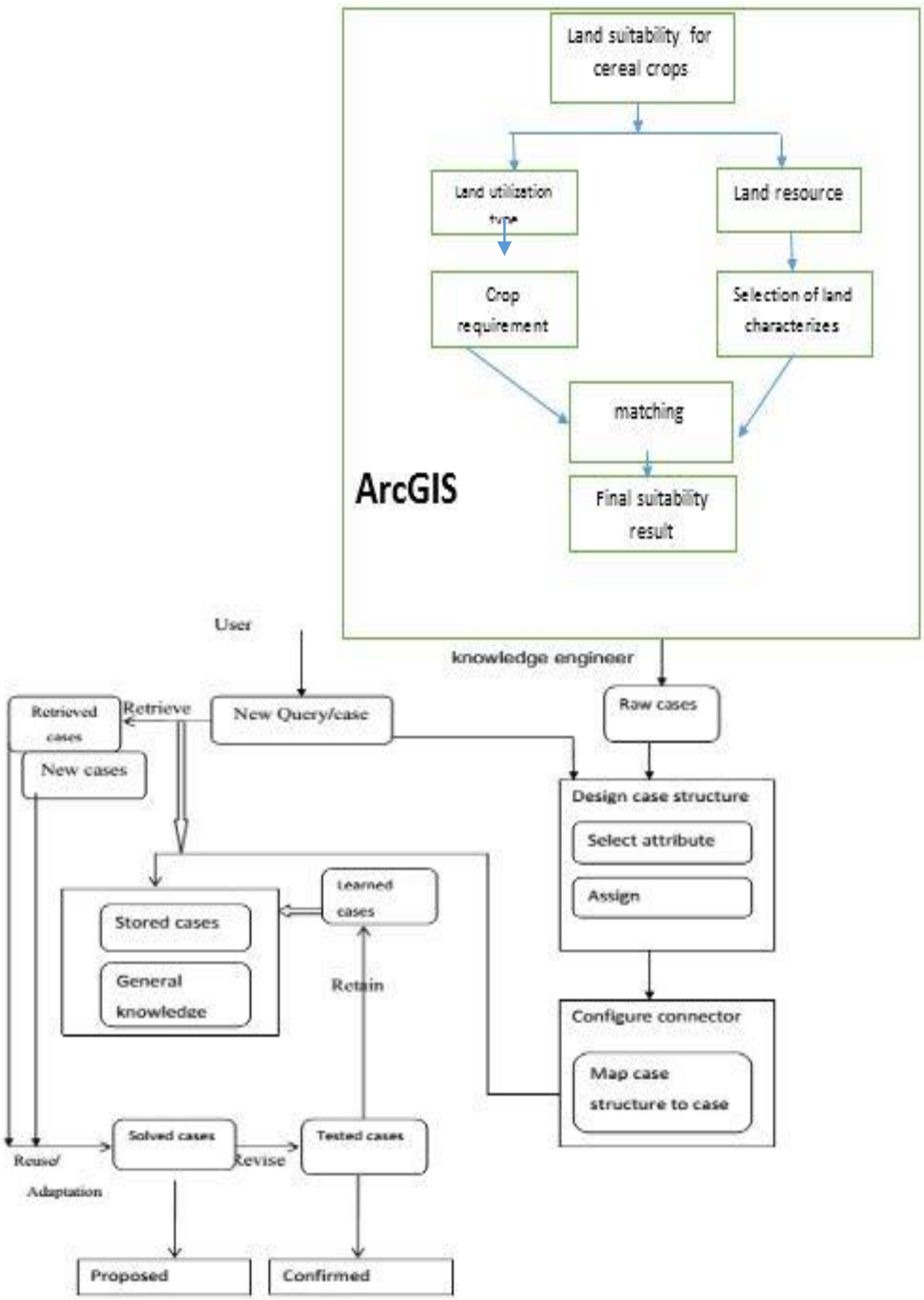

Figure 4.4.1: architecture of Case based reasoning system

To design case-based system, the researcher collects the relevant attributes from domain experts and FAO guideline which are generated as a tabular form in ArcGIS so as to represent the required knowledge. After processing of cases and having the selected attributes, assigning weight and important parameters for each attribute was the next performed step. This is because since all attributes are not equally important to land evaluation of crop. Once the CBR system is developed, farmer can get help easily through expert or any agricultural assistant who assign for farmer so as to use the system to choice their suitability for farm unit under cultivation given by interacting with the system in order to retrieve the best cases that can match with their query. When the user enters their query/case description through the user interface window, the system searches the best matching cases from the case base and retains the possible solution. If there is exact matching between the query and cases in the case base, the system recommends the most matched cereal crops for the farmer based on the given query. If the similarity between query and case is approximate, the proposed solution needs modification (adoption of solution) to fit the current problem described. At the end, the best modified solution should be stored into the case base for future use. The case base updates incrementally when the system learns from new case used by the user.

4.2. Case Based Reasoning System for land evaluation

To develop CBR system, JCOLIBRI case based reasoning framework is used. JCOLIBRI has been constructed as core modules to offer the basic functionality for developing CBR system. Developing a new case base recommender system is made by writing few Java classes that extend classes of the framework and configure some XML files. To start the JCOLIBRI graphical user interface (GUI) application tool, launch the main window by clicking on **JColibriGUI.bat** file and it becomes ready to use as shown below in figure 4.2.1

Figure 4.2.1: JCOLIBRI main windows

In this study, the development of CBR system for land evaluation for cereal crops process can be divided in the following sub processes which enable to achieve the objectives of the research.

4.2.1. Building the Case Base

As stated in the objective of this study, one of specific objectives was building case base reasoning to provide the most important similar cases to support farmer in the process of matching crop requirements with land use for their farm unit. As a result of this, the researcher collected the case from Ethiopia institute of agricultural research center which is process in ArcGIS base on the cereal crop requirement by the researcher. The acquired cases are used to build CBR system for land evaluation that is important to assist farmer in the process of matching land use with crop requirements. Cases are stored in the case base as a text file in structured format after the researcher has analyzed and interpreted the case. The case base is presented as a plain text comprising of N columns representing case attributes (A1, A2, A3, …, AN) and each M rows representing individual cases C ({C1, C2, C3, …, CM}) each attribute has a sequence of possible values associated to each column attribute A= {V1, V2, V3, …, Vk}. The case base consists of a set of cases that represents knowledge about land evaluation process. The researcher tried to collect all the cases from Ethiopia institute of agricultural research purposefully.

4.2.2. Case Representation

The case representation has been formulated in the way that easily represented in JCOLIBRI. Designing of such case structure helps to define the features available in the cases and used to measure the similarity between existing cases and the new case (query). The general application of this research is to retrieve similar cases to the query from the case base that can

Table 4.2.3.2: descriptions and weight of the selected attributes

guide farmer, solving problems of confusion and transforming a recommendation in the process of matching land use with crop requirements. Case base were structured to make the retrieval process efficient. This is done through case indexing process in the JCOLIBRI programming tool. Indexing refers to assigning index to the case for retrieval by comparing the existing case and the query given by the user.

4.2.3. Description and Weight of the Selected Attributes

The case in this research consists of thirteen descriptions/attributes that served to contain descriptions of the problem which is used to make decision by the system and two solutions attributes which holds solution for the recommendations. The following table shows the description of selected cases with their value and Weight.

Significant attributes			
Attribute Name	Data Type	Weight	Local Similarity
slope	Integer	1	Threshold
Soil erosions	string	1	Max string
Drainage	String	0.7	Max string
Soil depth	Integer	0.8	Threshold
PH value	Integer	0.9	Threshold
Organic matter	Double	0.8	Threshold
Texture	string	0.7	Max string
Average temperature	integer	0.9	Threshold
Average Rainfall	integer	1	Threshold
Farmer skills	String	0.5	Max string
Equipment	String	0.5	Max String
Profitability	String	1.0	Max string
Government policies	String	0.5	Max string
Solution			
Barley	String	1.0	Equal
Maize	String	1.0	Equal
Wheat	String	1.0	Equal
Sorghum	Integer	1.0	Equal
Tef	String	1.0	Equal

The above table shows the general description of attributes consisting of attribute name, data type, weight and local similarity. The most significant attributes to the problem domain such as slope; erosion, average Rainfall have the highest weight value of 1.0. These attributes are the most relevant in the process of matching land use with cereal crops requirements. Next to these, attributes like PH value, drainage, texture, temperature, soil depth and crop rotation has second most important factor which have great significance in the process of matching land use cereal crop requirements. The assignments of weights to each attribute indicates that attributes having high weight is the most relevant to the user in the process of land evaluation for cereal crops. The weight of each attribute has been assigned its value by domain experts at the time of attribute selection.

The local similarity of most attributes is maximum string. This is due to the similarity between query and cases can be calculated with maximum string length. Few attributes such as PH value have equal similarity weight since it needs exact match of cases and new query. After identifying relevant attributes of the case, the next task is definition of appropriate similarity measure in JCOLIBRI. JCOLIBRI follows both local and global similarity measures. Local similarity measure divides the similarity definition into a set of local similarity of each attribute whereas global similarity calculates the final similarity measure. Different types of local and global similarities are used in this research.

Local similarities include the following

- ✓ **Equal: -** The input query and cases in the case base must match to get the result; if there is no match between input query and cases, matching will fail.
- ✓ **Interval:** - Exact match is not required in this similarity. When it is assigned to the attributes, JCOLIBRI reminds this interval value in the searching from the case to get the similarity.

Global similarity

- ✓ **Average: -** it is a type of global similarity that considers the average of all attribute similarity values.

4.2.3. Managing Case structure

A case is composed of three components: description (describes the problem), solution (represents a possible solution approach) and result (reveals if the proposed solution is able to solve the problem). Description and solution are collections of simple or compound attributes, permitting us to build a hierarchical case structure.

One of the most important features of JCOLIBRI is managing or defining case structure of the system easily by case structure window. The selected attributes were added to the description window of case structure and their properties are assigned for each attribute to the right side of the window. After this, the program generates a java code automatically and saves in XML format. As shown in the following figure 4.2.3.1, data type, weight and local similarity of the selected attribute can be configured on the right side of the case structure window. The window is divided into two parts. The left side contains the structure of the attributes as a tree. We can generate description as well as solution cases by clicking “add simple” button and remove unwanted attributes by clicking on “remove” button. The right side contains the property values of the attribute which can be configured as the property of cases.

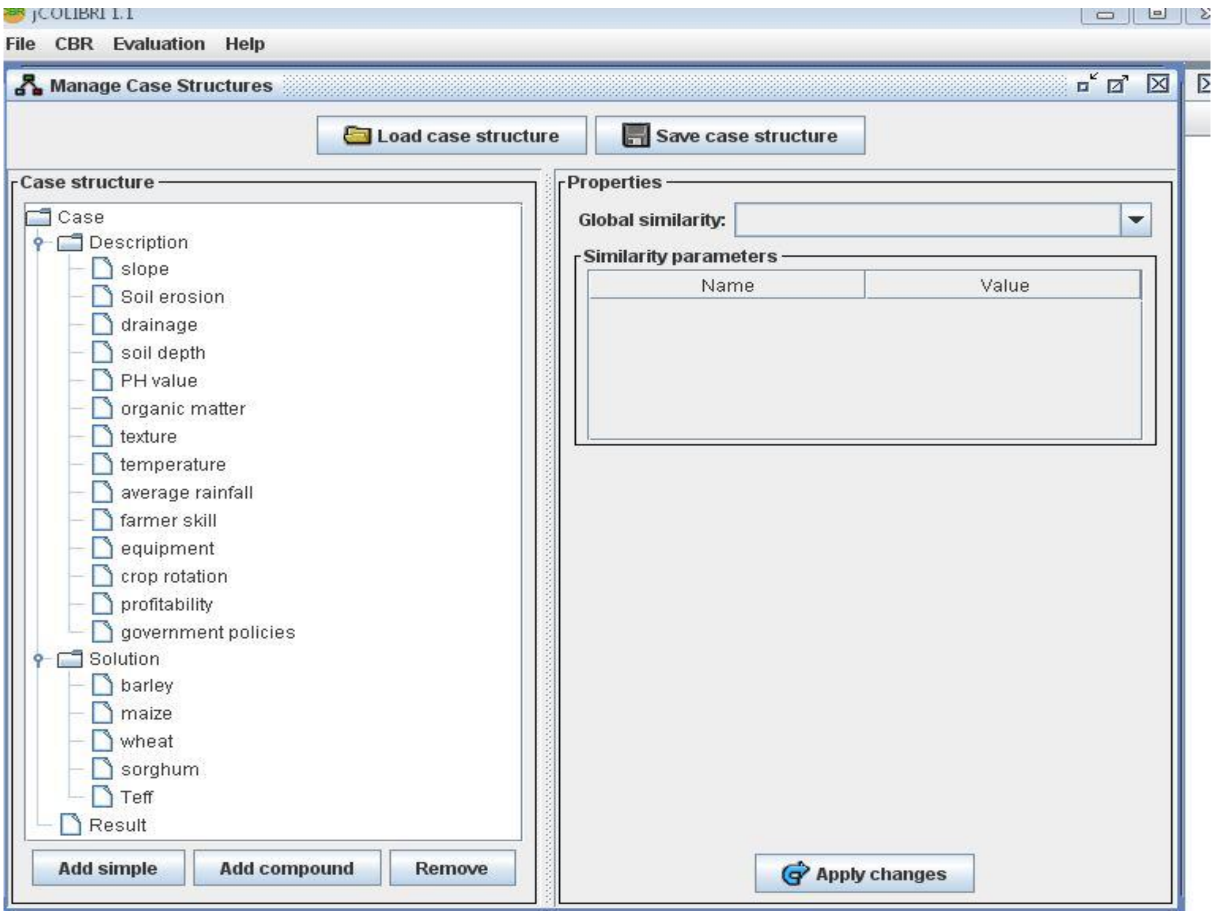

Figure 5.1: Managing Case Structure

4.2.4. Managing Connectors

Once case structures are configured in JCOLIBRI, the case based reasoning systems must access the stored cases from case base. JCOLIBRI supports both SQL database and plain text file to store its cases base.

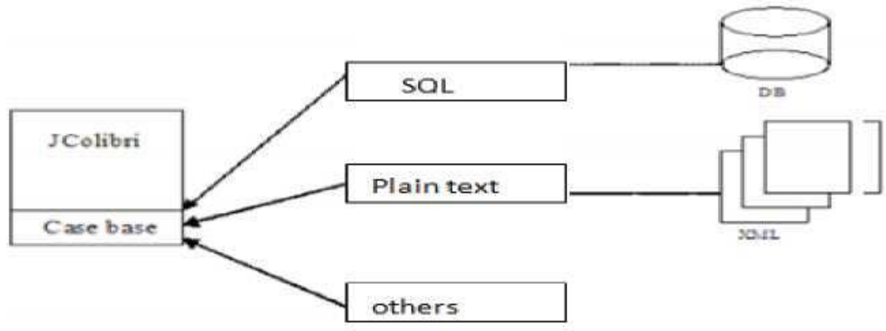

Figure 4.2.4.1: JCOLIBRI case base schemas

In this research plain text connector is used as a case base storage. The connector maps the case structure to its column from plain text file which is saved in **.txt** file format and later saved as XML file like that of case structure.

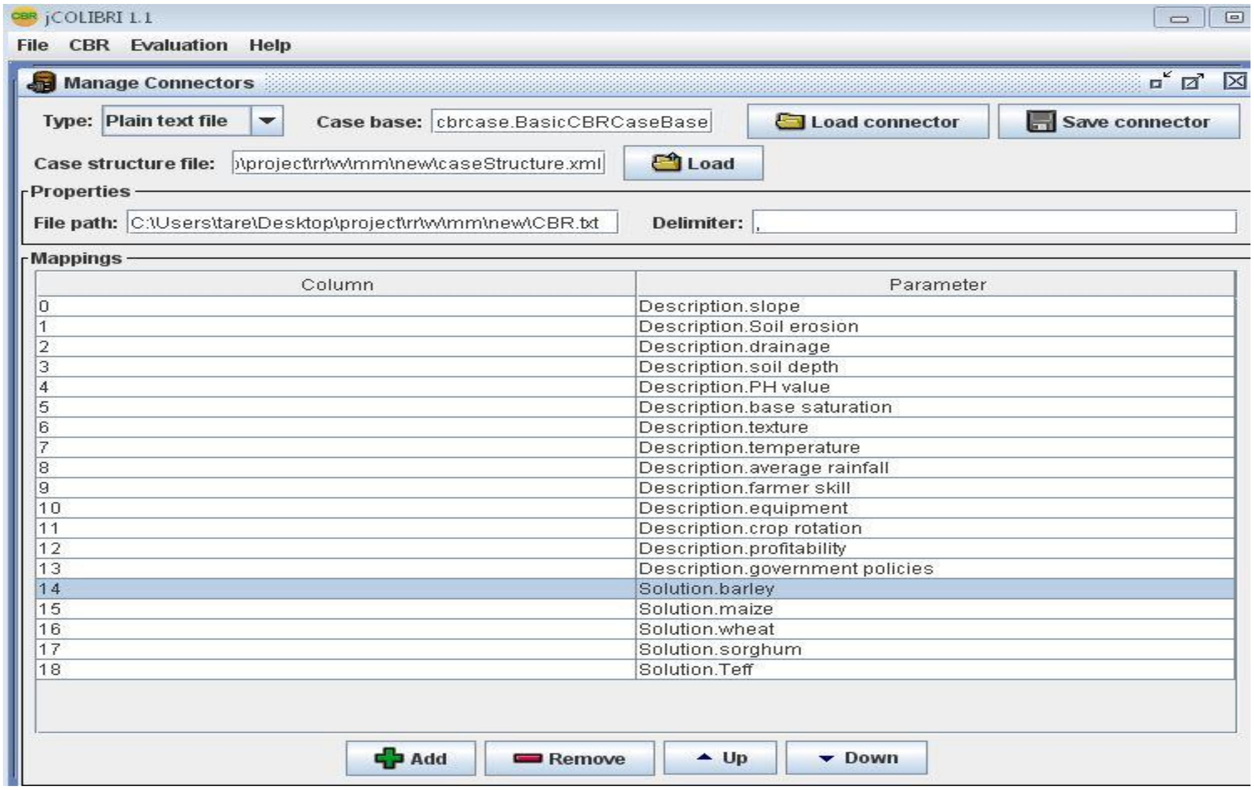

Figure 4.2.4.2: Managing Connectors

One of the most important tasks in managing connectors is specifying the correct path of case structure and file path. The case structure path is used to access and match attributes from case structure and file path is used to specify the **.txt** file that contains the case base. Delimiter of this connector uses comma (,) to separate value of each attribute in the case.

4.2.5. Managing Tasks/ Methods

JCOLIBRI is organized into packages. These packages can perform and execute tasks and methods decomposition process. For the development of case base recommender system prototype, the researcher used core package task. These core packages are Pre Cycle, main CBR cycle and Post Cycle. The detail of each tasks and methods can be discussed separately as follows.

4.2.5.1. Managing Tasks

After configuring the connector and case structure, the next task is selecting tasks and methods of application. jCOLIBRI has two types of task packages, namely, Core packages and User defined package tasks. For the development of CBR land crop matching prototype, the researcher used core package tasks. A core package contains all classes that represent core functionality of a CBR application such as the domain model, case bases, similarity functions and retrieval algorithms. Core packages also have predefined tasks and methods that used to configure new system by reusing the tasks rather than using tasks or methods defined by the system developer itself like user defined packages, because defined tasks and methods by user itself for every system is time taking and complex. Different core packages are available, which are the most important packets in JCOLIBRI. are PreCycle, main CBR cycle and PostCycle. The component of core packages is the final and important step for creating a new application where the CBR application is configured. The left side of Figure 4.2.5.1.1 shows PreCycle, CBR Cycle and PostCycle. The main components of core packets are discussed as follows

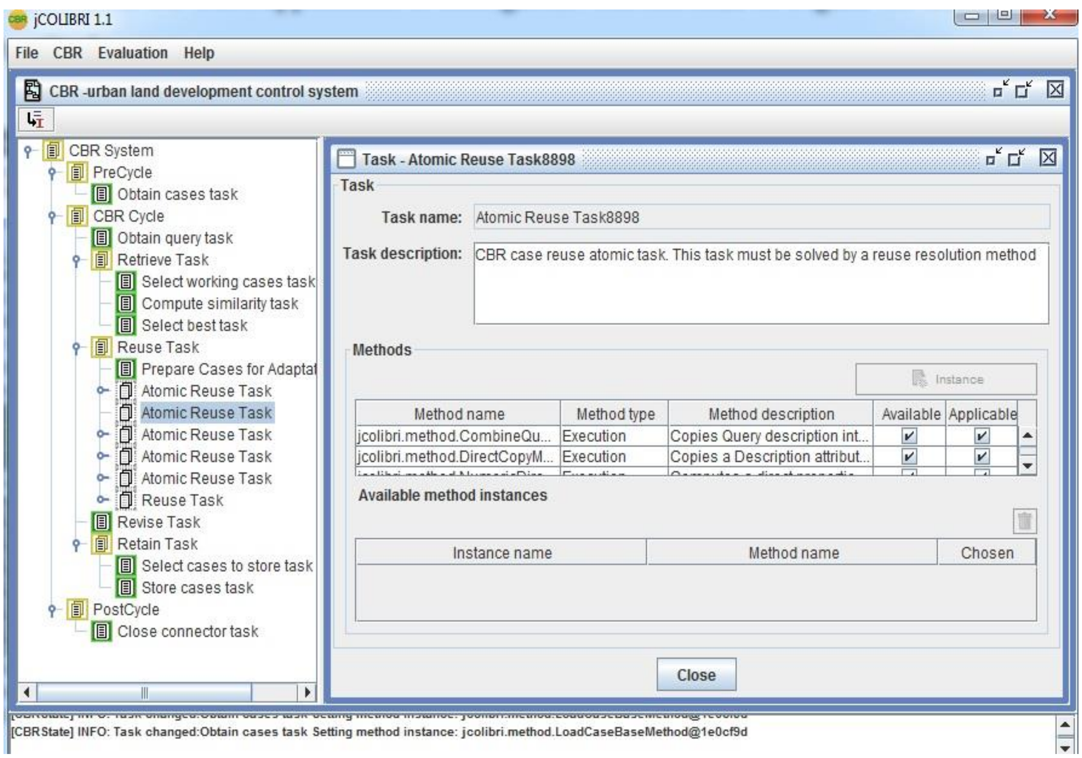

Figure 4.2.5.6 .1: Configure the CBR Application

- **PreCycle: -** This part of the task retrieves data or cases from case base before execution of the main cycles.

- **Main CBR cycle: -** it retrieves the most similar tasks and describes the typical cycle task at the highest level and obtains the query. Even if it is case based reasoning cycle part, it is also used to retrieve the most similar task for CBR land crop matching system from stored cases; reuse previously stored cases as knowledge to solve the problem; revise the proposed solution and lastly retain the experience. In this cycle, there are other subtasks and most of the tasks requires specific path of the case structure. The case retrieval phase of CBR cycle involves finding similar cases with the query using selected algorithm. After searching and selecting most similar cases with the new problem, the solution of problems will be displayed to the user. This task has sub tasks again, select working cases task which selects working cases from case base and stores into current case base; compute similarity tasks which is used to compute the similarity of cases between existing cases and new query; select best case which returns best cases from case base with high degree of similarity with new case.

Case Similarity, Matching and Ranking The main goal of case base reasoning land cereal crop matching system is to retrieve the best similar cases to the query from case base and selecting the nearest similar case. Selecting the best similar case is usually performed by means of some evaluation heuristic functions or distances, which are possibly domain dependent. JCOLIBRI uses the nearest neighbor algorithm as case retrieve one technique. Nearest neighbor algorithm used to measure the similarity between the existing cases and the new case/queries, and then return the search results in their ranked order. The local similarity function measures the similarity between each and every simple attribute values in the case base with new cases queries. The similarity score will be assigned based on the matching weighted sum features from those simple attributes.

The average score of each attribute between existing case and query are computed and the similarity between stored case and the query result is assigned to the object. Finally, the maximum degree of similarity among retrieved cases is displayed in their ranked order.

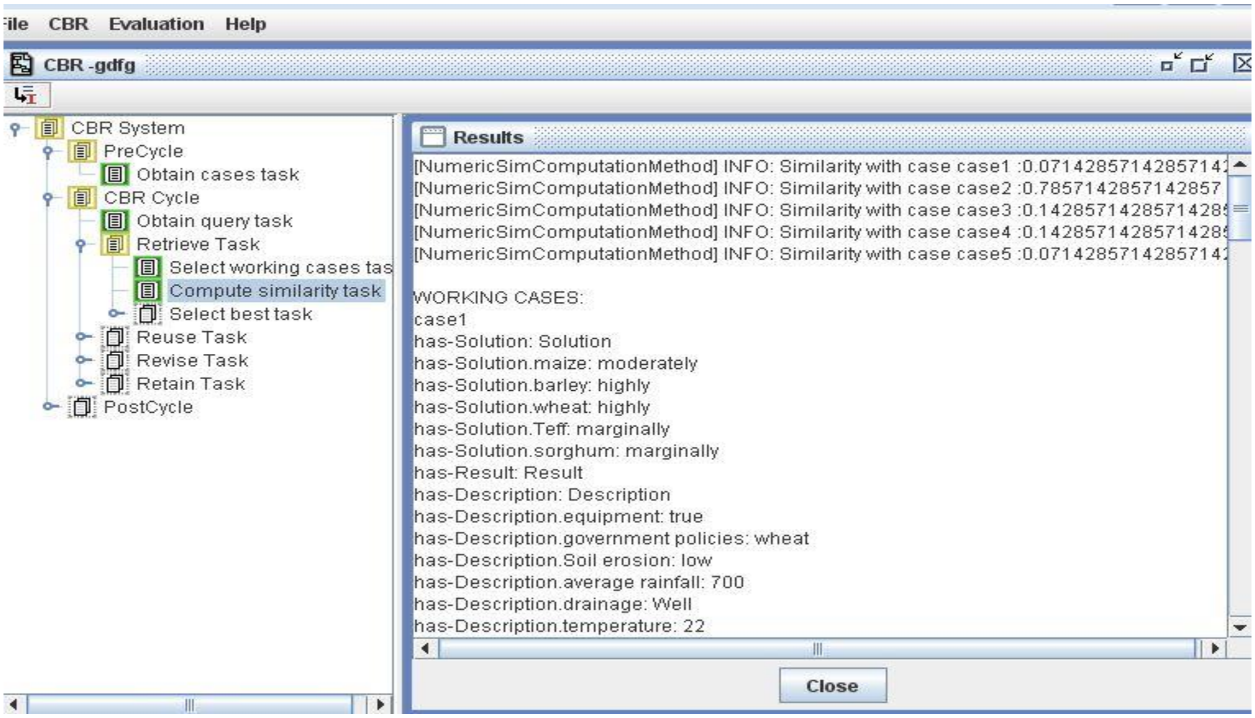

Figure 4.2.5.1.2: case similarities between case base and query

Reuse/ Adaptation: During case retrieval, once one or more very similar case is identified, the solutions are selected for this particular problem to meet the requirement of new solution. This reuse stage generates the proposed solution for the problem. There are situations where cases are not similar with the new case. At this time, this new case can be stored in the case base and will be reused by other farmer for the next time. The system can learn at every entry of new case and new users adopt this knowledge for land cereal crop matching process.

Revise task: This stage is the evaluation stage about the selected solution in the reuse phase. After selecting the most similar cases from search result, the solution for problem should be confirmed and validated before the solution is stored as a case for future use.

Retain task: after having confirmation in revision phase, the problem together with its solution will be stored in case base. These tasks are after validation and confirmation of retrieved case solutions at reuse and revise phases.

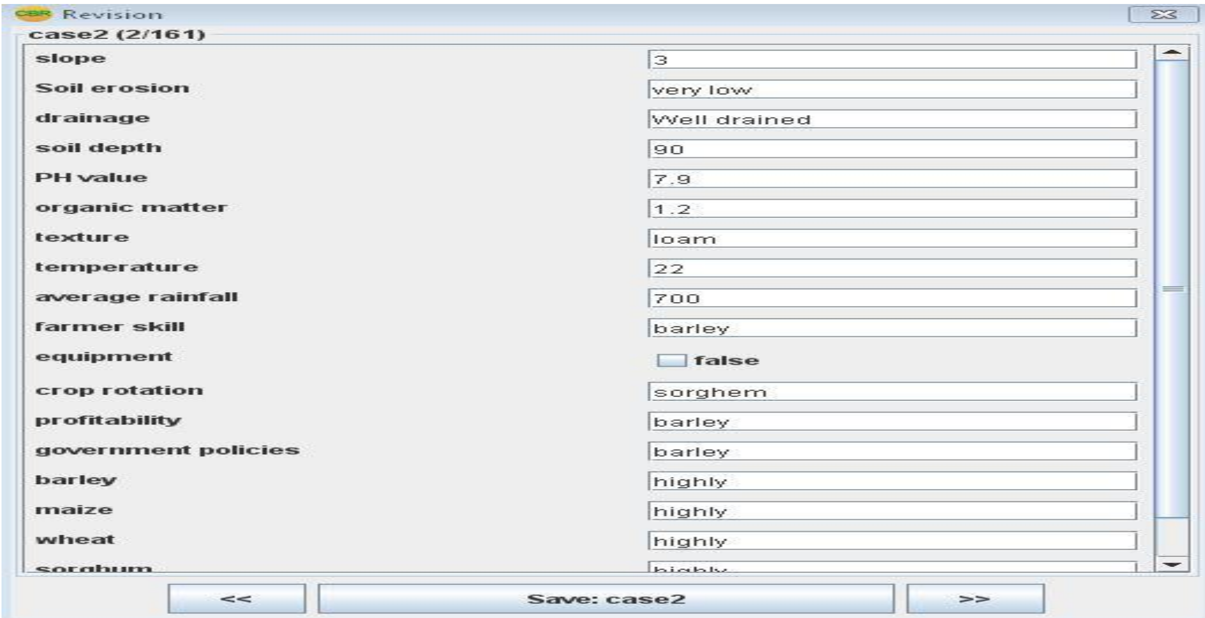

Figure 4.2.5.1.3: revise tasks

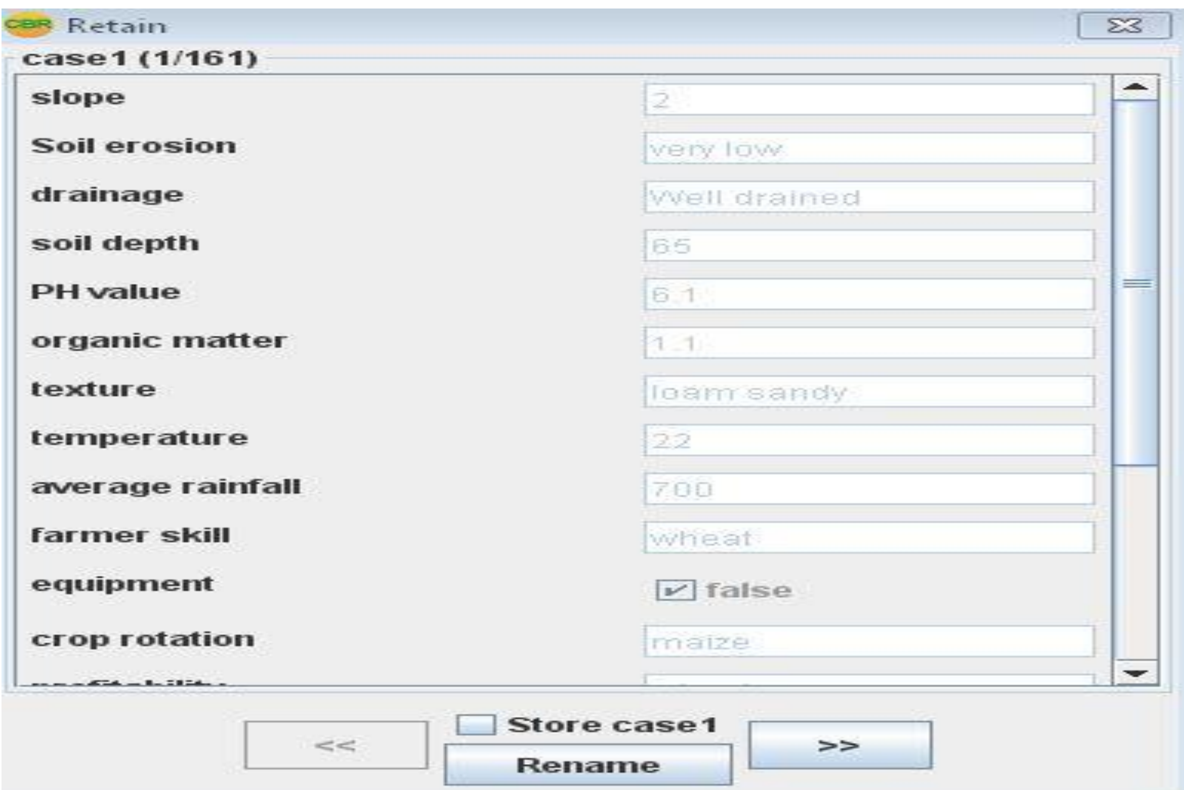

Figure 4.2.5.1.4: retain tasks

Post Cycle: - this cycle contains tasks that executes after main CBR cycle. Case retrieve begins with case description and ends with best matching of the case with the new case. The subtask of postCycle task called close connector task will close the connector and save the case base.

4.2.5.2. Managing Methods

The method packages store classes that resolve the task. These classes can resolve the CBR cycle using programming or using graphical user interface (GUI). All tasks in JCOLIBRI should have their own methods to be assigned in order to achieve its recommendation goal.

LoadCaseBaseMethod: This method returns the whole available cases from the case base to designer and use connector as parameter to retrieve case base.

ConfigureQueryMethod: This method obtains and configures the query. It displays the graphical user interface window by receiving case structures as an input to request query and to receive cases from the case base.

SelectAllCaseMethod: Selects working cases from case base and store them into current context. It allows displaying all the available cases from the case base to the result window. **NumericSumComputationalMethod:** Computes similarity tasks between cases and query. **ManualRevisonMethod**: Manual revision method enables users to modify cases in the query window.

SelectSomeMethod: select best of found cases. It returns most similar value of the top best selected case. Tasks in JCOLIBRI can be solved with different methods as listed above. Choosing the most appropriate method for the task is the role of researcher in the designing of case base land cereal crop matching system. For this research, only few of them are selected and discussed which are appropriate for recommendation system. Figure 4.2.5.2.1 shows the main window of cycle of JCOLIBRI tasks and methods. As shown below, pre cycle, main CBR cycle and post cycle are on the left side of the window. When the designer selects any task from these cycles, the configuration method windows displayed on the right side and appropriate inputs can be selected according to the situation. These inputs are parameters for new instances.

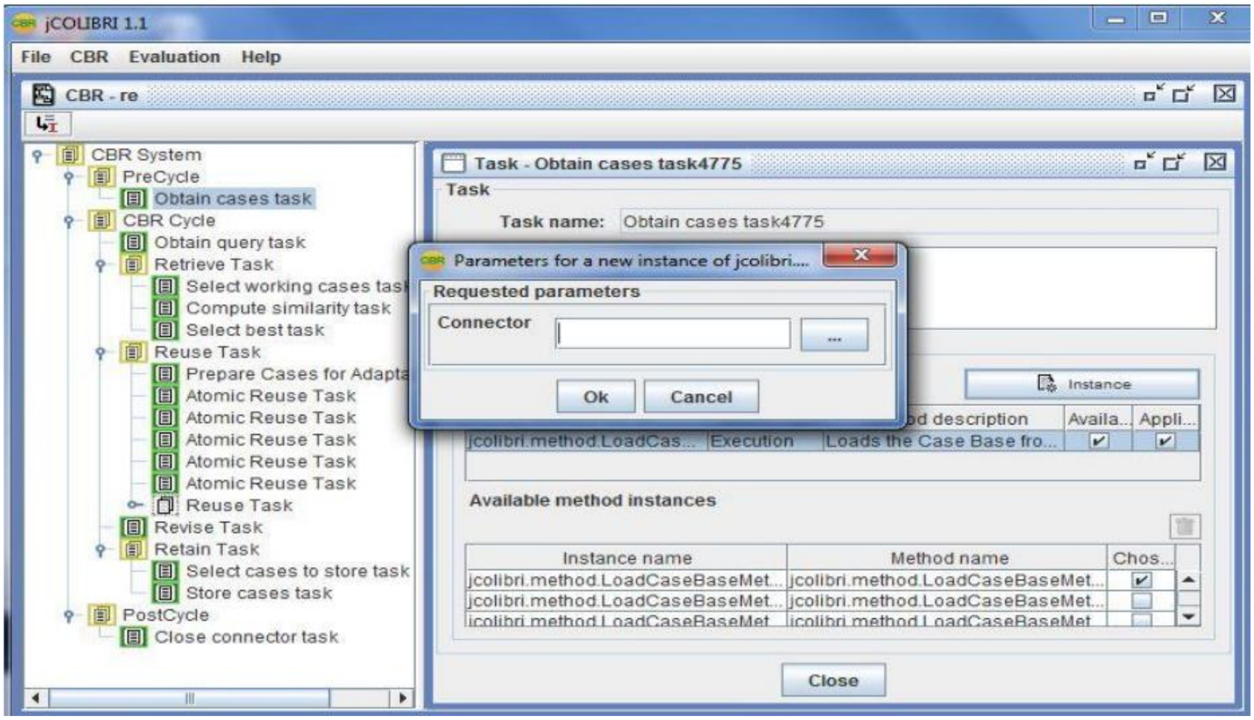

Figure 4.2.5.2.1: Managing Tasks/Methods

4.2.6. Deploy the case base reasoning system

After defining and configuring all the necessary steps required designing case base land cereal crop matching system in JCOLIBRI, testing the system application is the next step as shown in figure 4.2.6.1.

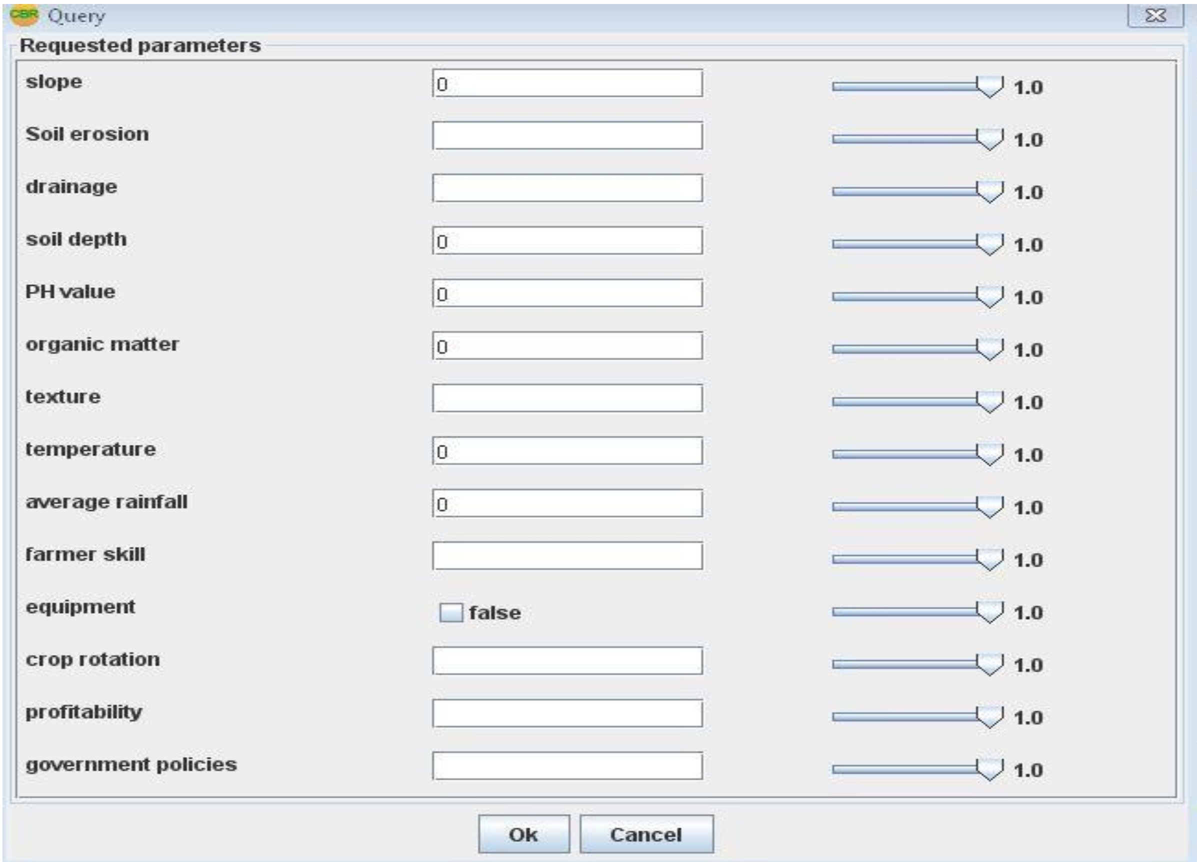

Figure 4.2.6.1: Case entry windows

Expert who provide crop advise service for the farmer are required to enter the query on the in the query request parameter according to the instruction. The service providers are required to enter existing or current land condition so as to get up-to-date information. The parameters which the system automatically fills by the number zero indicate that the filed could only accept no other values but only integers. After entering the query, at the bottom of the screen they will see the results of similar previous cases and the recommend land suitability for barley, maize, wheat, sorghum and tef.

Crop rotation is a process in which different crops are cultivated on a rotational basis to refurbish soil nutrients and to break the pest life cycle. Advice would be given to avoid the previous crop grown on the farm unit. This information is obtaining through asking the farmers(users) about previous crops which are cultivated on the farm unit, such as if the user have already grown the crop wheat in the previous season the advisor or expert would recommend for the currently farm unit to cultivate maize or sorghum crops. Hence, the user has an option to compare their choice (previous or current grown) with the choices suggested by the CBR land cereal crop system. The other factor machines available If the user provides “true” to machinery availability, the crops which do not require machinery would be eliminated by the system and the same for government police. The factors for farmer skills, machinery availability and government support are available for assessment by the system and their use by the farmer is purely optional. The user of the land cereal crop matching system can include or exclude any combination of available factors and the resulting crop choices can be observed on beside of the entered query.

The important point to stress is that the user actually decides which factors need to be taken into consideration by the assessment procedure and chooses options for each factor. When all the choices are made, the user can include Market index or current market price and market demand for a particular crop to so as to calculate profitability then take profitability into account. After evaluate overall factors then final suitable crop result appear under each cereal Crops.

CHAPTER FIVE

PERFORMANCE EVALUATION

Construction of case base is the first and most important step in developing case base reasoning system for land cereal crop matching application. After selecting fourteen important attributes, constructing cases is the most challenging task in developing case base reasoning system land cereal crop matching. As discussed in the chapter four, these attributes are selected by consulting with land evaluation and reviewing different document. The researcher tried to organize interview questions to get appropriate attributes which can help the information provider to farmer so as to make the right decision. As replied by domain experts and reviewed different documents, slope, erosion, drainage, soil depth, PH value, organic matter, texture, farmer skills, equipment, crop rotation, profitability and government are the most important criteria to be considered for matching land cereal crop. After selecting these important attributes, the next task was constructing the case base. These cases are organized based on important attributes. The researcher organized and processed these cases in ArcGIS software based on the crop requirement. The cases consist of various degree of complexity in terms of size of cases and attributes, the amount of parameters, weight and type of constructs. For this research a total of 161 cases were used to build the case base and to test the prototype. Therefore, these cases are important for user in decision making in the process land suitability analysis for farm unit under cultivation.

5.1. Case Similarity Testing

An experiment is made to evaluate how the system match a query/new cases with the cases from the case base. For this experiment, the researcher uses three experimental groups. These three cases are randomly selected from the case base reasoning land cereal crop matching. The first group is made up of cases from the case base. The second group consists of cases which are made by modifying one of the attribute values of the case from the case base, while the third group is made up of cases which have two modified attribute values. Each test case is presented to the system individually to evaluate the performance of the similarity measures. Table 5.1.1 below shows sample queries that are used in this experiment with their values.

Table 5.1.1: Sample queries utilized in case similarity experiment with its values

	slope	Soil erosion	Drainage	Soil depth	PH value	Organic matter	texture	Average temperature	Average rain fall	Farmer skills	Equipment	Crop rotation	profitability	Government policies
Query1	12	Low	well	65	6.1	55	Loam sandy	22	700	wheat	true	Maize	tef	Wheat
Query2	8	Very low	Moderate	105	6.8	85	loam	27	650	maize	false	Wheat	barley	Wheat
Query3	2	Low	well	120	7.1	59	clay	25	450	barley	true	Maize	sorghum	Barley

Based on the above attributes listed in table 5.1.1, the next step is test each group three queries are prepared with a total of nine queries. After the query is provided to the system the similarity of the query with respect to the case are generated as shown in table 5.1.2.

Table 5.1.2: Query similarity with their corresponding cases from the case base

Query	Description of Query	With respect to case	Degree of similarity
Query 1	The same value for all attributes	Case 5	1.0
Query 2	A value of attribute "PH value" is changed	Case 5	0.94
Query 3	Values of attribute "PH value" and "Equipment" is changed	Case 5	0.79
Query 4	The same value for all attributes	Case 78	1.0
Query 5	A value of attribute "slope" is changed.	Case 78	0.86
Query 6	Values of attribute "slope" and "drainage" is changed.	Case 78	0.78
Query 7	The same value for all attributes	Case 155	1.0
Query 8	A value of attribute "soil erosion" is changed.	Case 155	0.79
Query 9	Values of attribute "soil erosion" and "average temperature" is changed.	Case 155	0.76

The case similarity test result of this experiment shows that when the test case has attributes value the same as a case stored in the case base, the degree of similarity (global similarity) becomes 1.0 (i.e. exact match) as in query 1, query 4, and query 7 as shown in table 5.1.2. On the other hand, the degree of similarity decreases when there is a change in one or more attribute values of the test case as compared to a case from the case base. When attribute values with higher weight value changes the degree of similarity highly decreases. According to query similarity with their corresponding case when attribute values with higher weight value changes the degree of similarity highly decreases. Consequently, this declination in degree of similarity due to difference in weight of attributes has significant effect on retrieval result of the prototype system.

5.2. Testing the main Cycles and Evaluating the Performance of the System

To check the validity and performance of the CBRLCCM system to domain experts, the functionality of CBRLCCM system main cycles and effectiveness of the prototype should be tested with selected cases. The effectiveness of the prototype is measured with recall and precision using test cases. In addition to that, the performance of the system was evaluated from the users' side with users' acceptance testing. With users' acceptance testing, potential users' of the system rate the applicability of the system in the process of land cereal crop matching.

5.2.1. Evaluation of the Retrieval and Reuse Process

Retrieval of previously stored cases from the case base to solve new problem by measuring the similarity of stored case and new query is the first step in JCOLIBRI for CBRLCCM system application. Retrieval of similar cases from existing cases to the new case is followed by the reuse of similar solutions with solutions of previously solved problem. Since the implementation tool JCOLIBRI uses nearest neighbor retrieval algorithm, retrieval of cases is performed using this algorithm. Nearest-neighbor retrieval technique is used to measure similarity between the source case and the case which we are searching. The nearest neighbor algorithm measures the similarity of stored cases with a new input case, based on matching a weighted sum of features. When a new case doesn't exactly match with old cases, then this algorithm will return nearest match from case base. But the retrieval time of this algorithm increases linearly as the case in the case base increases (Lang & Lau, 2002).

The nearest neighbor algorithm can be represented in the following equation (Watson, 1994).

$$\mathrm{NN}\,(\mathrm{I},\mathrm{R}) = \frac{\sum_{i=0}^{n} wi \times \mathrm{sim}(f_i^I, f_i^R)}{\sum_{i=0}^{n} wi}$$

Where:

w is the importance weighing of an attribute, I is the target case, R is source case ***i*** is individual attributes from 1 to n, ***sim*** is the local similarity function, f_j^I and f_j^R are the values for attribute ***i*** in the input case (I) and case in the case base (R) respectively and ***n*** is the number of attributes in the case base.

Similar cases to the new cases are retrieved with appropriate ranking order during retrieval process. The user of the system can use the recommended case which is retrieved based on the solution cases in a way that can fit to the new query. The process of retrieval and reuse of cases is successfully implemented in CBRLCCM system application as shown in the following figure 5.2.1.1

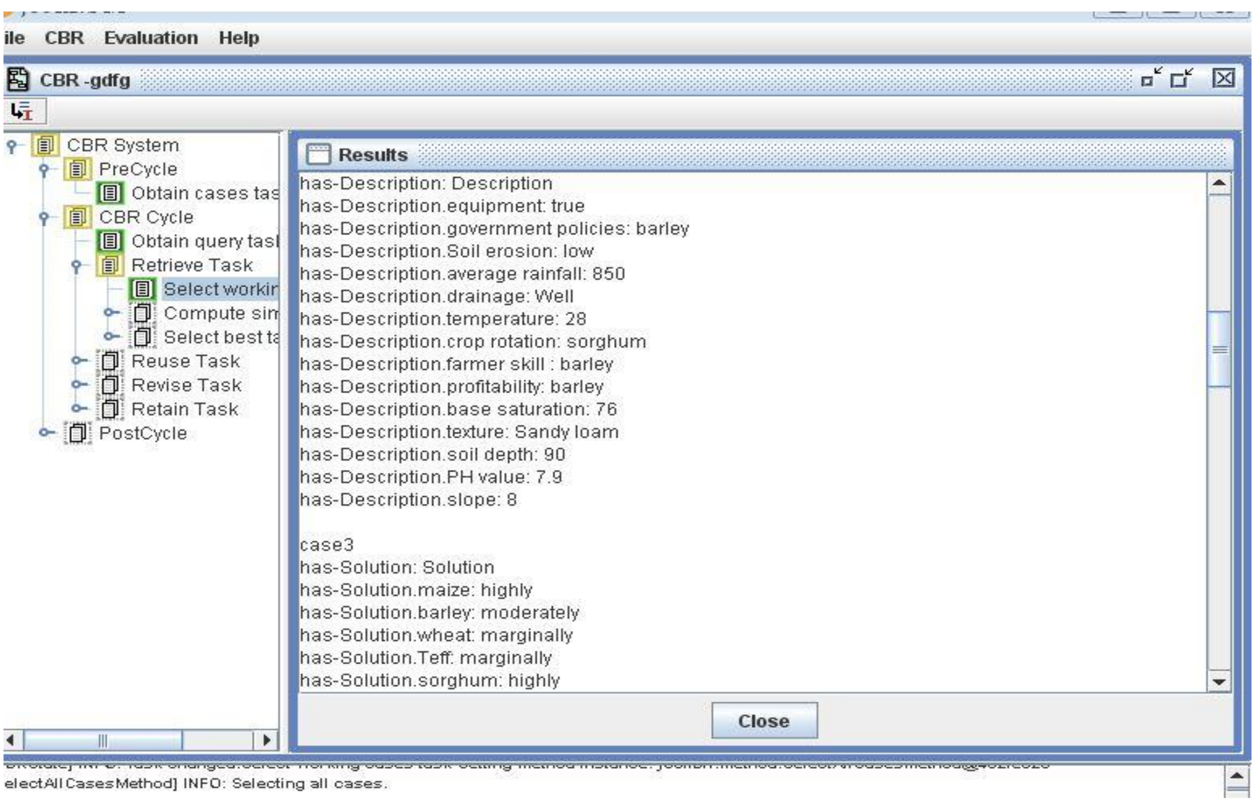

Figure 5.2.1.1: Retrieval case

As shown in the above, the system calculated the similarity of new case and existing cases and displayed the most similar cases to the new query. This is the recommended solution to the farmer based on the given case.

Precision and recall are the commonly used measures of performance of the retrieval process. Precision is the proportion of search results that are relevant to the query and recall is the ability of the retrieval system to retrieve all relevant cases to a given new query from the cases base.

$$\text{Precision} = \frac{\text{number of relevant cases retrieved}}{\text{Total Number of cases retrieved}}$$

$$\text{Recall} = \frac{\text{number of relevant cases retrieved}}{\text{Number of relevant cases in the case base}}$$

$$F = 2 \cdot \frac{\text{precision} \cdot \text{recall}}{\text{precision} + \text{recall}}$$

To start the evaluation process, for each selected test cases the relevant cases from the stored case base should be identified. The researcher determined six test cases for the system performance evaluation by confirming with the domain experts. Relevant cases for each test case are selected by the domain experts from the case base as it is shown in Table: 5.2.1.1.

Table 3.2.1.1: test cases selected by domain experts

Test case	Relevant cases from the case base
Case1	Case86, case21, case128, case66, case5, case39
Case2	Case26, case40, case67, case70, case133
Case3	Case63, case44, case30, case19, case71, case67, case9

Case4	Case78, case99, case12,
Case5	Case121, case22, case13, case266, case39, case47, case 13, case8
Case6	case32, case55, case52, case219

After the identification of the relevant cases to the test cases by the domain expert, precision and recall values were calculated with threshold interval to test the prototype using the test case query to know the performance of the system. There is no standard threshold for degree of similarity that has been used for retrieving relevant cases. That is why different researchers use different case similarity threshold to measure the performance of their system. And the researcher used a threshold value of [0.8, 1.0], which is identified after experimenting to have a better precision value. With this threshold, after computing the similarity values a minimum of 3 total cases were retrieved for test case two and a maximum of 8 total cases were retrieved for test case 3. This total cases retrieved include both relevant and not relevant cases retrieved during an experimentation as shown in (table 5.2.2.2) below.

Table 5.2.1.2: Performance Measurement of case base land crop matching system using Precision and Recall

Test cases	Relevant cases suggested by domain experts	Relevant cases retrieved by the system	Total cases retrieved by the system	Recall	Precision	F-measure
Test case1	6	5	7	0.83	0.71	0.76923
Test case2	5	4	6	0.8	0.66	0.72727
Test case3	7	5	8	0.71	0.625	0.66667
Test case4	3	3	4	1	0.75	0.85714

Test case5	8	5	5	0.625	0.71	0.76923
Test case6	4	3	4	0.75	0.67	0.75
Average				0.79	0.75	0.76

As shown in table 5.2.2.2 above, recall of each test case can be calculated by dividing the number of relevant retrieved cases with total relevant cases. For example, for the first test case, from six relevant cases selected by the domain experts the system retrieved only six cases so that the recall for Test case 1 is 0.83. And the rest recall values can be calculated in the similar manner as it is shown in Table 5.2.1.1.

Precision each test case can be calculated, it is relevant cases retrieved by total number of retrieved cases, which total numbers of retrieved cases contains both relevant and non-relevant retrieved cases within the threshold value used; i.e. [0.8, 1]. For example, Test case1 contains five relevant cases retrieved and two not relevant cases retrieved from the case base, with a total of seven retrieved cases. And so, the precision value is 0.71; and for the rest test cases, precision values are calculated in the same manner as shown in Table 5.2.2.

As shown in table 5.4, the calculated recall values for each test case is above 75%, which shows the ability of retrieval of the prototype CBR system to obtain most of the relevant cases from the case base is good. In evaluating the performance of the CBRLCCM prototype system, with recall values, has got an average recall value of 79%, which indicated a higher recall value; hence, it is clear on that the prototype system could obtain most of relevant cases from the case base. Therefore, the prototype system CBRLCCM has a capacity to retrieve relevant cases that allow effective in the process of land cereal crop matching. On the other hand, the prototype system retrieved relevant cases to the system with an average of 75% precision. The precision value of the system is not as expected by the researcher due to few number of cases used. As the number of cases increased, the precision value of the system will also increase and better performance will score in retrieving relevant cases. Although, it was difficult to attain the ideal 100%

precision and recall values in practice. As it is shown in Table 5.2.2, precision and recall values for the prototype system have been an average value of 75% and 79%, respectively. The value for F-measure which was 76% also showed good performance of the prototype system. Generally, precision, recall and F-measure average values shown us the average performance of the system as good and could be used to assist land evaluators and crop advisors' experts in their day to day decision making activities in the process of suitable analysis for cereal crops.

5.2.2. Case Revision and Solution Adaptation Testing

The purpose of testing adaptation of solutions is to evaluate the systems capability to reuse existing cases from the case base. The system loads case base at the PreCycle stage of JCOLIBRI framework and then selects working cases from the case base. These working cases always stored in to current context at the retrieval stage. The next stage is reusing the cases that are stored in the working memory and this reuse stage can be used by the next user of the system at another time. The adaptation process of CBRLCCM system is successful as the case features of the previous and new case have similar or less contradiction attribute values. Adaptation process will not be performed as the attribute values of the previous and new cases have more dissimilarity or totally different from the previous cases. In such cases, the adaptation process has to be edited and performed manually by a human domain expert in the revision stage as shown in the following figure 5.2.2.1.

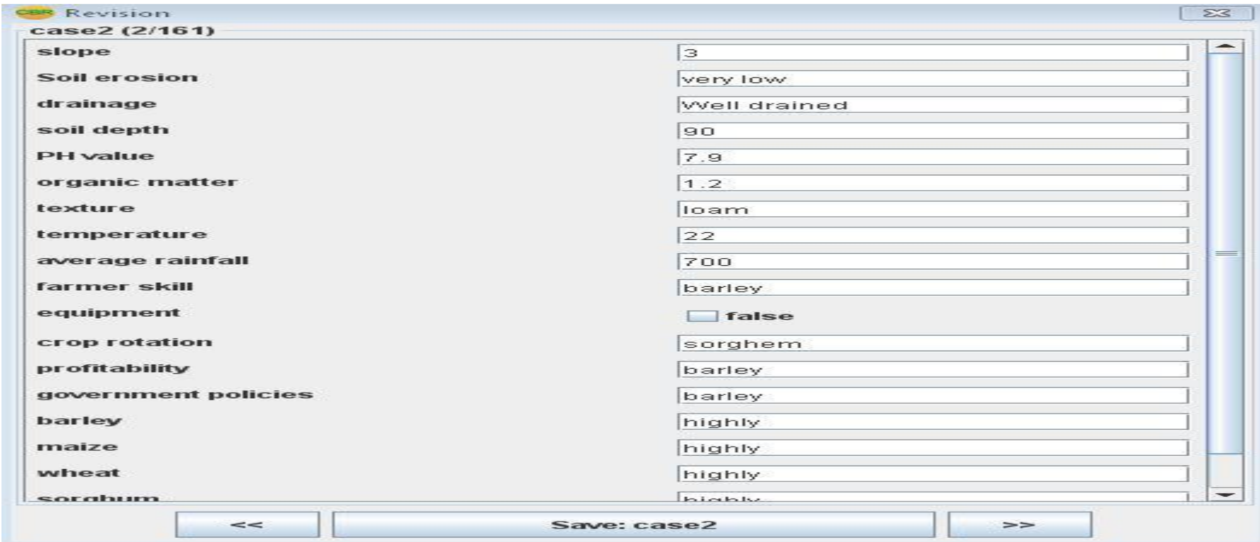

Figure 5.2.2.1: case revision testing

5.2.3. Case Retaining testing

The last cycle in JCOLIBRI framework is retaining which is an important step to store new cases together with the existing case. These new cases are used as an input or case for the next land cereal crop matching. There might be case or attributes which doesn't exist in CBRLCCM system.

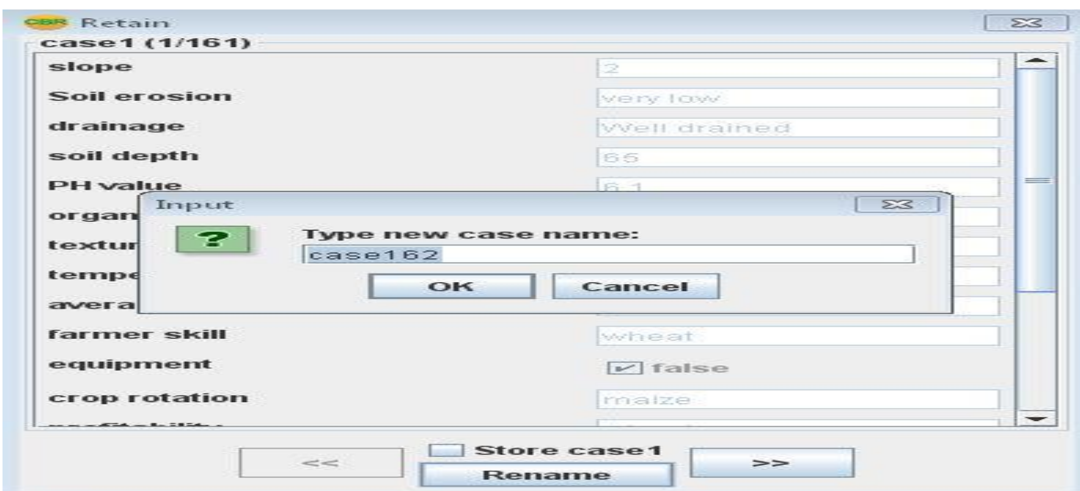

Figure 5.2.3.1: Case retaining windows

At this time, new cases which are new farm unit attribute which doesn't exist in case base library to the farmer and expert which are provide crop advice for the farmer solution process can be stored and reused by other users after some days later. This can be processed with retain stage after revision case. Manual indexing is important to assign the index value of new case as shown below in figure 5.2.3.1.

5.2.4. User Acceptance Testing

The validity of the CBRLCCM system is tested using user feedback to check its applicability in the process of land crop matching. The potential users of the system are expert who provide a service and information for a farmer under cultivation. To evaluate the applicability of the prototype, feedback was collected from domain expert who have worker in Ethiopia institute of agricultural research center in Addis Ababa. Domain experts and development agents were selected purposely from EIAR and Jimma zone of kersa Woreda respectively to evaluate the acceptance and performance of the CBRLCCM system. The type of questionnaires distributed for feedback collection from the evaluators was closed ended and open ended questionnaires focusing on easiness, attractiveness, time efficiency, and accuracy.

Evaluators were allowed to rate the options using checkbox questions. A check box question is similar to a multiple question except that it allows respondents to satisfy many of choices as part of their answer rather than just one choice. Since the system need installation of the JCOLIBRI software, only five land evaluators and three development agents were selected to evaluate the system.

The options of the check box questions are excellent, very good, good, fair, and poor for these closed ended questions which is called Likert scale questionnaire having five point for each question. Therefore, for easiness of analyzing the relative performance of the prototype based on the user evaluation after the interaction with the system, the researcher assigned numeric value for each of the options given in words. The values are given as Excellent = 5, Very good = 4, Good = 3, Fair = 2, and Poor = 1. The Table below indicates the feedbacks obtained from the domain experts (evaluators) on systems, interaction as calculated based on the given scale.

Table 5.2.4.1: performance evaluation by domain expert

No.	Evaluation Parameters	Performance Value					
		Poor	Fair	Good	Very Good	Excellent	Average
1	Adequacy and clarity of the system in recommendation process			1	3	2	4.2
2	Relevancy of the attributes in matching land cereal crops				2	4	4.7
3	Is the user interface of prototype interactive			1	4	1	4
4	Ease of use				5	1	4.2
5	Relevance of retrieved cases in the decision making to support recommendation system				3	3	4.5
6	Is the system efficient in time and memory		1	2	2	1	3.8
7	Fitness of the final solution to the new case			1	4	1	4
8	Significance of the system in domain area			1	3	2	4.3
		Total average					4.2

As shown in table 5.2.4.1, 16.67 % of the domain expert respondents' rate adequacy and clarity of the system as good, 50% rate as very good and the remaining 33.33% of the respondent's rate as excellent.

At the same time, relevancy of attributes to represent land cereal crop matching process has been rated by domain experts as 33.33% very good and 66.67% as excellent. In the case of interactive, 16.67% of respondents rate the interactive of user interface as good, 66.67% rate as very good and 16.67% rate as excellent. In the case of easy use of the system, 83.3% of the respondent rate the system as very good and 16.67% of them rate as excellent. Similarly, relevance of the retrieved cases in decision making to support users by recommending most similar cases rated as very good by 33.33% of the respondents whereas the remaining 66.67% of the respondents rate it as excellent. 16.67% of the respondent rates the efficiency of the system in time and memory as fair, 33.3% of the respondent rate as Good, 33.33% them as very good and the remaining 16.67% as excellent. In the case of fitness of the final retrieved solution to the new problem at hand, 16.67% of the respondent rate as Good, 66.67% of them rate as very Good and 16.67% of the respondent rate the fitness of final retrieved solutions as very good. The last evaluation parameter which deals about the applicability of the system to domain area also evaluated. 33.33% of the respondent rate the applicability of the prototype in their domain area as good, 50% rated as very good and the remaining 16.67% of the respondent rate as excellent.

As shown from the results in the above table 5.2.4.2, domain experts assigned less value for third criteria. This less rating might be the result of testing query used by domain experts. Finally, based on the evaluation of all the domain experts, the average performance of the prototype is 4.2(84%) which indicate the performance of the system is very good. From this performance, the researcher deduced that the prototype of the system has promising applicability in the process land cereal crop matching.

Also, the respondents confirm that this kind of system can reduce existed knowledge gap in agriculture domain in such a way that most of development agent (DA) workers at Woreda level are less qualified to identify suitable land for cereal crops. Therefore, by consulting and accessing experienced experts knowledge which is stored in the cased based form, development agents can assist and advice farmers in the process of identifying suitable land for cereal crop.

In addition to the closed ended questions, the evaluators were provided with open ended questions to forward their suggestions and opinions. These questions focusing on how the CBRLCCM is differ from the human experts or crop advisor in the process of evaluation of land use for cereal crop requirements. The contribution, strength and limitation of CBRLCCM in solving problems in the domain areas is main area of concern for the evaluation process of this prototype case based system.

For the first open ended questions, the evaluators responded that the CBRLCCM can solve problems based on the stored knowledge in the form of case based in time and cost wise but the human expert/crop advisor uses field manuals and laboratory examination which may take time and delay of timely decisions. The next open ended questions asked for the evaluators was that, do you believe that can a CBRLCCM handles suitable land evaluation for cereal crops tasks? The evaluators responded „yes" and some of the respondents stated that it can evaluate in a better way with time and cost efficiency. The other open ended question asked to the evaluators was that, can CBRLCCM contributes in land evaluation for cereal crops by assisting agriculture experts and development agent (DA) workers in the domain area? The evaluator responded that this system exactly contributes a lot, especially for those less experienced experts and development agent (DA) workers in giving timely decisions in the area thy have been working. It can be used as a training tool in the areas where shortage of skilled experts is available.

Lastly, the evaluators suggest the strength and the limitation of the prototype as they evaluate from its performance during the visual interaction with the system:

- Users who lack computer skills and access might not implement it, especially those who have no computer access.

The evaluators confirm some the strength of the system and its applicability in the domain area

as:

- This kind of prototype case based reasoning system helps to solve problems in the areas where experienced and skilled Agriculture experts are unavailable.
- The system is very helpful to solve problems timely with accumulated knowledge by based on the existing or current physical factor as well as socio-economic factor for land evaluation of cereal crops in remote areas.
- The system can reduce the existed knowledge gap observed in remote areas where skilled agricultural professionals are not available. By being a sharing and training tool for development agent (DA) workers, it can improve the skill of development agent (DA)

workers in land evaluation for suitable cereal crops identification and decision making at farmer's hand.

5.3. Discussion

As it is discussed in the evaluation section, the proposed system of precision, recall, f-measure and user acceptance values for the prototype system have been achieves an average value of 75%, 79%, 76%, and 84% respectively. The user acceptance testing result is better as compared to (Ejigu,2012) who developed KBS for cereal crop diagnosis and treatment particularly for wheat and barley crops that achieved 65%, 72% and 80.9% overall system performance of precision, recall and F-measure respectively by using interviews and documents analysis as means of knowledge acquisition technique which in turn indicates that using integrated (manual and automated) knowledge acquisition techniques is better than using manual knowledge acquisition techniques. Also according to (Ali, 2009) integrating case based reasoning and geographic information systems in a planning support system which achieved 80% overall performance. This research has been proofed that ArcGIS are the best tools for handling and generating geo-spatial data and integrating with CBR for providing a geo-spatial planning support in city planning practice through a practical example. However, in these approaches CBR only uses local inference, which doesn't consider external model or rule and Similar problems may not have similar solutions. In some situations, CBR technique must derive external rules/models to empower itself. Therefore, factor which are important for land evaluation analysis are spatial information which could be handled, analyzed and generated in GIS effectively and integrating with CBR for providing a geo-spatial information in the process of land evaluation for suitable cereal crops. In addition, CBRLCCM is taking in account for local as well as external rule or model so as to empower the system.

Furthermore, according to Gizachew (2015) Land suitability mapping and analysis is a prerequisite to achieving optimum utilization of the available land resources. The objective of this study was to spatially evaluate land suitability for groundnut and sweet potato crops in the east Amhara region, Ethiopia which was based on FAO guidelines. Geographical Information System (GIS) was used to create land suitability map. The criteria for crop suitability analysis were soil depth, soil texture, pH, organic carbon and temperature. Crop suitability map was made by matching between reclassified land characteristics with crop requirements using GIS model builder. The main restricting factors for these research land suitability analysis was that it doesn't consider the dynamic condition of different factor which is important for decision support system in agricultural land use planning. GIS can evaluate based on the existing or current condition of specific place land quality however; this conditions are dynamic it could be changed over time. In addition, it neglects social and economic factors. Socioeconomic conditions of farmers have a considerable effect on land use decisions. To alleviate these CBRLCCM has considered aforementioned dynamic factors which were neglected by Gizachew's system. In addition, it can produce suitability analysis class for five different type of cereal crops simultaneously. This research adopted a knowledge base approach to farm level agricultural planning that combines case-based reasoning systems with GIS tools which includes different types of factors such as qualitative, quantitative and heuristic factors that are associated with land evaluation for choosing an optimal cereal crop for a farm unit.

The current study has raised three research questions to be answered at the end of this research work. The first research question of this study was "what kinds of case are acquired for land cereal crop matching decision?" To answer these question different documents such as FAO guidelines have been intensively reviewed, and domain expert's interview were used and figured out the physical and environmental parameter, social attributes and economic indicator as necessary factors for best possible outcome in the process of land evaluation for suitable cereal crops. Spatial information like maps and land related attributive characteristics are incorporated into GIS based data of land suitability evaluation as a system.

Secondly "what are the most suitable techniques for the acquired case modeling and case representation?" To answer this question, physical and environmental parameters, socio-economic factors are modeled using hierarchical tree structure and feature value format representation technique was implemented so as to develop case based reasoning system for land cereal crops.

Finally, after developing the prototype system, testing of the CBRLCCM is done to evaluate the performance of the system. Based on prototype testing, 75%, 79%, 76% and 84% average values for precision, recall, F-measure and user acceptance is registered respectively which indicates that the proposed system is prominent in provision of crop advisory decisions by matching land with suitable cereal crops on particular farm unit.

CHAPTER SIX
CONCLUSION AND RECOMMENDATION

6.1. Conclusion

In terms of caloric intake, cereals dominate the diets of Ethiopian households. However, cereal production in Ethiopia is constrained by several factors. According to many studies it is identified that the main reason for low production is lack of skill on selection of suitable land for cereal crop due to these the farmers don't know what kind of crop they should cultivate for their farm units instead; they are cultivating crops in traditional methods rather than following scientific approach. To develop the case based system, the knowledge was acquired from domain experts and documented sources through interview and document analysis respectively.

Both tacit and explicit knowledge for the study was acquired from domain experts, reviewing different documents and FAO guideline were used and this study find out that physical and environmental parameter, social attributes and economic indicator as necessary factors for best possible outcome in the process of land evaluation for suitable cereal crops from Ethiopian institute of agricultural research center. Relevant attributes in the form of case structure which will have direct impact for decision was also identified. This extracted relevant knowledge were modeled using hierarchical conceptual modeling method and the model has been converted. For representing the knowledge feature-value pair format constructed for matching land use with cereal crop requirements. After the acquired knowledge is modeled and represented the case based reasoning is developed. This research adopted a case base reasoning approach to farm level agricultural planning that combines cased base reasoning systems with GIS tools. The case base is developed using JCOLIBRI case based framework tool which is the most compatible and reliable tool to develop CBR system. Where ArcGIS was used to preparing and handling spatial and non-spatial so as to jCOLIBRI used to utilize the GIS database as a case library, indexing of these cases and retrieval processes. Spatial and non-spatial attributes were generated in ArcGIS as a tabular input for the jCOLIBRI. Cases were represented with attribute-value format. The prototype system CBRLCCM is developed by using jCOLIBRI Programming tool. Since the system stores the new case within the existing cases, new case can be used as a case base for the next time.

To assess its performance and user's acceptance in the domain area, the CBRLCCM was evaluated using user feedback through visual interaction method which scored average value of 84% which is a higher acceptability from the end user. Moreover, system is encouraging since the retrieval performance of the prototype registers an average value of 79% recall, 75% precision and 76% F-measure. As a result; it can be concluded that, CBRLCCM can be used in supporting decisions in land suitable analysis for cereal crops.

Furthermore, this research approach is concerned with land evaluation as this can represent different types of factors which are qualitative, quantitative and heuristic factors that are associated with land evaluation for choosing an optimal crop for a farm unit. The CBRLCCM enables the easy integration of these factors thus enabling collaborative decision making process. The Prototype CBRLCCM system developed and tested in this research work is an important tool in timely crop advisor who can provide assistant to the farmer as well as advising agriculture experts. Such system can be really useful in rural areas where a shortage of agricultural expert is available who works in closer to the farmers.

The system is also highly appreciated by domain experts in assisting agriculture experts and development agents in areas skilled experts are not accessible. By being a sharing and training tool for development agent (DA) workers, the proposed system can improve the skill of development agent (DA) workers in land cereal crop matching and decision making at farm hands. By replicating this CBRLCCM, it can reduce the existed knowledge gap observed in remote areas where transferring skilled experts is difficult. The idea behind developing CBRLCCM system is that, it can enable many people to benefit from the knowledge of one expert in domain specific problems. So, case based reasoning matching land cereal crop system developed in this research is applicable and promising to apply the judgment and experiences of domain experts in land suitable analysis for crop and providing the descriptions for farm unit under cultivation.

5.2. Recommendations

The system achieves its objectives by demonstrating the applicability of case based reasoning by developing a case-based reasoning approach for matching land use with cereal crop requirements hopeful of performance and user acceptance. This thesis research is the promising study for further research works to fully implement the system in the agricultural domain area. As a result, the following recommendations are given based on the observed opportunities and uncover areas by this research. These recommendations are made for further investigations to fully implement the functionality of the prototype or to develop a new system in the domain area.

At present, the system suggests suitable cereal crop choice for a given time. It can be further extended, with inputs from climate prediction models, to predict future land use choices, what crops would be suitable after five, ten or twenty years, for a land unit based on predicted changes in climate.

The developed system assists planners and decision makers in the selection of appropriate scenario for only common cereal crop which cultivate in Ethiopia. For future work, the researcher recommends to develop a system which includes all kind of crops which cultivate in Ethiopia.

Additional land uses with their knowledge base approach to reflect their land use requirements can be added, so that the functionality of the system can be extended to industry site selection, ecological evaluation and impact assessment purposes.

The researcher has represented few cases in the knowledge base. This is due to that land suitability analysis cases found distributed in different region that challenging to integrating together. The system can be more accurate if it includes all alternative case with the solution.

The researcher recommends that it is better to create a web-based case based reasoning system that connects with various knowledge bases to access different cases in real-time to deliver agriculture recommendations. This can be achieved by having multiple knowledge bases, one for each scientific discipline, connected by a communication network which would serve the system and by this way, domain experts can keep knowledge bases updated and accurate. Such web-based systems are necessary in even non-agricultural fields such as Medicine, Law and Finance as the web capability enable to deliver advices and suggestions anytime and anywhere.

The use of nearest algorithm increases linearity of retrieval time when there are many cases and it returns the nearest match even with dissimilarity cases in the source and new case. In the future it is recommended to use other retrieval algorithms such as template retrieval that returns all cases that fits certain parameter.

The performance of the system can be improved if hybrid approach is employed by combining rules, cases and models since these rules, cases and models have complementary strength. For the future, it is better to integrate these approaches to make knowledge base system more successful.

Reference

Ali, K. (2009). Integrating case based reasoning and geographic information systems in a planning support system : çesme peninsula study, Izamir.

Abrham, T. (2009). Survey on farmer's participation in agriculture technology transfer Ethiopian Agriculture Research Organization, Ethiopia.

Abu-Naser, S.S. Kashkash, K.A. and Fayyad, M. (2008). Developing an Expert System for Plant Disease Diagnosis. Journal of Artificial Intelligence, 1: 78-85.

Aamodt, A. and Plaza, E. (1994). Case-based reasoning: Foundational issues, methodological variations, and system approaches. AI Communications, 7(i).

Beek, K. (1980). From soil survey interpretation to land evaluation. European conference on land evaluation LNAI 2401, Springer, Veriag, 236-305.

Belayhun, H. (2010). Present and Future Trends of Natural Resources (Land and Water) Management in Ethiopian Agriculture. Report paper on Workshop by MOWR and ILRI, Addis Ababa, Ethiopia.

Bonzano, A., Cunningham, P. and Smyth, B. (1997). "Using Introspective Learning to Improve Retrieval in CBR: A Case Study in Air Traffic Control," in Proceedings of second International Conference on Case-based Reasoning, RI, USA, 291-302.

Brooks, F. (1996). "The Computer Scientist as Toolsmith II." Communications of the ACM

39 (3): 61-68.

Burrough, P.A., R.A. McDonnell and R. McDonnell. (1998). Principles of geographical information systems Oxford university press Oxford.

Calu, F. (2009). Major Pests, Diseases and Weeds of Cereal Crops: retrieved October2011, fromhttp://www.pesticides.gov.uk/.

Chamberlin, J. and Schmidt, E. (2011). Ethiopian Agriculture: A Dynamic Geographic Perspective. Development Strategy and Governance Division, International Food Policy Research Institute. Ethiopia Strategy Support Program II, Ethiopia.

Chen, Y. L., Hsu, C. L., & Chou, S. C. (2003). Constructing a multi-valued and multi labeled decision tree. Expert Systems with Applications, 25, 199–209.

Davidson, D. A. (1992). The Evaluation of Land Resources. Stirling University press, USA, NewYork.

Dent, D. and A. Young. 1993. Soil survey and land evaluationE & FN Spon.

Diao, X., and Alejandro, N. (2010). Growth options and poverty reduction in Ethiopia – An economic wide model analysis. Food Policy 32 (7): 205-228.

Elaalem, M. (2010). The Application of Land Evaluation Techniques in Jeffara Plain in Libya using Fuzzy Methods.

Ebrahim, E. (2014). Land Suitability Assessment for Sorghum and Maize Crops Using as SLA and GIS Approach in, Dera Woreda, Ethiopia.

FAO. (1976). A framework for land evaluation Food and Agricultural Organization of the United Nations Rome-Italy.

FAO. (1983). A framework for land evaluation Food and Agricultural Organization of the United Nations Rome-Italy.

FAO. (2007). LAND EVALUATION: Towards a revised framework, Land and water discussion paper. Food and Agricultural Organization of the United Nations, Rome, Italy.

Fasil, K. (2002). Analysis of Yield Gap for Wheat Cultivation in the Highlands of North Ethiopia. Ph.D. Thesis, Gent University, Belgium.

Flowerdew, R. (1991). Spatial data integration. Geographical Information Systems 1: 375.

Gatheru, M. & Maingi, P.M. (2010). Evaluation of Land Suitability for Maize Using Geographic Information Systems (GIS): A Case Study for Two Maize Varieties in Machakos District. Proceeding at the 6th National ERJSSH 1 (1), September-October 2014 Fertilizer Conference, Loresho, Nairobi, 20-21st August, 309-315.

Gizachew, A. (2015). A Geographic Information System Based Physical Land Suitability Evaluation to Groundnut and Sweet Potato in East Amhara , Highlands of Ethiopia, 5(1), 33–39.

Goodwin, C. J. (2005). Research in Psychology: Methods and Design. USA: John Wiley & Sons, Inc.

Hailu Beyene (2012). Adoption of improved tef and wheat production technologies in crop livestock mixed systems in Northern and Western Shewa Zones of Ethiopia. University of Pretoria, Pretoria.

Holt, A., & Benwell, G. (1996). Case Based Reasoning and Spatial Analysis. Urisa Journal.

Hogeveen, H, Noordhuizekstassen,E. N., Schreinemakers, J. F. and Brand, A.(1991).

Our Industry Today Development of an Integrated Knowledge-Based System for Management Support on Dairy Farms. Journal of Dairy Science Vol. 74, No. 12, 1991. Holland, A. (2011). Rule Based System: Intelligent Systems retrieved on February 16, 2012 from http://www.4c.ucc.ie/~aholland/udg/Girona_Lec2.pdf.

Iqbal, N. and Ashraf, M. (2006). Master Thesis Report. Evaluation of jCOLIBRI.Malardalen University, Vasteras, Sweden, Vasagatan 44, 72-123.

Leake, B. (1996). CBR in context: The present and future, Book: Case-Based Reasoning: Experiences, lessons, and future directions, Published: AAAI Press/MIT Press, 1-35.

Leake, B. and Wilson, C. (1994). "Categorizing Case-Base Maintenance: Dimensions and Directions," in Proceedings of 4th European Workshop, EWCBR-98, 197-207.

Kassa, B. and Degnet, A. (2009). Challenges Facing Agricultural Extension Agents: Case Study from South-western Ethiopia. Blackwell Publishing Ltd.UK.

Kaster, S. (2005). Supporting modeling and problem solving from precedent experiences: the role of workflows and case-based reasoning. Environmental Modeling & Software, 20 (6), 689-704

Kebede, G. (2014). Land suitable assessment for wheat and barley in Asela, Arsi, Ethiopia.

Kemelew, M. and Alemayehu, A. (2011). Diversity and Agronomic Potential of Barley (Hordeum vulgare L.) Landraces in Variable Production System, Ethiopia: World Journal of Agricultural Sciences 7 (5): 599-603, 2011 ISSN 1817-3047. IDOSI.

Kolodner, J. (1993). Case-Based Reasoning, Morgan Kaufmann Publishers: California.

Makfi, P. (2011). Introduction to knowledge modeling and neural network. Available at: http://www.makafi.com/KCM_intro.html.

Messing, I., M.H. Hoang Fagerstrom, L. Chen and B. Fu, 2003. Criteria for land suitability evaluation in a small catchment on the Loess Plateau in China. Catena, 54: 215-234.

Minale, K. (2009). The Contribution of Sustainable Agriculture and Land Management to Sustainable Development. Journal of sustainable development innovation briefs. United Nation Department of economics and Social Affairs publishing. University of Gothenburg, Sweden.

Mellor, J. and Dorosh, P. (2010). Agriculture and the Economic Transformation of Ethiopia Development Strategy and Governance Division: Ethiopia Strategy Support Program 2. International Food Policy Research Institute.

MulatDemeke (2009). Agricultural Technology, Economic Viability and poverty alleviation in Ethiopia, Addis Ababa University.

Malczewski, J. (2004). GIS-based land-use suitability analysis: a critical overview. Progress in Planning 62: 3-65.

Pal, S., Dillon, T. and Yeung, D. (2001). Soft Computing in Case Based Reasoning, (Eds.) Springer-Verlag (London) Ltd, 21 (9), 1-28.

Pal, Sankar, K. & Shiu, Simon C. 2004. Foundation of soft-Case Based Reasoning. A JOHN WILEY & SONS, INC., PUBLICATION.

Pirbalouti, A.G. (2015). GIS Based Land Suitability Assessment for German Chamomile Production. Bulgarian Journal of Agricultural Sci- ence 17 (1), 93-98.

Prentzas, J. and Hatzilygeroudis, I. (2007). Categorizing approaches combining rule-based and case-based reasoning, Journal Compilation Blackwell Publishing Ltd. 24(2).

Prentzas, J. and Hatzilygeroudis, I. (2002). Integrating Hybrid Rule-Based with case based reasoning. European conference on Case based reasoning LNAI 2416, Springer, Veriag, 336-349.

Prasad, G. and Babu, V. A. (2006). A Study on Various Expert Systems in Agriculture: Georgian Electronic Scientific Journal of Computer Science and Telecommunications. 2006, No. 4(11).

prentzas, j., and Hatzilygeroudis, l. (2007, 5 1). Categorizing approaches combining rule-based and case-based reasoning. Expert systems,24(2), 97-122.

Rabia, H. (2012). A GIS Based Land Suitability Assessment for Agricultural a Planning in Kilte Awulaelo District, Ethiopia. The 4th International Congress of ECSSS, EUROSOIL 2012, 2-6 June, Bari, Italy. 12-57.

Rafea, A., El-Azhari, S., Hassan, E. (1998). Integrating Multimedia with Expert Systems for Crop Production Management. Proceedings of the Second International IFAC Workshop on Artificial Intelligence in Agriculture, Wageningen, Netherlands.

Rashid, S. (2010). Variation in staple food prices: International Food Policy Research institute, Addis Ababa, Ethiopia.

Recio, A. and Diaz-Adugo, B. (2002). On developing a Distributed CBR Framework through semantic Web Services, University of Madrid, Spain.

Reddy, KP & Ankaiah, R. (2005). A framework of information technology-based agriculture information dissemination system to improve crop productivity, Current Science**,** vol. 88, no. 12, pp. 1905-1913.

Rossiter, D.G. (1996). A theoretical framework for land evaluation. Geoderma 72: 165-190.

Salem, A., Roushdy, M. and HodHod, R. (2007). A Case Based Expert System for Supporting Diagnosis of Heart Diseases. AiML Journal, 5 (1).

Sajja, P. S., &Akerkar, R. (2010). Knowledge-Based Systems for Development. In Advanced Knowledge Based Systems: Model, Applications & Research (pp. 1-11).

Sarma, S. Kr., Singh,Kh. R. & Singh, A. (2010). An Expert System for diagnosis of diseases in Rice Plant. International Journal of Artificial Intelligence, 1(1):26-31.

Shi, X., & Yeh A. (2001). Case-Based Reasoning (CBR) in Development Control. International Journal of Applied Earth Observation and Geoinformation, 3 (3), 238-251

Shiu, C. and Pal, S. (2004). Case-Based Reasoning: Concepts, Features and Soft Computing. Kluwer Academic Publishers. Manufactured in The United States. Applied Intelligence 21, 233–238.

Simon, H. (1996). The Sciences of the Artificial, Third Edition. Cambridge, MA, MIT Press.

Smyth, B. and Keane, A. (1998). "Adaptation-guided retrieval: questioning the similarity assumption in reasoning," Artificial Intelligence.

Teklu, F. (2005). Land Suitability Assessment for Sorghum and Maize Crops Using as SLA and GIS Approach in, Dera Woreda, Ethiopia.

Verheye, W. 2008. Land evaluation systems other than the FAO system. National Science Foundation Flaners/ Belgium and Gegraphy Department University Gent, Belgium II.

Watson, I. (1997). Applying Case-Based Reasoning: Techniques for Enterprise Systems. Retrieved from: http://www.elsevier.com/wps/find/bookdescription.cws_home/680525/description#description

APPENDIXES

Appendix I

Interview questions to Domain Experts

After introducing the objective of the study and confirm the respondents' willingness to participate in the study the interviewer records their answers for the following interview questions.

1. What factors used in typical land evaluation and crop suitability analysis process?
2. What are the main criteria for assessing land suitability for cereal crops?
3. What factors are given more consideration in the process of land evaluation for cereal crop suitability analysis?
4. How do you prioritize attributes in the process of land suitability analysis for cereal crops?
5. What are the identification techniques and procedures you applied to land evaluation analysis for cereal crop suitability analysis?
6. What tools and materials are used in the process of land evaluation for cereal crops suitability analysis?
7. Do you use standardized guidelines/manuals to evaluate cereal crop suitability analysis?
8. What are the steps undertaken to land evaluation for cereal crop suitability analysis?
9. What are the current problems in the land suitable analysis that affect the decision making process in Ethiopia?
10. What alternative methods should be taken so as to overcome these limitations in the process of land evaluation and cereal crop suitability analysis?
11. After evaluated of cereal crop suitability, what method or procedure do you used to advise the farmer?
12. What tools or system do you used to identify the land suitability analysis for cereal crops

Appendix II

Prototype Evaluation form for the Domain Expert

This is an evaluation form to be filled by land evaluator and crop advisor in order to evaluate the case based reasoning system for land evaluation for cereal crop suitability analysis. Description of the parameter values are as follows.

Performance value	1	2	3	4	5
Description	Poor	Fair	Good	Very good	Excellent

Instruction: please assign (X) on the appropriate value for the corresponding parameter of evaluation questions of the case based reasoning system for land cereal crop matching.

No.	Evaluation Parameters	Performance Value					
		Poor	Fair	Good	Very Good	Excellent	Average
1	Adequacy and clarity of the system in recommendation process						
2	Relevancy of the attributes in representing field of study selection process						
3	Is the user interface of prototype interactive						
4	Ease of use						
5	Relevance of retrieved cases in the decision making to support recommendation system						
6	Is the system efficient in time and memory						
7	Fitness of the final solution to the new case						
8	Significance of the system in domain area						

II. Please state your opinions and suggestions regarding the performance and applicability of the CBRLCCM in the domain area.

9. How is CBRLCCM differing in evaluating the cereal crop from the human experts (crop advisor)? ____

____

10. Do you believe that, can CBRLCCM prototype handle land evaluation for cereal crop tasks?

____

11. Do you think that, can CBRLCCM contribute in evaluating of lands for cereal crops by assisting agricultural experts and DA workers in the domain area?

____

12. What are the limitations of CBRLCCM as you see its performance when you are interacting with it?

____

13. What are the strengths of the CBRLCCM in solving problems in the domain areas?

____