LUMA-GIS Thesis nr 33

Using GeoTools to Implement the Multi-Criteria Evaluation Analysis - Weighted Linear Combination Model

Salem Jamal-Uddeen

2014
Department of
Physical Geography and Ecosystem Science
Centre for Geographical Information Systems
Lund University
Sölvegatan 12
S-223 62 Lund

Salem Jamal-Uddeen (2014). Using GeoTools to Implement the Multi-Criteria Evaluation Analysis - Weighted Linear Combination Model.
Master degree thesis, 30/ credits in Master in Geographical Information Sciences
Department of Physical Geography and Ecosystems Science, Lund University

Using GeoTools to Implement the Multi-Criteria Evaluation Analysis - Weighted Linear Combination Model

Author
Salem Jamal-Uddeen

Master thesis, 30 credits, in Geographical Information Sciences

Supervisor
Cecilia Olsson

Department of Physical Geography and Ecosystem Science, Lund University, Lund, Sweden.

To my parents, Mohammed Hassan and Sawsan
To my dear wife, Reham
To my wife's parents, Mohammed and Samar
To my lovely little daughter, Yara
To my sisters, Noor and Ghoffran
To my brothers, Ayman and Abdulrahman
To my close friends, Qurainawi, Shousharah, and Mazen
To every knowledge seeker, scientist and student

Table of contents

Acknowledgements

First of all, the first thanks and appreciation is to Almighty Allah who allowed me by his willing to finish this program and I hope that I could expand and spread the knowledge I gained from this program in every possible way.

Special thanks to Sweden, the Swedish people, Lund University, all the staff of LUMA-GIS program, Lund University GIS Centre, and to every person who has given any effort to make this program work and succeed.

Great thanks to Ms.Cecilia Olsson who supervised this thesis sincerely, provided support and guidance all over the process of the thesis, and has always been a great motivator to finish it.

Thanks to Eng.Khaled Masri from Standards Associates for his support and guidance during initialization of this thesis. He has not spared any possible help.

The greatest and most appreciating thanks to my parents, my father Mohammed Hassan and my mother Sawsan, for being always there for me with all their support and feelings, for raising me, and helping me reaching this extent of my life. If I wrote a thousand page of thanks to my parents, I would not suffice a bit of what they have offered me over the past thirty years of my life and over the future years.

I cannot express how much love, respect, and appreciation I bear in my heart to my wife, Reham, for being always beside me, holding my back, standing the hard times I have been through the whole program studying years, and for being patient when I was busy and away from her and from our new born daughter, Yara.

Finally, I would like to express my thanks and appreciation to my sisters, Noor and Ghoffran; my brothers, Ayman and Abdulrahman; and my close friends, Qurainawi, Mazen, and Shousharah for being always caring, looking after me for all my life, and for their attention to keep my life smooth to gain this degree.

List of Acronyms

API	Application Programming Interface
CSP	Concentrated Solar Power
DNI	Direct Normal Irradiance
ESRI	Environmental Systems Research Institute
FOSS	Free and Open Source Software
FOSS4G	Free and Open Source Software For Geospatial
GPL	General Public License
GRASS	Geographic Resources Analysis Support System
GIS	Geographic Information System
GUI	Graphical User Interface
IDE	Integrated Development Environment
Java SE SDK	Java Standard Edition Software Development Kit
JTS	Java Topology Suit
LGPL	Lesser General Public License
MADA	Multi-Attribute Decision Analysis
MCDA	Multi-Criteria Decision Analysis
MCDM	Multi-Criteria Decision Making
MCE	Multi Criteria Evaluation
MCEA	Multi Criteria Evaluation Analysis
MC-SDSS	Multi-Criteria Spatial Decision Support System
MODA	Multi-Objective Decision Analysis
NaN	Not a Number

OGC	Open Geospatial Consortium
OSGeo	Open Source Geospatial Foundation
OS	Open Source
SDI	Spatial Data Infrastructure
SDK	Software Development Kit
SRP	Score Range Procedure
SRS	Spatial Reference System
SWT	Swing Weights Technique
TIGER	Topologically Integrated Geographic Encoding and Referencing
USA	United States of America
WFS	Web Feature Service
WLC	Weighted Linear Combination
WMS	Web Map Service

List of Tables

List of Figures

Abstract

Optimization problems have always been arising in all fields originated from operational research and management science. Different models for solving optimization problems were developed. The wide application of optimization solutions reached many different disciplines. Since Geographic information Systems (GIS) appeared in the 1960s, there have been many early attempts to employ optimization models to solve spatial-related optimization problems. Starting in the early 1990s, a considerable interest of incorporating optimization models of Multi-Criteria Evaluation Analysis (MCEA) in GIS has started to appear through the increasing number of refereed articles and applications of MCEA and GIS. Among these applications was the problem of locating best site for some activity or facility.

Different models of MCEA have been used in different GIS applications. MCEA involves the analysis of different factors to solve the problem in hand. When this method of analysis is integrated with GIS, factors of spatial nature are employed to solve problems that have spatial aspects as well. In this thesis, a desktop computer application, MC-Analyst, has been developed using GeoTools (www.geotools.org), an open source GIS software library, to implement one model of the MCEA: Weighted Linear Combination (WLC).

MC-Analyst has been tested on a previous case study that located best sites to build solar farms. The case study was conducted on Colorado, USA. Factors involved in the study area were solar Direct Normal Irradiance (DNI), distance to transmission lines, distance to primary roads, distance to cities, population density, and federal lands. Federal lands layer was used as a constraints layer.

The output of MC-Analyst was examined by altering input criteria weights for the main layers of solar DNI, distance to transmission lines, and distance to primary roads in the GIS-MCEA/WLC model.

Different outputs of MC-Analyst have been analyzed and found as expected that when a higher weight had been given to one criterion, the final suitability result showed noticeable higher suitability in areas where that factor was considered optimal.

Raster and vector data processing capabilities of GeoTools were explored to perform basic GIS functions and it showed strong support for such functions. It was also found that in order for MC-Analyst to work as a standalone fully integrated Multi-Criteria Spatial Decision Support System (MC-SDSS), it would require other tasks that are found in common desktop GIS applications such as re-projection, clipping, and resampling. Therefore, re-using the source code of an open source

desktop GIS application such as uDig (a product of Refractions Research, Victoria, British Columbia, Canada (udig.refractions.net)) or gvSig (www.gvsig.org) and integrate the developed MCEA-WLC model with it to produce fully integrated MC-SDSS could be of great interest for future development for Free and Open Source GIS software (FOSS GIS) community. This would create an effective product that might compete with effective proprietary GIS software that implements MCEA models.

The study showed also that before choosing a software library to start with, the development team should be acquainted with other possible open source and free software applications and libraries in order for the development process to be smooth and to output a productive application.

1 Introduction

1.1 Background

Developing efficient software is affected by multiple factors such as available technology, target users, user requirements, and development teams' skills and experience. Free and Open Source Software (FOSS) GIS has been going through huge development efforts to reach its current popularity and wide use of today. Even though proprietary GIS software such as ArcGIS (a product of Environmental Systems Research Institute (ESRI), Redlands, California, USA) dominates the field in business and education, there are broad applications in which FOSS GIS was employed by commercial organization and education institutes alike. It has also gone far beyond that, some governmental bodies replaced its proprietary GIS software with a FOSS GIS alternative.

Applications on using FOSS GIS varied from simple desktop and web mapping up to public participation and web-based analysis (Jankovic and Milidragovic, 2013; Steiniger and Hay, 2009; and Steiniger and Hunter, 2012a). A spatial optimization problem or a spatial decision making problem such as suitability analysis to find best location is one of the major problems that arise frequently whenever an urban planning or a facility locating question is asked. The concentration of this study is the combination of FOSS GIS and suitability analysis through Multi-Criteria Decision Analysis (MCDA).

When it comes to site selection problems, there are many studies in GIS for different disciplines. While GIS provides the tools for analyzing spatial decision problems, MCDA provides the tools to structure the problem (Malczewski, 2006). GIS-MCDA has been applied in many application fields such as environmental planning and management, transportation planning and management, urban and regional planning, waste management, hydrology and water resource, agriculture and forestry, geology and natural hazard, and real estate and industrial facility management (Malczewski, 2010).

There is a large number of methods of MCDA. Greene et al. (2011) listed 23 methods that perform MCDA. Many of them are applied on a small number of alternatives due to computational or practical limitations (Greene et al., 2011).

According to Malczewski (2006), there are two main categories of GIS-MCDA methods based on the objective: 1) Single Objective Multi-Attribute that is also referred to as Multi-Attribute Decision Analysis (MADA), and 2) Multi-Objective Multi-Attribute that is also referred to as Multi-Objective Decision Analysis (MODA).

In the post-1990 era, there has been exponential growth of the number of published articles in refereed journals starting from less than 10 articles in 1991 up to 319 articles in 2004 (Malczewski, 2010). The nature of how MCDA models screen and solve the problem through sharing multiple decision makers helped communicating, sharing, and refining the decision making process (Malczewski, 2010). In addition, public participation in the decision-making process would increase democratization and decision sharing in the society (Malczewski, 2010).

Malczewski (1999) provided a very good reference about different GIS-MCDA methods such as Weighted Summation/Overlay, Ideal/Reference Point, and Analytical Hierarchical Process (AHP). According to Malczewski (2006), the method of Weighted Summation was the mostly used method (39.3% and 9.4% of MCDA published articles for the methods of Weighted Summation and AHP respectively).

The main reason behind this spread was that Weighted Summation methods are implemented using simple map algebra operations (such as multiplying and summing rasters) which are easy to understand and intuitively appealing to decision makers who are a primary contributor in every MCDA problem solving. The method that was implemented in MC-Analyst was the WLC that belongs to the Weighted Summation category of GIS-MCDA.

1.2 Aim

The main aim of this study was to use FOSS GIS through GeoTools to develop a custom desktop application. Secondary aims of this study were to address implementations of the WLC model in proprietary GIS and FOSS GIS and to study the feasibility of developing MCEA in FOSS GIS.

1.3 Method

The requirements of MC-Analyst were extracted from Malczewski (2000). In this article, Malczewski discussed different phases of applying MCEA-WLC model on raster data. After the requirements had been addressed, GeoTools library was explored, tested, and learned to fulfil those requirements. Finally, MC-Analyst was built gradually by the tested code.

To prove the liability of MC-Analyst, a previous case study of locating best site areas to build solar farms in Colorado State, USA, was employed for testing. Sensitivity analysis to ensure the proper results of MC-Analyst has been carried out through weights alterations. Weights alteration was accomplished by assigning different weights for the same set of criteria for every iteration, comparing, and analyzing the final suitability results for the alterations. Three weights alterations were carried out in this study.

1.4 Disposition

To reach the objectives of this study, many questions had to be answered and investigated along the process. The milestone works in this study could be listed in four points: 1) The feasibility of using FOSS in general and FOSS GIS in specific was researched, 2) A survey of the literature that researched previous efforts and implementations for integrating MCEA in GIS software was carried out, 3) Experiencing software development using a FOSS GIS library, GeoTools, in order to address strength and weakness points for development using GeoTools and FOSS GIS, and 4) Some recommendations concerning future development approaches to produce better integrated MCEA and FOSS GIS systems were addressed for future works.

MC-Analyst could be used for any GIS-MCEA/WLC problem such as locating schools, parks or any facility. As long as the input to the model conformed to the specifications needed by MC-Analyst, MC-Analyst could be used to solve best site problems of any discipline.

The weights alterations showed how the final results of the WLC model were strongly affected by the different weights assigned to each criterion and how expected results were output from MC-Analyst.

The study proved how FOSS GIS could be used by programmers who even did not have previous experience using FOSS GIS to develop GIS applications. Motivations behind the wide spread of FOSS GIS were listed. The thesis also addressed the major findings through the coding process such as the problems faced during the development using GeoTools and suggestion on how further work toward stronger integration of MCEA models and GIS could be approached.

2 Open Source, Free Software, and GIS

2.1 Definitions and Standards

The Open Source Initiative defines open source software as "software that can be freely used, changed, and shared (in modified or unmodified form) by anyone. Open source software is made by many people, and distributed under licenses that comply with the Open Source Definition." (Open Source Initiative Homepage, 2013)

Another concept about software is the Free Software. Free Software refers to the freedom of using, redistributing, and modifying of software. While Open Source (OS) refers to a development concept that the source code of the software is left open for reading and modifying, Free Software is a social concept. By combining both of these concepts, a new term emerged, Free and Open Source Software (FOSS). A very important misconception that people have about FOSS is that it must be given for free (without a price). On the contrary, the freedoms of FOSS do not place any restrictions whether or not FOSS could be delivered for free or sold at a price (Steiniger and Bocher, 2009).

FOSS communities bundled efforts to create the Open Source Geospatial Foundation (OSGeo) to support and build the highest quality open source geospatial software (Hall and Leahy, 2008). The OpenGIS Consortium (OGC) was founded in 1994 as a non-profit, international, voluntary consensus standards organization that is specialized in geospatial data and web services. The OGC consists of over 250 organizations from government, academia, industry and other groups (Hall and Leahy, 2008).

OGC has set many standards for data processing and interchanging such as Web Map Service (WMS) and Web Feature Service (WFS) which are the standards for exchanging maps (WMS) and vector data such as features (WFS) over the Internet. A complete list of supported standards could be found in the website of OGC at www.opengeospatial.org.

2.2 FOSS in Business

According to Hall and Leahy (2008), FOSS Business models do not necessarily mean that developers, users or consultants who develop, use or provide services for FOSS work for free. Whereas this is the understanding a large group of audience has about FOSS, this is not true. Free and Open in the term Free and Open Source

Software refer to the freedom of usage, modifying and redistributing the software. Actually, there are many channels from which freelancers, organizations, and companies can benefit and make revenue out of the business of FOSS. Services could be delivered to FOSS users/developers to gain income. For example, not all FOSS software has good documentation or material. Designing good tutorial materials or documentation and conducting special training over such software are among the channels from which income is gained out of FOSS. Providing consultation services for business enterprises that plan to adopt FOSS is another channel of income. Customizing, installation, and optimization all could be assumed as channels of income from FOSS.

2.3 Major FOSS Licenses

There are different licenses under which FOSS is licensed. The most common two types of FOSS licenses are General Public License (GPL) and Lesser General Public License (LGPL). GPL guarantees the freedom of licensees to redistribute, use, and modify the software (source code, object files, link files, and executables) in any way as long as the GPL license is granted on the derivative work or the distributed copies. GPL prevents proprietary software from mixing with code snippets from GPL licensed software which is the main difference between GPL and LGPL. LGPL allows proprietary software to mix with code snippets from LGPL-licensed software and keep their own software licensed under their own original separate license with the exception that any derivative work that is built on top and does not only mix with the LGPL-licensed library must still be LGPL-licensed. A complete list of open source licenses could be found at http://opensource.org/licenses.

2.4 Success Factors of FOSS

In spite of the wide spread of FOSS**,** adopting the usage of FOSS in business, education or any other discipline is not an easy matter specially if there is a transition from proprietary to FOSS software. In order for a FOSS product to be considered successful and competent, a number of criteria should be assessed. Ramsey (2007), a director in Refractions Research Inc., a development organization for FOSS and proprietary GIS software, has addressed these factors:

1. The project documentation: the project must be documented at tutorial level documentation for all user categories (administrators, users, and programmers). The documentation and the source code must be accessible through the web.

2. The development team: the development team members should be defined and contactable through mailing lists. The membership its self should be attainable through a merit-based process.
3. Software modularity: the project must be designed in a modular way that allows its extensibility without re-working of its internals. The extensibility method should be documented.
4. The development community: this refers to the number of core organizations in the development team, whether they are national or international, the funding of the team, and the volume of the developer mailing list.
5. The user community: this refers to the organizations which used the software and what experience they had with it.

2.5 Categories of Open Source GIS Software

There is a considerable work and efforts in developing FOSS GIS. This leads to developing FOSS GIS to meet various needs and applications in different disciplines. Different FOSS statistics literature and studies categorize FOSS GIS in five to seven categories. Steiniger and Hunter (2012b) categorized FOSS GIS as:

1. Desktop GIS: a GIS application that is installed on computers and used mostly for data creation, editing, analysis, map generation, reporting, and charting. Examples: GRASS (Geographic Resources Analysis Support System (grass.osgeo.org)), uDig (a product of Refractions Research, Victoria, British Columbia, Canada (udig.refractions.net)), gvSig (www.gvsig.org).
2. Spatial Database Management Systems (Spatial DBMS): a RDMS with modules that support storage and processing of spatial data to enrich the RDMS usability for spatial business and applications. Examples: PostGIS (www.postgis.net), MySQL with Spatial Extension (a product of Oracle, Redwood Shores, California, USA).
3. Web Map Server: a server application with concentration on serving spatial data such as maps and vector data over the Internet or the intranet through web services. Example: MapServer (www.mapserver.org).
4. Server GIS: While a Web Map Server serves data, a Server GIS serves data and processing services such as analyzing and geocoding. Examples: GeoServer (www.geoserver.org) and 52North WPS (www.52north.org).
5. Web GIS clients: clients that query Web Map Servers or Server GIS to display and process spatial data stored on remote locations that are only accessible via Internet or intranet. Example: OpenLayers (www.openlayers.org).

6. Mobile GIS: GIS applications used on mobile devices. Mostly found on the field for data acquisition and field handling. Examples: gvSIG Mobile (www.gvsig.org) and Quantum GIS (www.qgis.org) for Android.
7. GIS libraries and extensions: software libraries that are used by professional software developers to customize existing applications or build new applications for the geospatial business. Examples: GDAL (www.gdal.org) and GeoTools (www.geotools.org).

2.6 FOSS GIS Popularity, Combination, Motivations, and Limitations

FOSS GIS is becoming more popular and widely used. It is transiting from "early adopters" to "early majority" stage because of its capability of competing with proprietary software (Steiniger and Hunter, 2012b). Indicators for this are the increasing download rates, increasing participants at local FOSS4G conferences, and FOSS4G consulting companies with full order books (Steiniger and Hunter, 2012b). A couple of indicators referenced by Steiniger and Bocher (2009) included the number of FOSS projects, the number of governmentally funded projects, the number of downloads, and the increasing number of use cases of open source GIS software.

In 2006-2008, 20 entries have been added to the list of software projects in the website FreeGIS.org, which created 330 entries for software alone in 2008. On the date of writing this thesis, the website referenced 356 entries for software. Another indicator is the number of governmentally funded FOSS desktop GIS projects (4 out of 10 listed by Steiniger and Bocher (2009)).

It has even gone further that some governments have decided to replace their proprietary software with FOSS software. For example, gvSIG, which stands for "generalitat valenciana, Sistema d’Informacio´ Geogra`fica", is a FOSS desktop GIS project founded by the Regional Council for Infrastructures and Transportation (CIT) of Valencia (Spain) to replace ESRI’s ArcView, a proprietary software of similar functionality, in municipal authorities which is a step to switch all systems to a Linux-based computer infrastructure (Steiniger and Bocher, 2009).

Since there is no single FOSS GIS software that meets all the needs of a GIS solution, multiple components or projects of FOSS GIS integrate into each other to build the final solution. One of the major methodologies that helped this is the adherence of FOSS GIS to OGC standards. This interoperability could integrate software components that are as basic as primary software development components up to complete standalone products such as desktop GIS and GIS server software. Steiniger and Hunter (2012b) issued a study which revealed the usage of different FOSS GIS software such as desktop GIS, web map servers, spatial database management systems, and web map development toolkits to build a Spatial Data Infrastructure (SDI). This implementation was possible with limited financial budget and without worrying about legal constraints that proprietary software vendors place on their products. Combining different FOSS GIS (gvSig, Sextante (a free spatial data processing framework licensed under GPL2. (www.sextantegis.com)), GeoServer, and PostGIS) has also helped the local government in the city of TREBINJE - Bosnia and Herzegovina build their own decision-making system concerning improving the public lighting system (Jankovic and Milidragovic, 2013).

Transition from proprietary to FOSS GIS software might not always be associated with no costs. In some cases, there are costs for users training, installation, configuration, consultation, and customized prepared documentation (Steiniger and Bocher, 2009). Another reason why customers do not adopt FOSS GIS is that most of FOSS GIS is specialized towards some discipline, whereas commercial GIS software offers products for wider customer categories (Jankovic and Milidragovic, 2013). FOSS GIS itself has barriers that obstacle its development such as knowledge barriers, legal integration, forking, sunk costs, and technology immaturity (Nagy and Yassin, 2010). Steiniger and Hunter (2012b) explained why these barriers exist. Knowledge barriers exist when potential users lack the awareness of the availability or capabilities of FOSS GIS and its relevance to their needs. Forking occurs when an immature FOSS project branches out to smaller projects and the whole development stops because of the lack of integration between those smaller projects. Sunk costs emerge because of the prior investment in proprietary software.

2.7 FOSS GIS Software Libraries and GeoTools

Software libraries are the core building blocks for software. They are the product of applying software engineering methods that allow the combination and sharing of efforts and works from different development teams to enhance software interoperability, modularity, and reusability. There are hundreds of software libraries released on the Internet. While some software libraries did not find considerable interest by the community, other libraries have become cores to build other GIS servers, desktop GIS or libraries software. OSGeo adopts the most widely used and standard FOSS GIS. All of FOSS GIS projects that have major community are listed in OSGeo website (www.osgeo.org).

Among the geospatial libraries adopted by OSGeo is GeoTools. A description of GeoTools found in OSGeo website is "an open source Java GIS toolkit providing implementations of many Open Geospatial Consortium (OGC) specifications as they are developed. GeoTools is also associated with the GeoAPI project that creates geospatial Java interfaces." (OSGeo, 2013).

For its topology modelling, GeoTools uses the popular topology library of Java Topology Suite (JTS), a leading Java (www.java.com) geometry open source library (www.geotools.org). GeoTools is also used as the basic component to build GeoServer, an OSGeo sponsored GIS Server FOSS project and the popular desktop GIS application of uDig (www.geotools.org; and Turton, 2008).

More mature FOSS GIS projects are developed using C instead of Java which itself used to develop mature projects more than .NET (Microsoft .NET framework for software development) (Ramsey, 2007). This is because of the history that C++ and Java platforms have before the release of .NET. However, C++ has the lion's share of the developed FOSS GIS mature projects.

Steiniger and Hunter (2012b) considered GeoTools as:

1. Data input/output and conversion library.
2. Geometry Library: through utilizing JTS.
3. Projection library.
4. General Framework that provides other different functionalities.

For GIS libraries, there are indicators that GeoTools is becoming more popular, widely used, and successful software. Different releases of GeoTools 2.2.2 and 2.3.5 had 1545 and 4378 downloads in 2007 and 2008 respectively (www.geotools.org). GeoTools 10.0 released on 20-09-2013 had 3720 downloads in the period between Sep 2013 and Apr 2014 (www.geotools.org). Over the years, GeoTools has grown up from being a software library developed by one developer to a multinational project developed by more than 20 developers around the world (Turton, 2008). Among the indicators of GeoTools popularity is that it is used by three large open source projects: GeoServer, uDig, and GeoVISTA studio (an open software development environment designed by Pennsylvania State University for geospatial data (www.geovistastudio.psu.edu)) (Turton, 2008). GeoTools is considered a good start especially for students who would like to gain how GIS operations work behind the user interface without worrying about the restrictions of copyright and software patents (Turton, 2008). Advanced students can experiment the different methods of implementation to see the practical advantages and disadvantages of using GeoTools. This provides a true education that will help advance careers as well (Turton, 2008).

3 Multi-Criteria Evaluation Analysis in GIS Software

GIS functions are divided into two categories: fundamental and advanced (Malczewski, 2010). Fundamental functions such as measurement, reclassification, scalar and overlay operations, neighbourhood operations, and connectivity operations are available in most GIS software such as ArcGIS, IDRISI (a product of Clark Labs of Clark University, Massachusetts, USA), and GRASS (a FOSS GIS project (grass.osgeo.org)).

According to Malczewski (2010), most GIS software have limited capabilities to perform MCDA except for IDRISI, Common GIS (a product of Fraunhofer Institute Intelligent Analysis and Information Systems (IAIS), Sankt Augustin Germany), ILWIS (a project of World Institute for Conservation & Environment (www.ilwis.org)), and TNT-GIS (a product of MicroImages, Inc, Nebraska, USA). Thus, more advanced functions such as statistical and mathematical operations are needed in GIS software to provide better implementation for decision-making problems. Among IDRISI, Common GIS, ILWIS, and TNT-GIS, only ILWIS is an open source software distributed under GPL license. However, simple MCDA models such as the weighted summation model still could be built using the fundamental GIS functions such as map algebra operations (Malczewski, 2010).

Even though GIS and MCDA are two independent systems, Malczewski (2010) defined three components of a MC-SDSS which is the integration between GIS and MCDA systems. The components are: 1) Geographical data management and analysis toolbox which contains the tools of GIS, 2) MCDA toolbox which contains the tools for: a) Generating value structure, preferences modelling, and multi-criteria decision rules, and b) Performing sensitivity analysis, and 3) User interface which includes all the mechanisms by which commands, requests, and data are entered into the MC-SDSS.

This integration of GIS and MCDA systems could be classified into three categories based on the level of coupling between the two systems: 1) Loose-coupling, 2) Tight-coupling, and 3) Full integration (Malczewski, 2010). In loose-coupling, both GIS and MCDA have separate interfaces and the two systems communicate data through files. In tight-coupling, both systems share a common interface which contains a single model. Finally, the full-integration approach in which users develop their own routines in a generic programming language and those routines are then added to the GIS package in the system.

Most of MCDA techniques could be implemented in most GIS packages without custom programming (Malczewski, 1999). There are some implementations of MCDA models such as the "weighted overlay" tool in ArcGIS ArcMap. In addition, there are many free and commercial Add-On tools that could be installed into ArcGIS ArcMap that implement other MCDA techniques (Greene, 2011). IDRISI and CommonGIS, on the other hand, are the only GIS packages that have full integration with MCDA and implement different techniques of MCDA (Greene, 2011). All of ArcGIS, IDRISI, and CommonGIS are proprietary software which need paid licenses for use. CommonGIS has a research and educational purposes license but it is not by any means a FOSS license (www.iais.fraunhofer.de).

Other tools have been developed to work under proprietary software such as the free and open source tool MCDA4ArcMap which was developed by Rinner and Voss (2013) and the multi-criteria evaluation site selection tool developed by Eldrandaly (2011). Both of these tools are developed as Add-On to ArcMap. Thus, an installed version of ArcMap is needed before these tools could be used.

4 Methodology

4.1 Theoretical Concept

Weighted Linear Combination (WLC) method has four major phases: 1) Criteria definition, 2) Values normalization of criteria layers (commensurate value attributes), 3) Weights definition, and 4) Weighted layers combination (weighted overlaying).

In criteria definition phase, the criteria involved in the problem are addressed and data are collected. In values normalization phase, the raw values of criteria layers are transferred to the range of [0-1]. In weights definition phase, a weight is assigned for every criterion according to the criterion importance and influence on the final objective. In the weighted layers combination phase, a sum-overlay is performed on all the weighted criteria layers to produce the final suitability result.

Before proceeding in applying the four major phases of WLC, the study area boundaries, the spatial resolution of the model, and the reserved areas should be determined.

4.1.1 Criteria Definition

To decide whether or not a specific area of land is suitable for some facility or activity, criteria that help decision makers must be addressed. These criteria (attributes) that are represented as thematic GIS layers are supposed to meet some requirements. According to Malczewski (2000), the criteria have to be comprehensive, measureable, complete, non-redundant, and minimal. A comprehensive attribute is an attribute that indicates the achievement of an objective that is related to the decision problem. A measureable attribute is an attribute that can be evaluated using a number that measures the preferences of the decision maker in the problem. Attributes are considered complete if they cover all aspects of the problem. Non-redundant attributes are attributes that are independent of each other in such a way that every single attribute represents one aspect of the problem. The minimal attributes concept refers to that attribute should be as few as possible but without affecting the attributes being "comprehensive".

Not all Multi-Criteria Evaluation Analysis (MCEA) studies that investigate the same problem share the same set of criteria. This is normal because criteria selection is based on different factors such as the objective of the study, the experts' analysis and expertise, and availability of the data. For example, a solar farm that is connected to a grid must be close enough to transmission lines, while solar farms that are built to supply a plant must be close enough to that plant.

4.1.2 Values Normalization of Criteria Layers

Criteria layers represent different attributes that have different ranges or scales of values. WLC requires that values contained in different attribute layers to be transformed to comparable units (Malczewski, 2000). In order to incorporate these layers into a GIS-MADA model, all layers must have the same unit; a range of values from [0-1]. This is performed through normalization.

There are many approaches for attributes' values normalization. Normalization transfers raw attribute values from the raw scale to [0-1] scale. The most common method is Score Range Procedure (SRP) (Malczewski, 2000). In SRP, the minimum raw value is subtracted from all values in the attribute and then scaled so that 0 corresponds to the minimum raw value and 1 corresponds to the maximum raw value. SRP formula is showed in Eq. 1.

SRP Normalize Value = (Raw Value - Minimum Raw Value)/(Maximum Raw Value - Minimum Raw Value) (Eq. 1)

SRP method of linearity might not be suitable for all attributes. One example where SRP might not meet the need is the opposite distance analogy that considers "closer is better" which means that the minimum value in the raw attribute values should be represented as 1 and the maximum value should be represented as 0 on the normalization scale. In addition, there are other cases in which multi scale or qualitative normalization is required. This occurs where linearity does not model the decision maker preferences toward the attribute. An example on this in finding the best location for a solar farm problem is the preferences of decision makers to place a solar farm on a distance that is not too close to and not too far from cities. In this case, values that exist on the middle of the numerical range of the distance attribute are assigned 1. For attributes with non-numeric values, such as land cover, SRP does not work either. In such cases, the use of value functions is better (Hepner, 1984). Such functions must be manually defined by the user because every single raw value or set of values from the attribute must be considered. Thus, there is no single representation for a normalization function. In MC-Analyst's implementation, the user is allowed to choose among different kinds of normalization methods.

A sensitive step before applying the normalization procedure on the attributes is to drop areas that are excluded from the selection problem. Usually those areas are called constraints areas. For example archaeological sites if there is an attribute about archaeology incorporated in the study. This exclusion of areas will remove the raw attribute values of these areas that otherwise could have affected the normalization range if they had been included (Malczewski, 2000).

4.1.3. Defining Weights

Since weights reflect the decision maker preference toward the criteria, weights definition is the most important component of the MCDA model. There are several ways to create the weights of attributes such as pair-comparison (a method used in the Analytical Hierarchical Process (AHP)) and ranking methods. Weights in the WLC could range from [0-1] or [0-100]. A common practice is to directly assign the weights to the attributes. Direct assigning approach of weights definition is criticized because it ignores the change in attribute values (Malczewski, 2000). It completely ignores the difference between the minimum value and the maximum value of the attribute. Consider for example an area with topography that changes from 2500-2510m. This attribute may represent a considerably high but not changing topography (only 10m difference between minimum and maximum topography values). In this example, the topography is stable over the whole area which means that it does not make big difference in the decision. While another area with topography that changes from 1000-1250m (250m change) should be more considerable by decision makers and is more affecting on the result.

A method that concerns about the amount of change in attribute values is the Swing Weights Technique (SWT) (Malczewski, 2000). The method starts by listing the attributes and asking the decision maker to maximize only one attribute to its optimal value. This attribute is now considered the highest priority or first most important attribute. Then, the decision maker is asked to maximize another attribute to its optimal value. This attribute is now considered the second highest priority or second most important on the list of attributes. Following the same approach, all attributes are ordered by their importance or priority in descending order. Now, after ordering all the attributes, the first most important attribute is assigned P_1, a priority value of 100, given that the attribute raw values range from A_{1-min} to A_{1-max} (the minimum and the maximum raw values of the first most important attribute respectively). Then, the decision maker rates P_2, the priority of the second most important attribute, given that the attribute raw values range from A_{2-min} to A_{2-max} (the minimum and the maximum raw values of the second most important attribute respectively). Using the same approach, all less important attributes are assigned their priorities after taking into consideration the change in their raw values from minimum to maximum. It is important to make sure that only the first attribute is allowed to have 100 as priority. After assigning all the priorities, every attribute's weight is calculated by dividing the attribute's priority by the sum of all priorities.

An example of three attributes with priorities of $P_1 = 100$, $P_2 = 60$, and $P_3 = 30$, the first, second, and third attributes are given the final weights of W_1, W_2, and W_3 that are calculated as:

$$W_1 = 100/ (100+60+30) = 0.523 \quad \text{(Eq. 2)}$$

$$W_2 = 60/(100+60+30) = 0.315 \quad \text{(Eq. 3)}$$

$$W_3 = 30/(100+60+30) = 0.157 \quad \text{(Eq. 4)}$$

Those final weights will be used as multiplication factors for the criteria when applying the final step of "combination".

4.1.4. Weighted Layers Combination

The combination of attributes is achieved through sum-overlaying. Since the attributes represent phenomena that should be comprehensive and complete all over the study area, every attribute should have a value over the whole study area. This is the reason why most of MCDA problems use raster data as input to the model. In addition, the combination of attribute layers is performed using GIS overlay operation. After normalization and weights definition are performed, every attribute layer is multiplied by its corresponding weight and the multiplied layers are sum-overlaid. The combination is given by:

$$\text{Score} = \sum_{i=1}^{n} W_i * X_i \quad \text{(Eq. 5)}$$

where *Score* is the score for the area or cell, *n* is the number of attributes or criteria, W_i is the weight assigned to the *ith* attribute, and X_i is the normalized value of the *ith* attribute. The result of the combination is the suitability raster whose cells' values range from [0-1].

Fig. 4.1 shows a detailed flow chart of the steps that were performed in the methodology. Even though the flow chart shows more steps than explained in the methodology, only milestone steps of the methodology were stated.

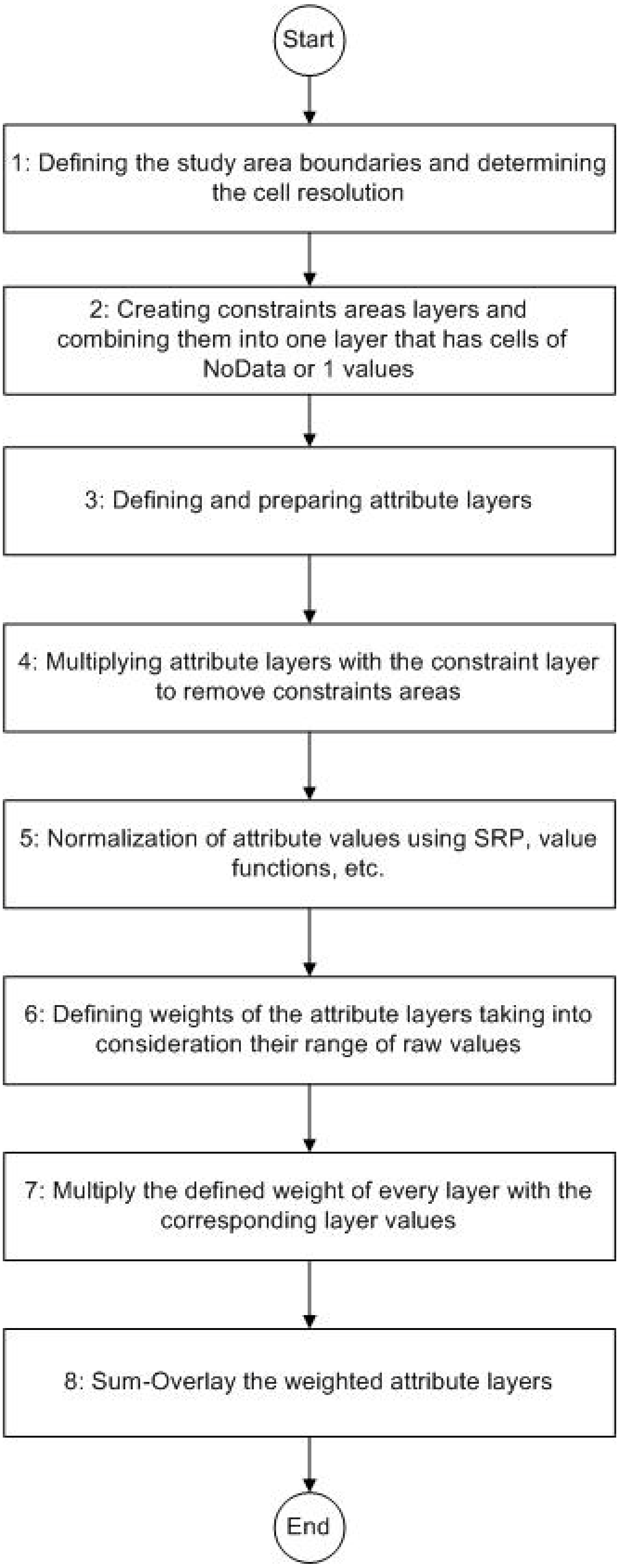


Fig. 4.1: The main flow chart of the Weighted Linear Combination (WLC) methodology (author's figure).

4.2 MC-Analyst

Malczewski (2000) defined the WLC procedure through six steps: 1) Defining the set of attributes (criteria), 2) Defining the set of feasible alternatives, 3) Generating commensurate attribute maps, 4) Assigning attribute weights, 5) Combining attribute maps and weights, and 6) Ranking the grid cells. Even though Fig. 4.1 showed more than six steps but the additional steps are tasks that are performed behind the scene in order to implement the WLC procedure properly. Actually, the steps in Fig. 4.1 are addressed in Malczewski (2000) in between the major aforementioned six steps.

MC-Analyst implemented the WLC procedure from steps 3 through 5. Before moving to the WLC procedure, two wizards are demonstrated: 1) Buffering and Rasterization and 2) Reclassification.

4.2.1 Business Requirements

The resource of requirements collection was the outlines addressed by Malczewski (2000). Malczewski (2000) methodology requires general GIS processing tasks such as raster resampling and clipping. These were not considered as must-have requirements of MC-Analyst since the scope of this study is the unique operations carried out in the MCEA model itself such as SCP normalization, weights definition, and rasters combination.

4.2.1.1 Reserved (Constraints) Areas Layer

In every problem of evaluating suitability across an area, portions of that area could be excluded out for some reasons such as being these areas wild life reserves or federal lands. These areas are combined into one layer, the constraints layer. The constraints layer will work as a filter by which all criteria layers are multiplied. This multiplication will exclude the reserved areas out of all of the criteria layers.

4.2.1.2 Normalization of Cell Values of Criteria Raster Layers

In order for the criteria layers to be compared and summed, cell values must be normalized (transferred) into the range of [0-1] where 0 corresponds to no suitability and 1 corresponds to best suitability. This normalization is carried out using SRP and value functions.

4.2.1.3 Weights Definition

MC-Analyst must allow the user to define the weights according to the SWT. MC-Analyst must show the range of raw values of every criterion raster at this stage in order to keep the user aware of the raw values range before the normalization phase. Only one criterion is allowed to have the priority of 100. Final weights are derived from the priorities as in the example equations Eq. 2-4 in section 4.1.3. Defining Weights.

4.2.1.4 Weights and Criteria Rasters Multiplication

For every defined weight, MC-Analyst must multiply that weight with the assigned criterion raster. The result is weighted criterion raster layers.

4.2.1.5 Weighted Criteria Raster Combination

In the last phase, MC-Analyst must sum-overlay all the previously weighted criteria rasters resulting out the final suitability raster with its cell values ranging between [0-1].

4.3 Menus of MC-Analyst

4.3.1 Introduction

Since WLC resolves spatial suitability, spatial proximity and reclassification of spatial data are frequently applied in WLC problems. Two wizards were developed for this purpose: 1) Multi-Buffering and Rasterization Wizard and 2) Reclassification Wizard.

4.3.2 Multi-Buffering and Rasterization

This wizard is accessible through the menu "Data" => "Buffer...". The wizard should look like Fig. 4.2. This wizard takes as an input many parameters. The primary inputs are a shape file and buffering levels. The output of this wizard is a raster whose cells are the buffering distances provided by the user and 0 elsewhere.

Before using this wizard, the shape file should be under a projected coordinate system; this is because the wizard considers that the buffering distances use the projected coordinate.

The fields of this wizard are:

1. Input file: the field where the input file full path and name appears. In order to select an input file, the user can click the "open" button next to this field and choose the shape file.
2. Output folder: this field allows the user to choose the output folder where the output file will be saved. The output file will have the same name of the input file followed by the text "-buffered".
3. Cell resolution: the spatial resolution of the result raster using the unit of the projected coordinate system of the input shape file.
4. Containing Extent Layer (Optional): this field could be left empty but if it is assigned a raster/shape file, then the output raster will have the same extent of this extent layer file. The maximum buffering distance entered should be inside the Containing Extent Layer otherwise the wizard would not work.
5. Extent Layer SRS (Optional): This is to choose among the Spatial Reference System under which the Containing Extent Layer is referenced (a requirement when the Containing Extent Layer is a raster layer).
6. Distances: a table where the user adds and removes the buffering distances by clicking the button "Add Buffering Distance" and "Remove Buffering Distance", respectively.
7. Buffer: the button to start the buffering process.

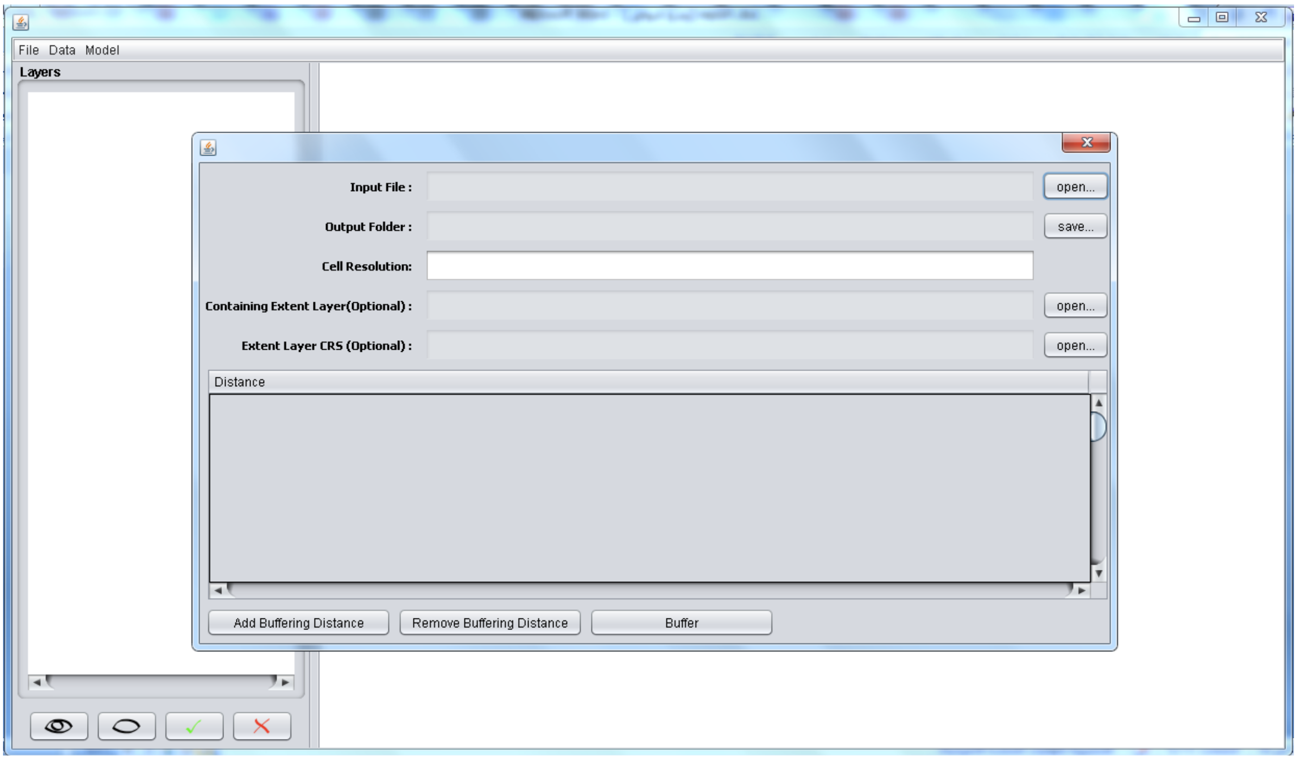


Fig. 4.2: Multi-Buffering and Rasterization Wizard

Multi-Buffering and Rasterization procedure works as follows:

1. Buffering distances are sorted in descending order.
2. A shape file of the largest buffer is created.
3. A raster with the largest buffering extent is created along with the entered raster spatial resolution. If an extent layer (raster or shape file) is chosen then the created raster borrows the extent of the extent layer instead of the extent of the maximum buffered features collection.
4. Rasterize the shape file of the largest buffer on the previously created raster. This raster is called the result raster.
5. For all remaining buffering distances:
    a. Buffer the features up to the current buffering distance.
    b. Rasterize the newly created buffered features. This raster is called the buffered raster.
    c. Now the buffered raster has cells with either the value of the buffering distance or 0. The cells with values of "buffering distance" are changed to 0 and all other cells are changed to 1. This raster is called "the mask raster".
    d. Now, the result raster is multiplied with mask raster. The output is the result raster with the value of 0 for the cells that lie inside the current buffering distance. This multiplication of the result raster with the mask raster resets the inner areas of the overlapped buffered features to 0 so that when the buffered rasters are sum-overlaid, cell values are not accumulated by the buffering distances, instead, cells will have values of the minimum buffering value of the overlapped buffered features over the cells.
    e. Now, the masked result raster and the buffered raster are sum-overlaid.
    f. Move to process the next smaller buffering distance.

The Rasterization procedure works as follows:

1. The spatial extent of the buffered shape file is read and a corresponding raster is created and referenced to that extent with the defined cell resolution. The number of rows and columns in the created raster is determined by:

$$\text{Rows Count} = (\text{Vertical Extent Length}) / \text{Cell Height} \quad \text{(Eq. 6)}$$

$$\text{Columns Count} = (\text{Horizontal Extent Width}) / \text{Cell Width} \quad \text{(Eq. 7)}$$

where *Vertical Extent Length* is the absolute difference between top and low boundaries of the projected coordinates system of the original shape file. *Horizontal Extent Width* is the absolute difference between left and right boundaries of the projected coordinates of the original shape file.

2. All of the cells in the raster are assigned the value of 0.
3. Now for every feature in the shape file :
   a. The attribute to be rasterized is read and assigned to the colour of the graphics object that is writing the feature.
   b. The coordinates of the feature are read.
   c. A feature is drawn on the raster using the previously read coordinates.

It is obvious that overlapped feature resolution approach is "last". It means that if there are any overlapped features, then the last one processed will have its value written in the mutual cells.

This rasterization approach is ideal for points, lines and non-overlapped polygons feature rasterization.

A flow chart of Multi-Buffering and Rasterization process is shown in Fig. 4.3. Fig. 4.4 shows an example of the primary roads in Colorado, USA mapped and buffered at 10 km and 30 km buffer distances using this wizard.

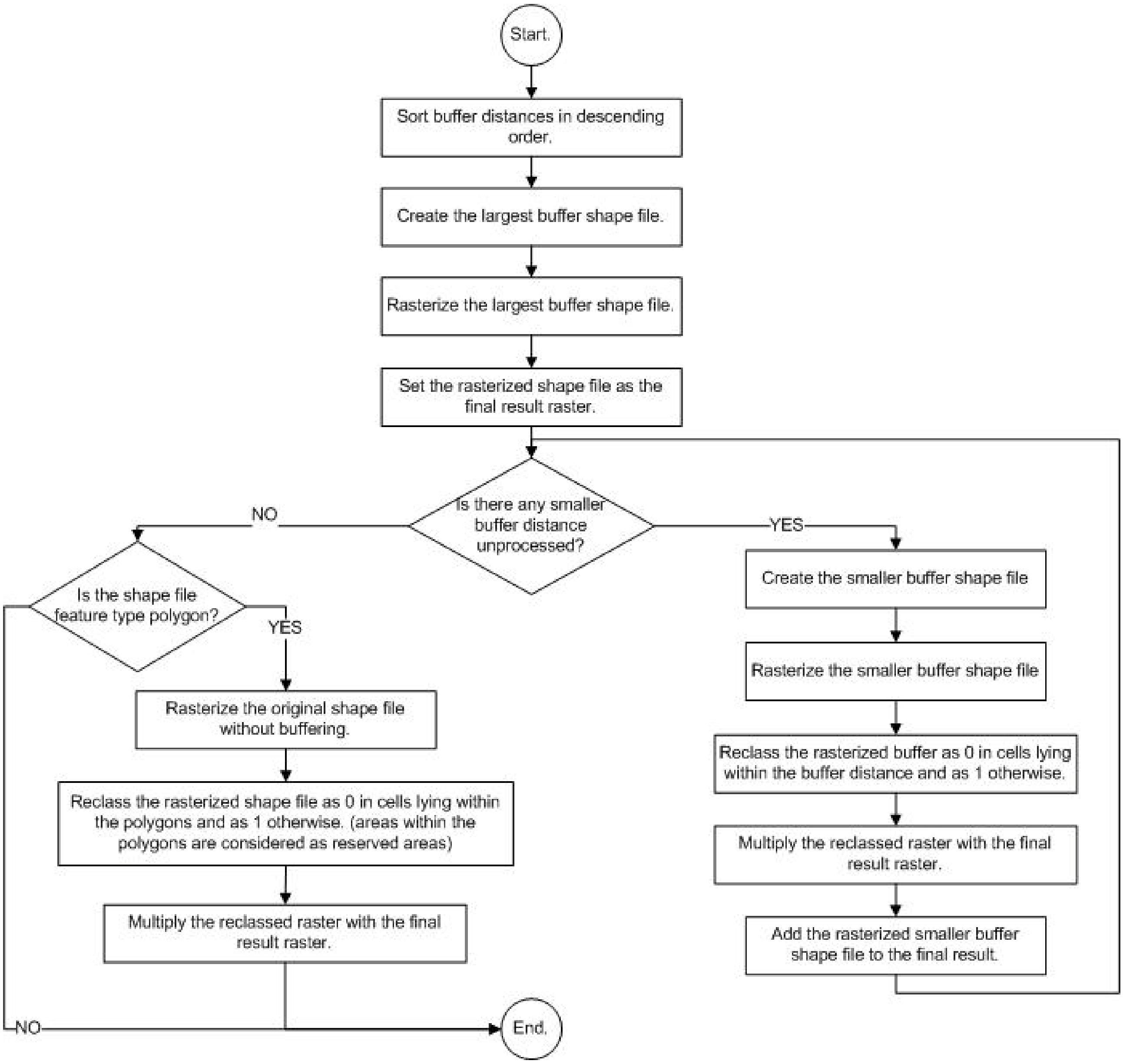


Fig. 4.3: Multi-buffering and rasterization process flow chart.

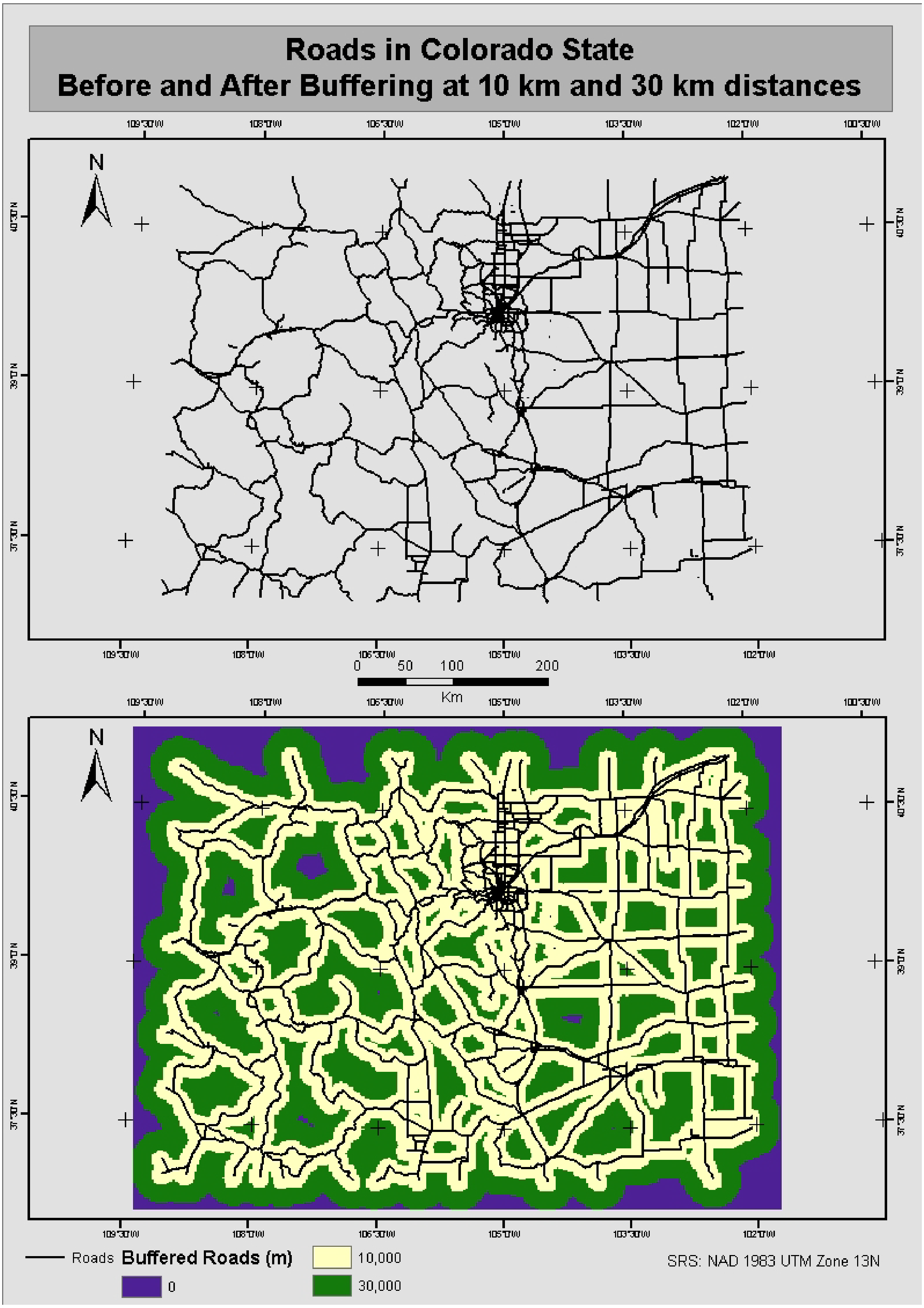
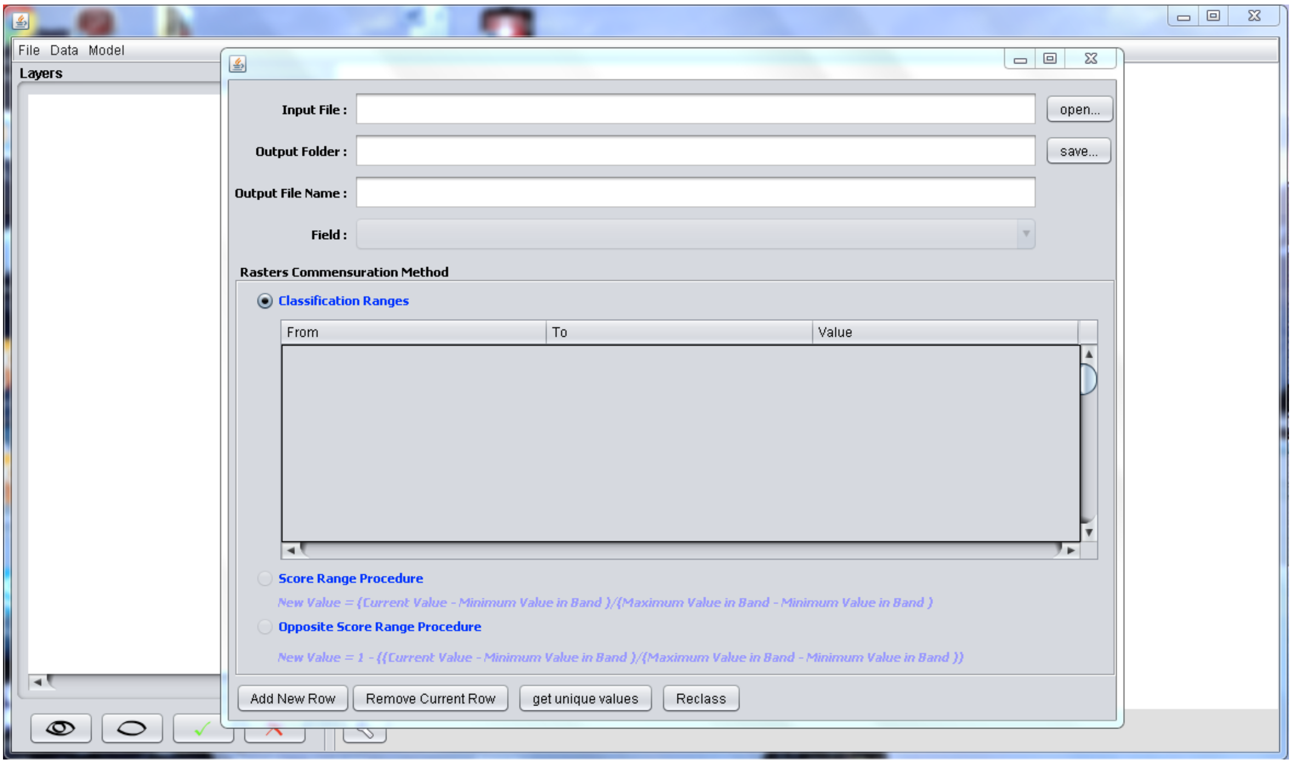


Fig. 4.4: Primary roads in Colorado, USA. Before and after buffering at 10 km and 30 km distances

4.3.3 Reclassification

Reclassification task converts raster cell values from one range to another. It maps ranges of values to one value. (reclassification that uses SRP is also available only when used from inside the GIS/MCE WLC model as demonstrated later).



New Value = {Current Value - Minimum Value in Band }/{Maximum Value in Band - Minimum Value in Band }

New Value = 1 - {{Current Value - Minimum Value in Band }/{Maximum Value in Band - Minimum Value in Band }}

Fig. 4.5: Reclassification Wizard

Fig. 4.5 shows the reclassification wizard. As the figure shows, reclassification has a set of input fields which are:

1. Input file: the input file. It could be a shape or a raster file.
2. Output folder: the output folder where the output reclassified file will be saved.
3. Output file name and extension: used when the input is a raster file. It is the name of the output file. ".tif" should be added to the suffix.
4. Field: active only with shape files. This is a list from which the user chooses the field to get the values upon which to reclassify the shape file features. It shows only fields of type numeric and string.
5. Raster commensurate methods: reclassification methods offered by the wizard. When the wizard is run as a standalone, it offers only one commensurate method, "The Classification Ranges", which is a mapping between ranges or unique values from the source to unique values on the

output raster. The other two methods are the Score Range Procedure and its opposite. The score range procedure and its opposite are given by:

$$\text{New Value} = (\text{Current Value} - \text{Minimum Value in Band})/(\text{Maximum Value in Band} - \text{Minimum Value in Band}) \qquad (\text{Eq. } 8)$$

$$\text{New Value} = 1 - ((\text{Current Value} - \text{Minimum Value in Band})/(\text{Maximum Value in Band} - \text{Minimum Value in Band})) \qquad (\text{Eq. } 9)$$

where *New Value* is the new value of the cell in the output reclassified raster, *Current Value* is the value of the same cell in the source layer, *Minimum Value in Band* is the minimum value in band 0 of the raster, and *Maximum Value in Band* is the maximum value in band 0 of the raster.

6. Add New Row button: when the user clicks this button a new row for range classification is added to the classification table.
7. Remove Current Row button: when the user clicks this button, the current selected row is removed from the classification table.
8. Get Unique Values button: when the user click this button, the classification table is populated with the unique values from the input file. If the number of unique values exceeds 30 values, then the table is populated with 5 equally divided ranges.
9. Reclass button: when the user clicks this button, the reclassification process starts.

Fig. 4.6 shows the flow chart of the reclassification wizard. Fig. 4.7 shows a map of concentrated solar power in Colorado that has been reclassified using SRP normalization method. The reclassification result, which is a .tif file, is one of the files that the WLC model outputs when run. Fig. 5.2 shows the map of concentrated solar power in Colorado before SRP reclassification.

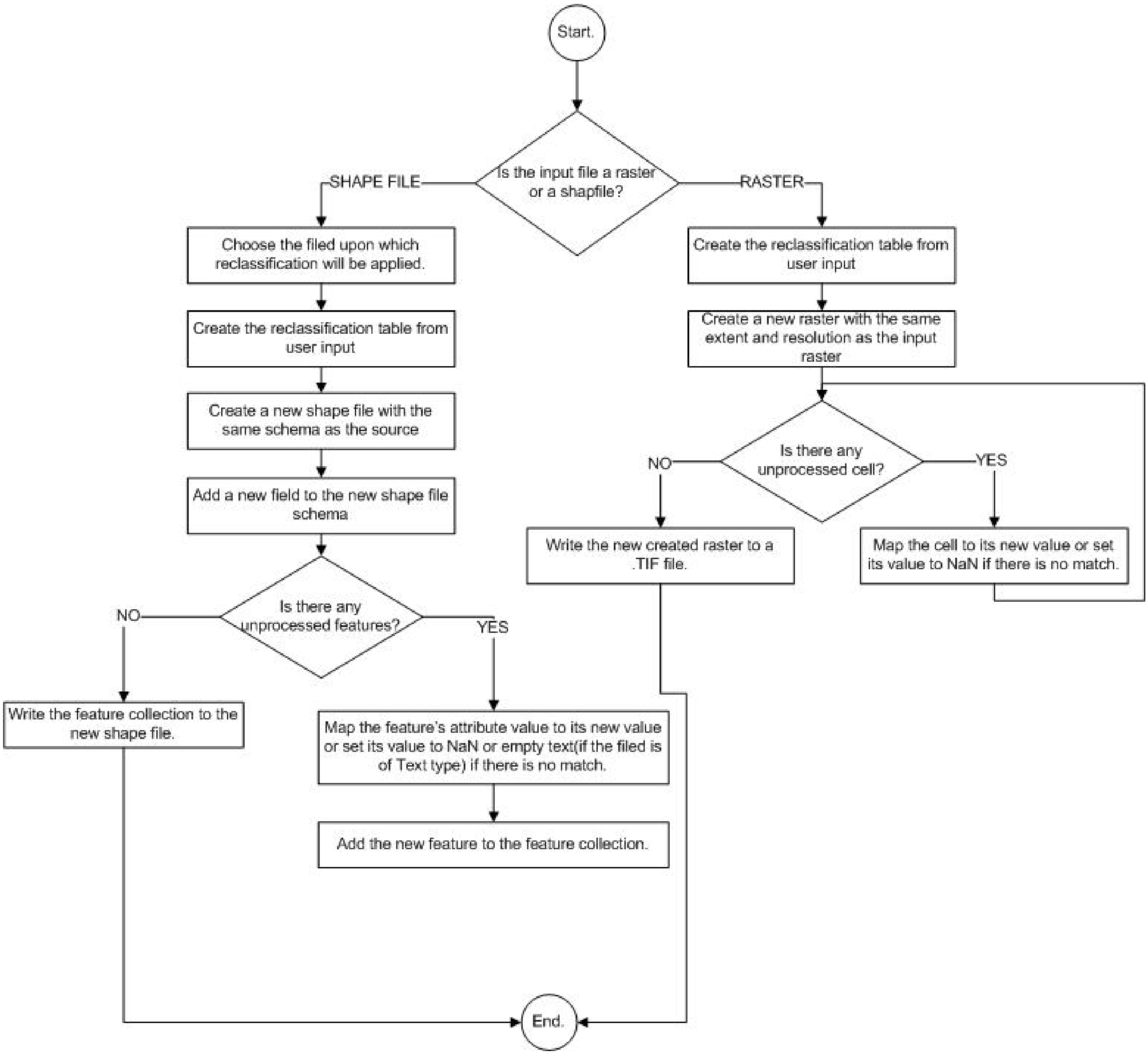


Fig. 4.6: Flow chart of the reclassification process.

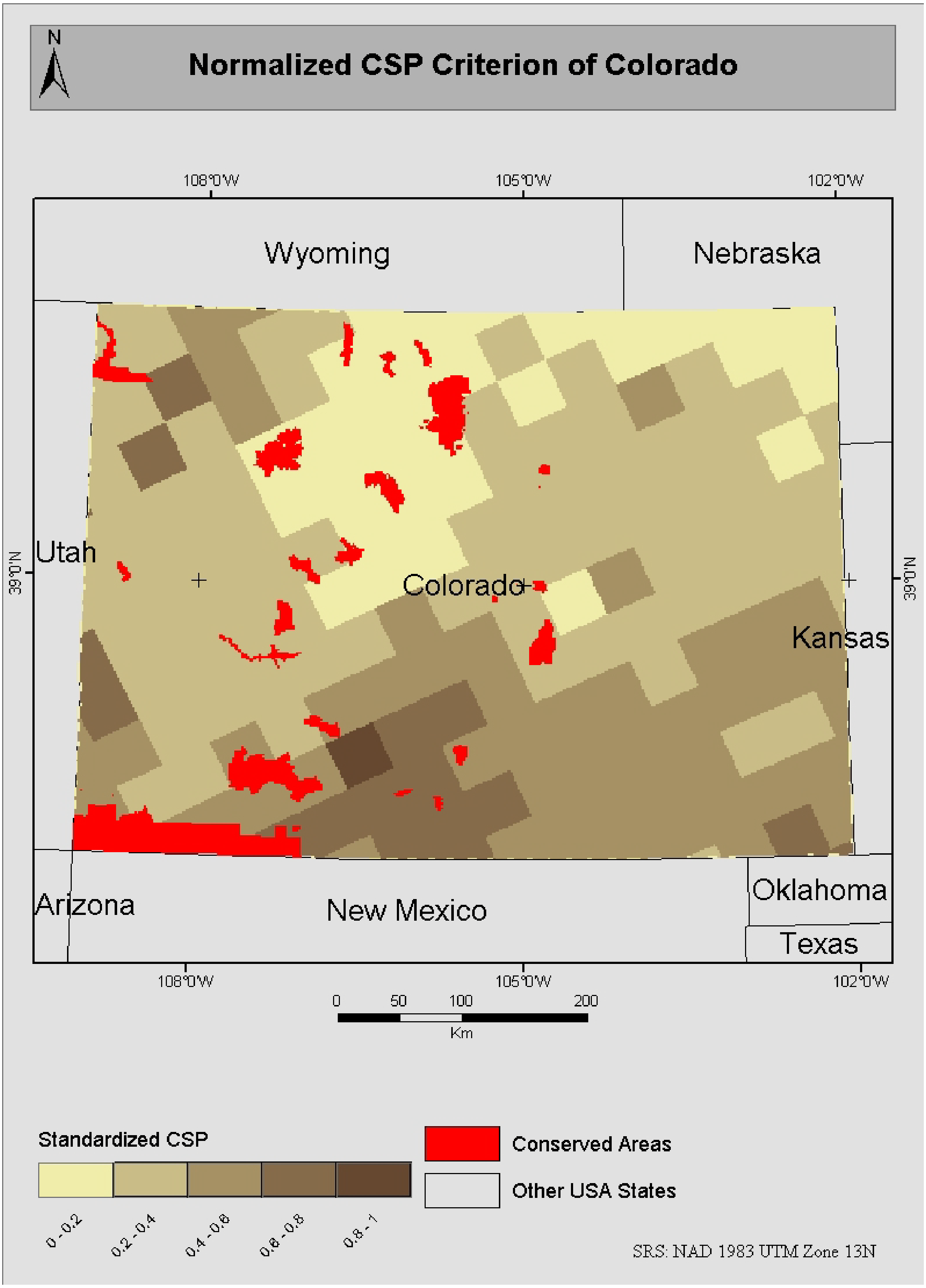


Fig. 4.7: Reclassified raster of Concentrated Solar Power (CSP) in Colorado using Score Range Procedure normalization method.

4.3.4 The WLC Wizard

The WLC wizard consists of four steps: 1) Selecting the global variable of the model such as the constraints raster and the spatial reference system (SRS), 2) Selecting the criteria rasters, 3) Reclassification of the criteria rasters, and 4) Defining the weights.

4.3.4.1 The Global Variables

This step receives the following input:

1. The cell resolution of the model: the spatial cell resolution at which the constraints, the criteria, and the output rasters will be resampled. The resampling method is "nearest neighbour". A pre-assumption is that the SRS of all the layers of the model is a projected coordinate system.
2. Model output folder: the folder in which all output rasters created during the processing are stored.
3. Coordinate Reference System: the SRS of the used criteria rasters.
4. Constraints Map: the raster used as a constraints layer for the model. Its cells values are either 1 or NaN (Not a Number).

Fig. 4.8 shows the screen of the global variables.

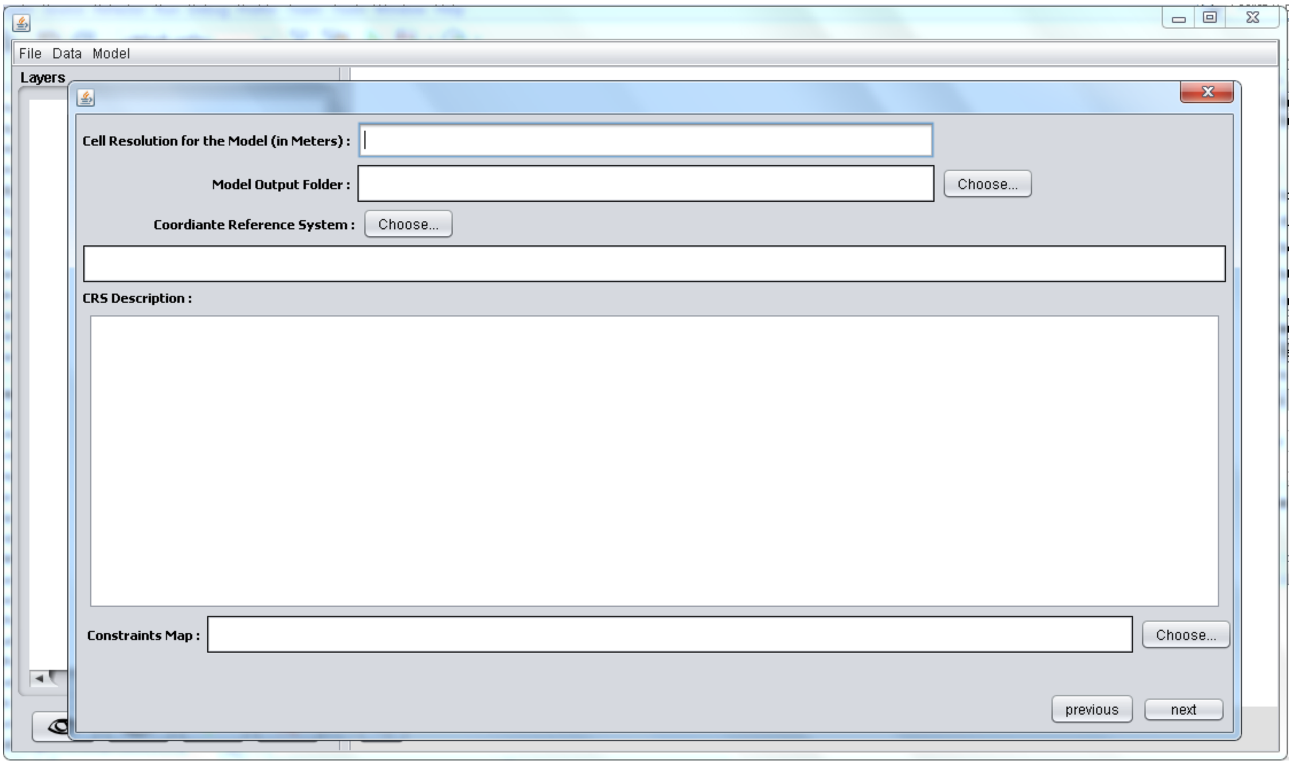


Fig. 4.8: The first step of the WLC wizard - The global variables.

4.3.4.2 Selecting the Criteria Rasters

From this window, all criteria layers are added. After the user clicks next on this step, the first processing step starts as the following: 1) The constraints raster is resampled using the nearest neighbour method at the resolution entered in the previous step, 2) The constraints raster "NoData" value is changed into "NaN" (Not a Number), 3) The criteria rasters are resampled as well, 4) The criteria rasters "NoData" value is changed into "NaN", 5) The criteria rasters are multiplied by the constraints rasters and the multiplication results are written.

Fig. 4.9 shows the flow chart of this process. Fig. 4.10 shows this screen during processing the case study in this thesis.

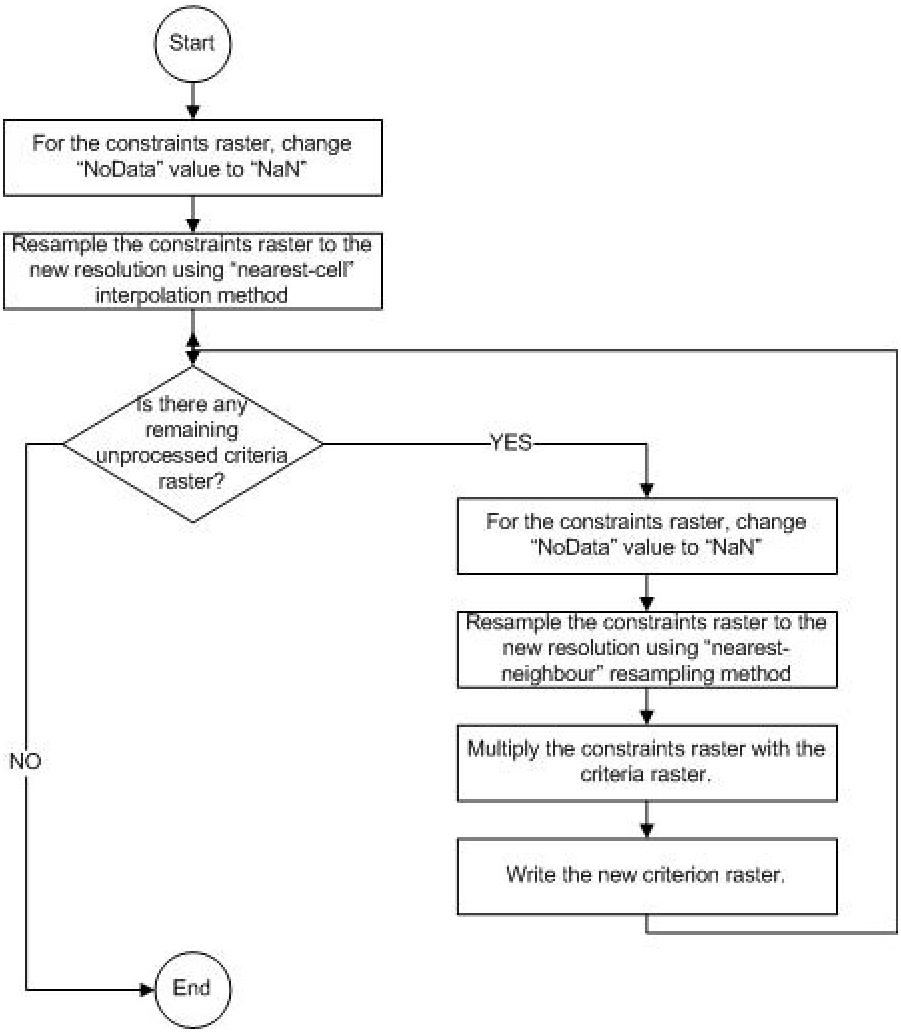
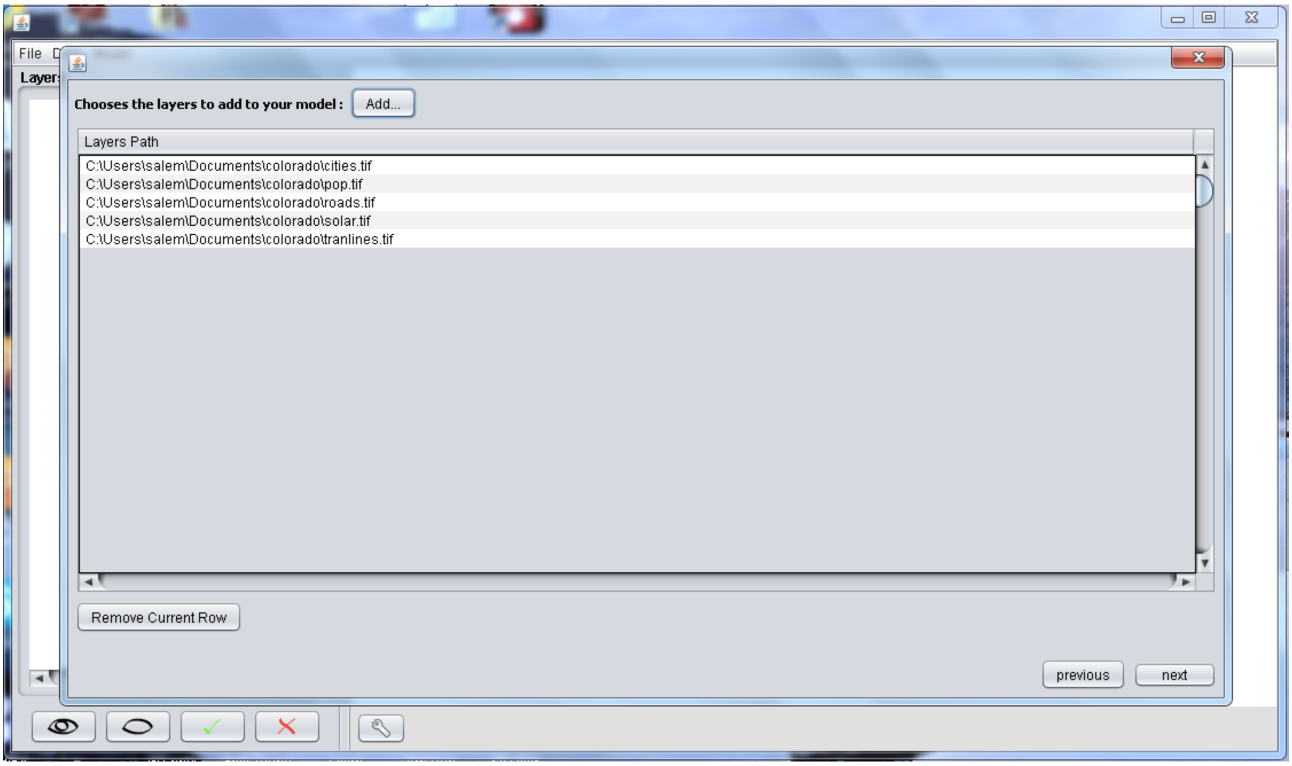


Fig. 4.9: The flowchart of resampling the constraints raster and the criteria rasters and the multiplication of criteria rasters by the constraints raster.



Chooses the layers to add to your model :

Layers Path
C:\Users\salem\Documents\colorado\cities.tif
C:\Users\salem\Documents\colorado\pop.tif
C:\Users\salem\Documents\colorado\roads.tif
C:\Users\salem\Documents\colorado\solar.tif
C:\Users\salem\Documents\colorado\tranlines.tif

Fig. 4.10: The second step of the WLC wizard - selecting the criteria layers.

4.3.4.3 Reclassification of the Criteria Raster (Normalization)

This screen was demonstrated before in section 4.3.3 Reclassification.

4.3.4.4 Definition the Weights

In this window, the range of original values of every layer is displayed, the priority for every layer is defined, and the final weights are derived. The user is allowed to enter the priorities in the third column, then after clicking "derive weights" button, all final weights are calculated and placed in the fourth column. The first and the second columns show the range of values for the layers. MC-Analyst validates the derived weights final values according to SWT (examples shown in Eq. 2-4).

In the event that the derived weights did not sum up to 1, the wizard shows an error. This screen of the wizard contains a help button that is once clicked a pop-up window will show an example of how the weights are derived. The displayed range for every criteria values before the normalization process keeps the user aware of the change in original criteria values as stated in the SWT.

Fig. 4.11 shows the flow chart of the weights validation and derivation. Fig. 4.12 shows the flow chart of how the weighted criteria rasters are calculated and combined to produce the final suitability result. The weights definition screen of the wizard is showed in Fig. 4.13.

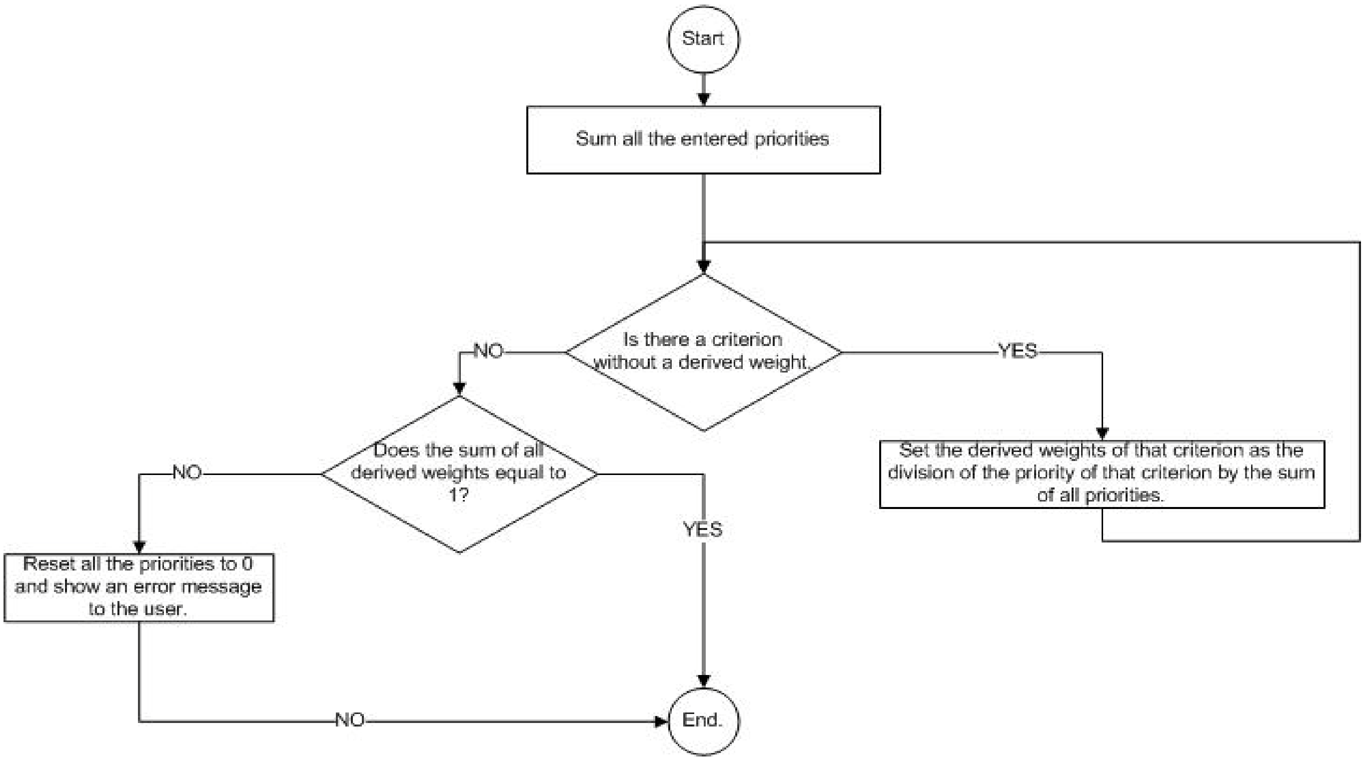


Fig. 4.11: A flow chart showing how weights are validated and derived.

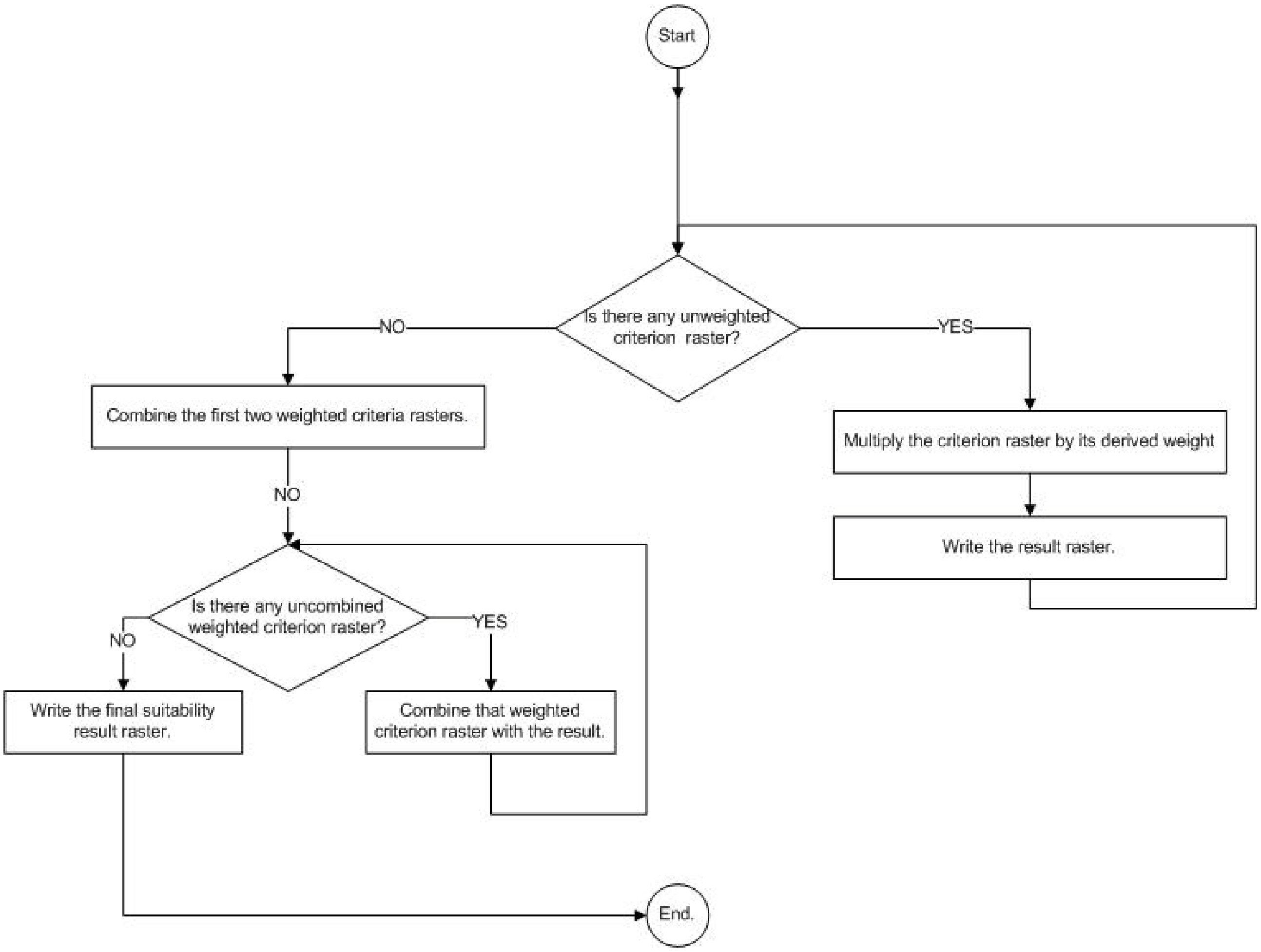
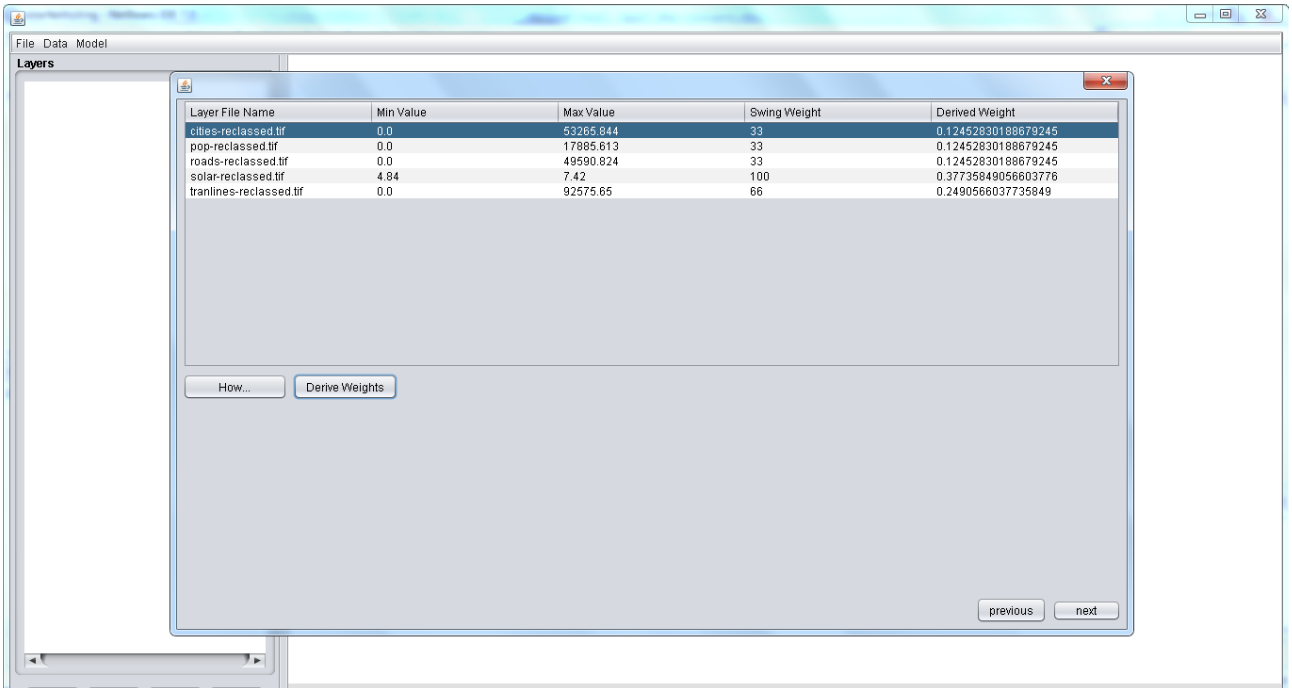


Fig. 4.12: A flow chart showing how weighted rasters are produced and combined to produce the final suitability raster.



Layer File Name	Min Value	Max Value	Swing Weight	Derived Weight
cities-reclassed.tif	0.0	53265.844	33	0.12452830188679245
pop-reclassed.tif	0.0	17885.613	33	0.12452830188679245
roads-reclassed.tif	0.0	49590.824	33	0.12452830188679245
solar-reclassed.tif	4.84	7.42	100	0.37735849056603776
tranlines-reclassed.tif	0.0	92575.65	66	0.2490566037735849

Fig. 4.13: The last step of the WLC wizard - The Weights Definition.

After this window, the final result path is displayed and the wizard is finished.

4.4 Implementation of MC-Analyst

MC-Analyst was developed using Java Standard Edition (Java SE) 1.7.0.17 Software Development Kit (SDK), GeoTools 10.3 which is the available version from www.geotools.org at the time of development, and for the development environment, the free open source NetBeans 7.3 (www.netbeans.org) Integrated Development Environment (IDE) was used to ease the process of development, code compilation and user interface design.

4.5 The License of MC-Analyst

GeoTools is licensed under LGPL. MC-Analyst uses GeoTools to achieve the work but does not mix any code with it nor there have been any changes or modifications on the copy of GeoTools coupled with MC-Analyst. Actually, MC-Analyst used the compiled java files (Oracle, 2014) (Java byte code in JAR file formats) of GeoTools. As to MC-Analyst, it is licensed under GPL. However, the distribution of MC-Analyst must provide some way to the end user to access GeoTools. This was accomplished by providing a link to GeoTools. This was mentioned in the "readme" file distributed with MC-Analyst.

4.6 The Source Code

Appendix A lists a review of the mostly used classes of GeoTools in MC-Analyst. A complete NetBeans project of MC-Analyst was associated with the thesis.

5 The Example Case Study

5.1 Data and the Study Area

In order to test MC-Analyst, a case study was adopted from a previous study carried out by Janke (2010) in which the best site location for solar and wind farms was investigated in Colorado state of USA. The case study demonstrated here was the one that is related to the solar farms only. Even though MC-Analyst can analyze both cases (solar and wind farms) that both of the cases share the input and format of the input (all final formats of the criteria are rasters) except for one single criteria (wind data layer and solar data layer) that is exclusive to every case.

Employing this case study was just an example on how MC-Analyst could be used and was not intended to be considered as a final result upon which decisions could be taken in real scenario. Any other case study could have been used to demonstrate MC-Analyst as well.

The criteria involved in this example were solar DNI (Direct Normal Irradiance), distance to transmission lines, distance to primary roads (hereafter simply referred to as roads), distance to cities, and population density. The conserved areas used were the federal lands.

In the main test of MC-Analyst, solar irradiation layer was assigned the highest weight, the second highest weight was assigned to distance to transmission lines layer, and all remaining layers were assigned similar weights (Table 5.1). In another test, the highest weight was assigned to distance to transmission lines layer (Table 5.2). At the last test, the highest weight was assigned to distance to roads layer (Table 5.3). Outputs from the three tests were compared.

Colorado State covers 270,000 km^2 that ranges approximately from 102^o-110^{o}W longitude and from 37^o-41^{o}N latitude. Fig. 5.1 shows the location of Colorado State.

DNI (Direct Normal Irradiance) is the component of solar irradiation that reaches a surface of the earth (the direction of sun irradiation is normal to the surface) without any atmospheric losses due to scattering or absorption and it is especially important for Concentrated Solar Power systems (Solar and PV Data at www.solargis.info). DNI data of Colorado was downloaded from the National Renewable Energy Laboratory (www.nrel.gov). This data set provides monthly average and annual average for total solar resource. Original data had a spatial resolution of 40 x 40 km in size and a time window between 1985 and 1991.

However, in the example, the annual average has been used. Solar annual average data is shown in Fig. 5.2.

According to Department of Local Affairs of Colorado (State of Colorado, Department of Local Affairs, 2010), the census counted approximately 5 million population. Cities GIS data was downloaded from US Census Bureau (United States Census Bureau, US Department of Commerce, Topologically Integrated Geographic Encoding and Referencing (TIGER Products), 2013). Fig. 5.3 shows the map of population and cities distribution.

Roads GIS data of 2013 was downloaded from US Census Bureau (United States Census Bureau, US Department of Commerce, Topologically Integrated Geographic Encoding and Referencing (TIGER Products), 2013). Federal lands and transmission lines GIS data was downloaded from (Colorado Energy, Colorado Renewable Resource Database). Fig. 5.4 shows the primary roads, federal lands, and transmission lines of Colorado.

Fig. 5.1: Colorado State, USA.

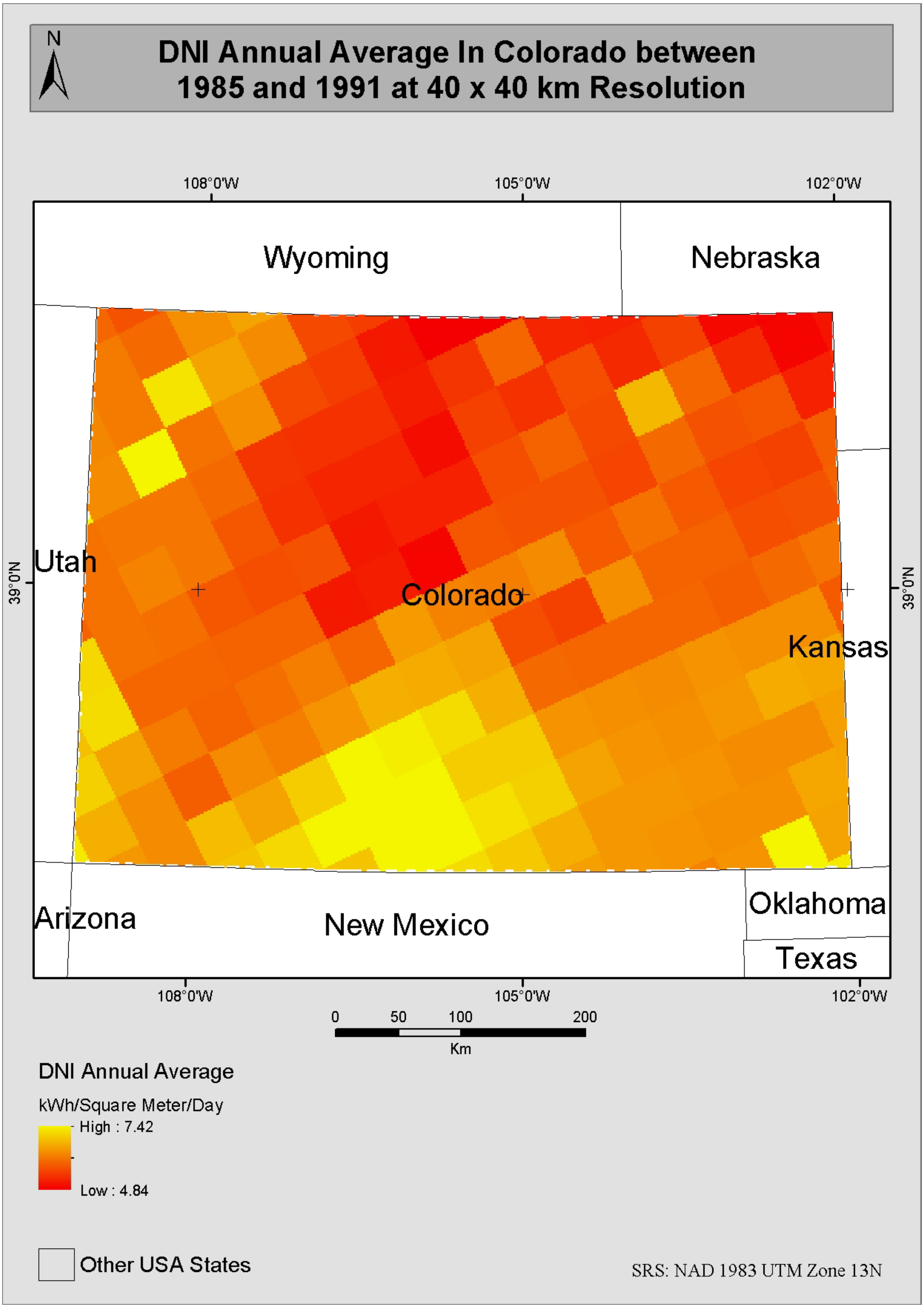


Fig.5.2: DNI annual average solar resource in Colorado for years 1985-1991.

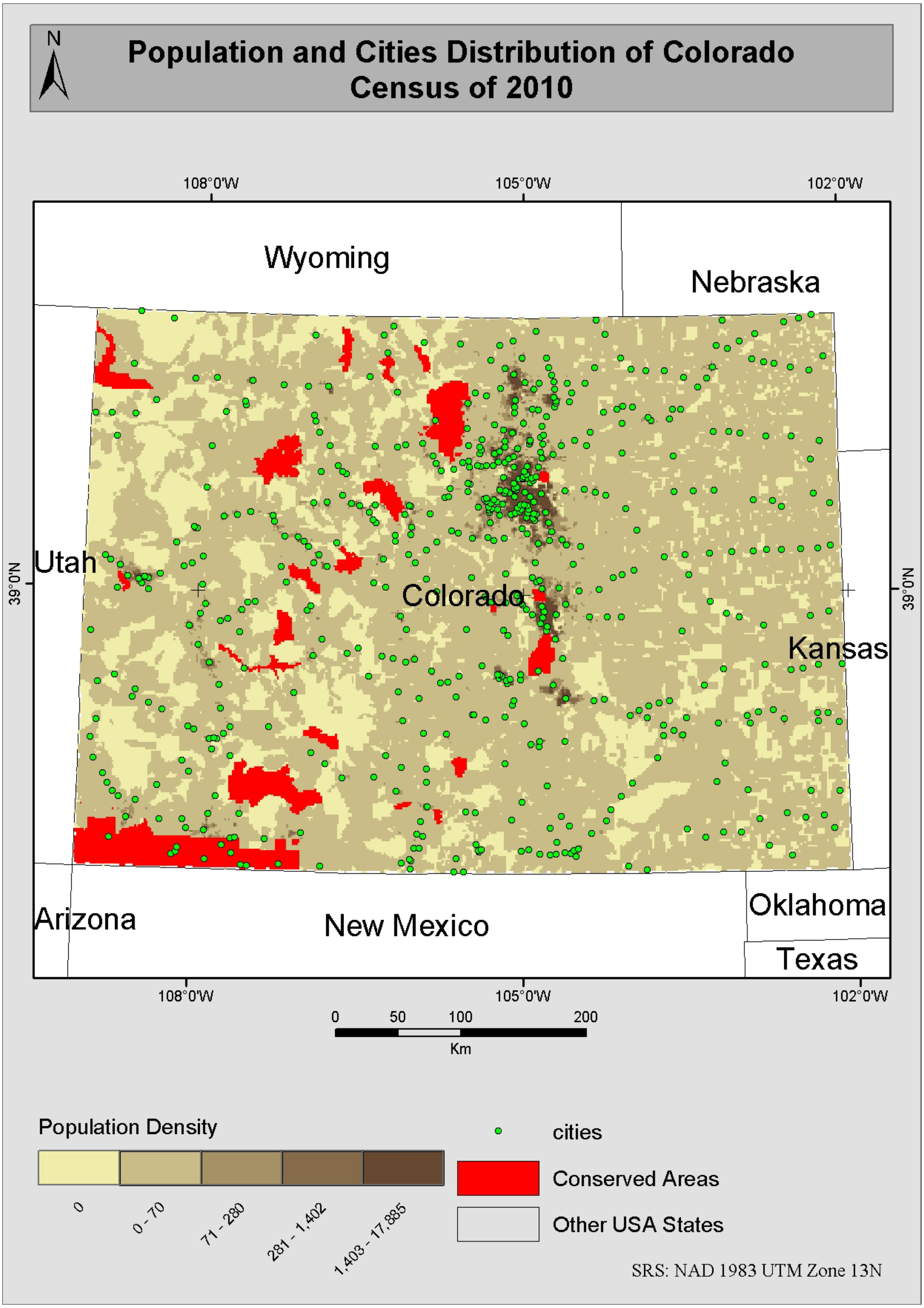


Fig. 5.3: Population and cities distribution of Colorado census of 2010.

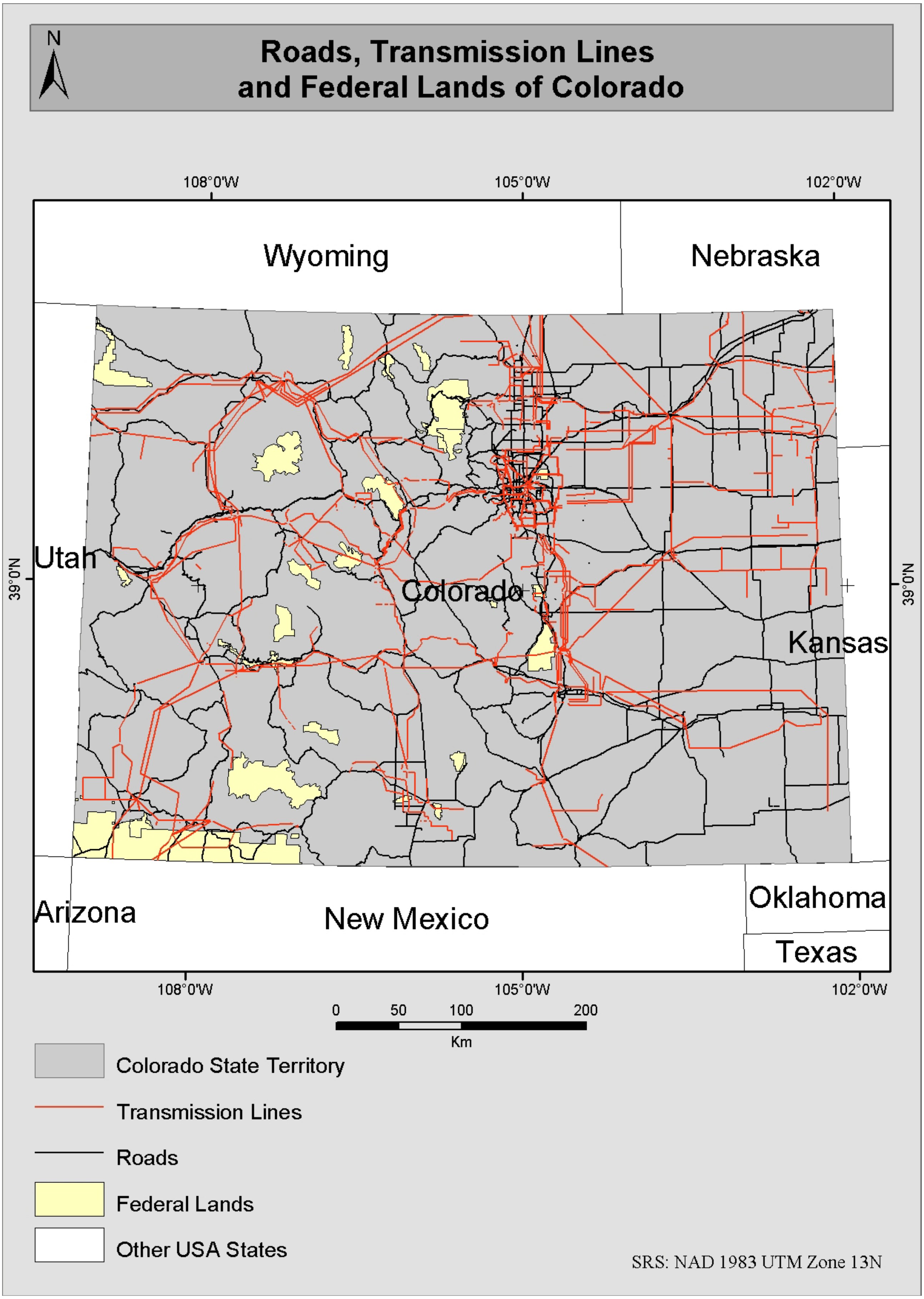


Fig. 5.4: Roads, transmission lines, and federal lands of Colorado.

5.2 Applying the model

5.2.1 General Methodology

There is a large number of GIS-MCEA studies (Malczewski, 2006). Three main motivations were behind choosing Janke (2010): 1) The used methodology was simple and straightforward in terms of weights assignment, 2) The number of used criteria was small, and 3) The availability of data.

Janke (2010) demonstrated a simple methodology that used direct weights assignment. Therefore, it was a good motivation that weights could be easily approximated in applying the case. The considerably few criteria (five criteria) also allowed for easier interpretation and comparison of the suitability results that were output from the three weights alterations applied on the criteria layers. All data of Colorado was available online.

Janke (2010) chose a resolution of 1500 m for his study. This was the same spatial resolution applied in the test here. Since the study area covered all Colorado, higher resolution will bring a more accurate result but will highly load the performance. The spatial resolution depends on the objective of the study.

After data had been collected, data preparation tasks such as clipping and re-projecting were applied. Federal lands layer was used as constraints layer. Criteria layers were multiplied by the constraints layer to exclude reserved areas. Weights definition, criteria layers weighting (multiplication of raster layers by the defined weights), and weighted layers combination were applied then.

5.2.2 Data

The main objective of the case study determines the criteria that should be involved. Solar DNI is one kind of solar energy used specially for Concentrated Solar Power (CSP) systems that use lenses to accumulate solar irradiance (www.nrel.gov). Thus, the study was implicitly intended to evaluate suitability for CSP systems. Involving distance to transmission lines and distance to roads criteria account for cost of site's connectivity to the current infrastructure and site's accessibility from current roads. The other two criteria were distance to cities and population density. The site would be considered more suitable if it had less population density and further from cities. These two criteria account for the impact on the public.

Even though Janke (2010) included land cover layer in his study, the criteria did not account for other environmental factors such as humidity, species migration routes, natural risk of floods and earth quakes, temperature, precipitation rate, distance to rivers and water bodies, distance to wild life areas, and distance to environmentally sensitive areas. For economic factors, criteria such as land slope, aspect, distance to power grids, and industrial areas could have been included. For conserved areas, land types such as private lands, schools, museums, and parks could have been included in the constraints layer as well.

Therefore, the involved criteria accounted only for some economical costs (accessibility (roads) and infrastructure (transmission lines)) and influence on public life quality (cities and population). A real life scenario case study should include more factors. However, this is dependent on the objective of the study.

5.2.3 Data Preparation

Before using the data as input to MC-Analyst, a series of data preparation tasks was performed such as projection, clipping, rasterization, resampling, and generating Euclidean distance rasters. These operations were performed using ArcGIS. Further operations such as removing reserved areas of the model (federal lands in this case study), normalization of raw values, weights definition, and rasters combination were performed using MC-Analyst. Fig. 5.5 shows the flow chart of applying the model tasks with a legend that clarifies which tasks were carried out by MC-Analyst.

Data was projected to North American Datum, Universal Transverse Mercator, Zone 13N. Clipping of the data was performed by the state boundaries to make sure that all data layers are of equal extent. Rasterization of vector data was performed at 1500m resolution. For all criteria layers, MC-Analyst replaces the No Data values with Not a Number (NaN) to make sure that the No Data value is similar in the process among all layers. NaN value is reclassified to 0 in normalization.

For proximity layers (layers that defines criteria of distance such as roads, cities, and transmission lines), distance rasters were created using ArcGIS tool Euclidean Distance. Those distance rasters are used as the input of MC-Analyst.

Federal lands areas are considered as conserved areas and should be excluded out of all layers of the study area before the normalization phase is applied. Thus, in MC-Analyst, the federal lands layer was input as constraints layer, which is a raster of cells with values of 1 or NaN, where cells with 1 are allowed areas and cells with NaN are conserved areas. Then, multiplication of the constraints raster with all criteria rasters is performed after unifying the No Data value to NaN in all criteria rasters. This procedure excludes the values of the excluded areas off the criteria layers before the normalization phase. Therefore, it is important to include all excluded areas in one raster, that is the constraints raster. Practitioners usually exclude these areas off the final suitability raster, which is a source of error practice (Malczewski, 2006).

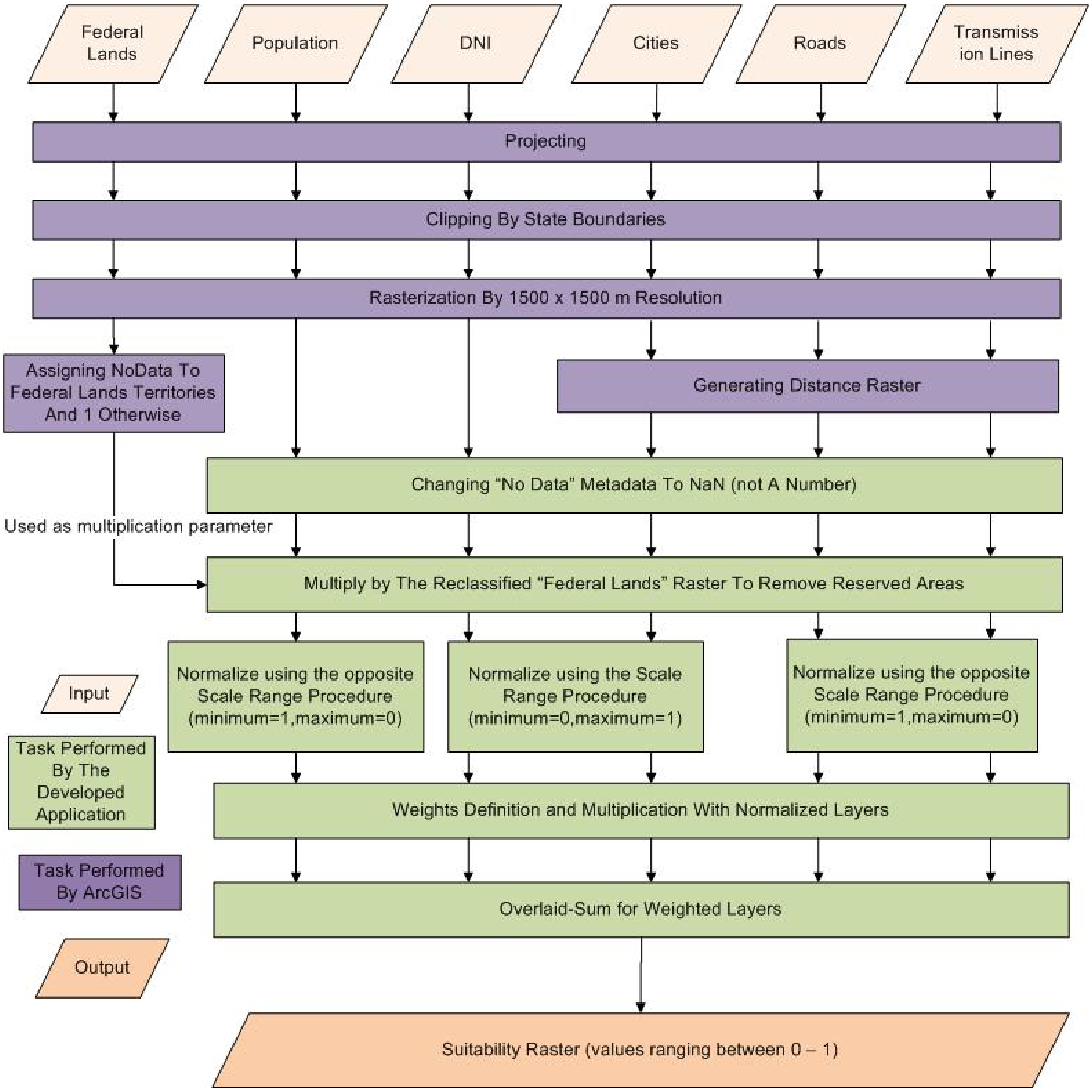


Fig. 5.5: The flow chart of tasks performed during applying the model process.

5.2.4 Normalization

The input of this phase is the resized and constrained criteria layers. Normalization is performed using one of the three normalization methods stated in section "4.1.2 Values Normalization of Criteria Layers". For this case study, only two methods were used, Score Range Procedure (SCP) which projects [minimum-maximum] on [0-1] and its opposite which projects [minimum-maximum] on [1-0]. SCP was applied on the criteria rasters of DNI and distance to cities. This is logical because the larger volume of solar resource of the location and the further the location is from cities, the higher suitability it gains. For criteria rasters of population, distance to roads, and distance to transmission lines, the opposite SRP was applied. Because ideal conditions of the candidate locations require that the location should be far from higher population areas, closer to roads, and closer to transmission lines. Fig. 5.6-10 show normalized values for criteria: solar CSP, distance to transmission lines, distance to cites, distance to roads, and population density respectively.

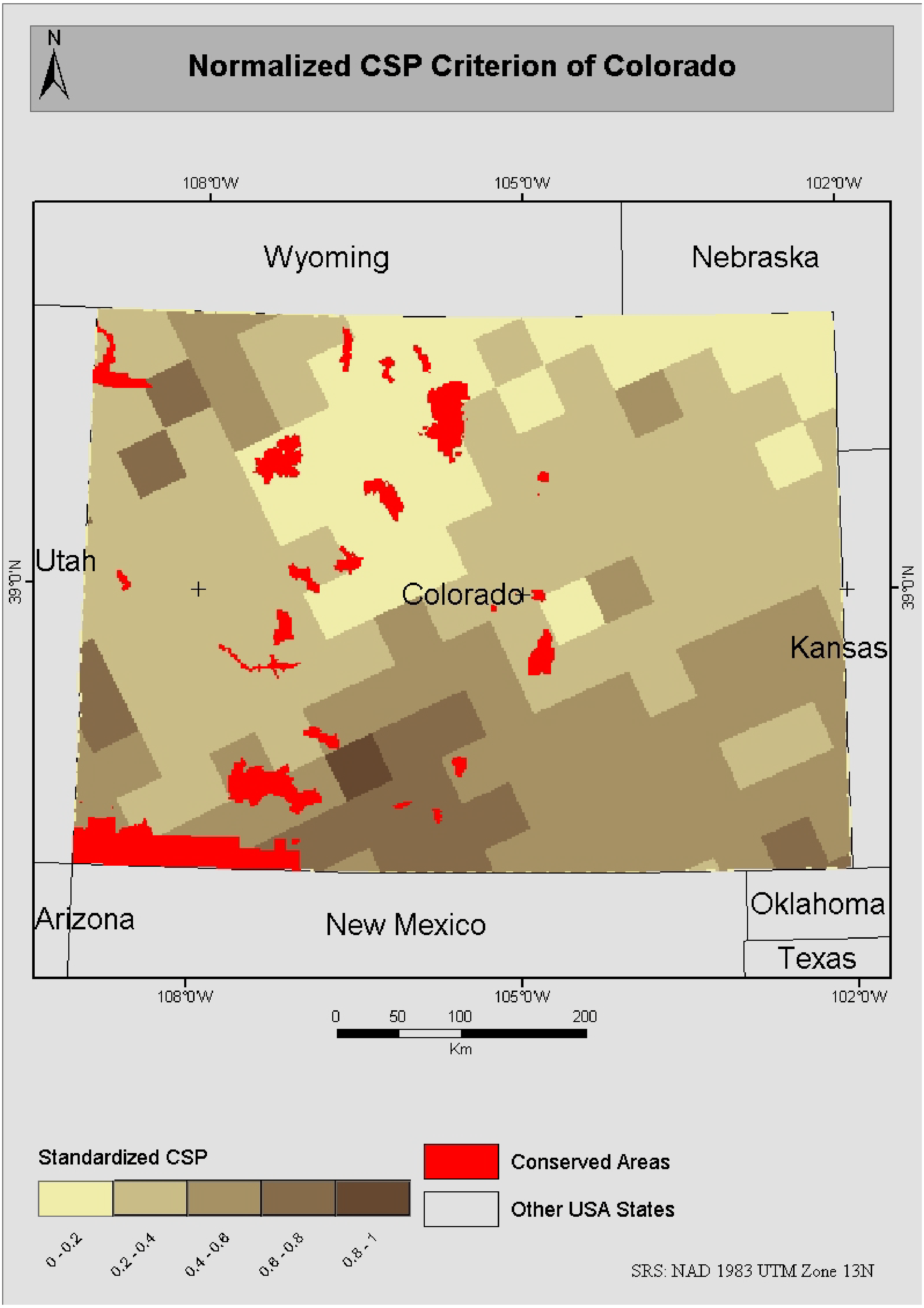


Fig. 5.6: Normalized Concentrated Solar Power (CSP) criterion in Colorado.

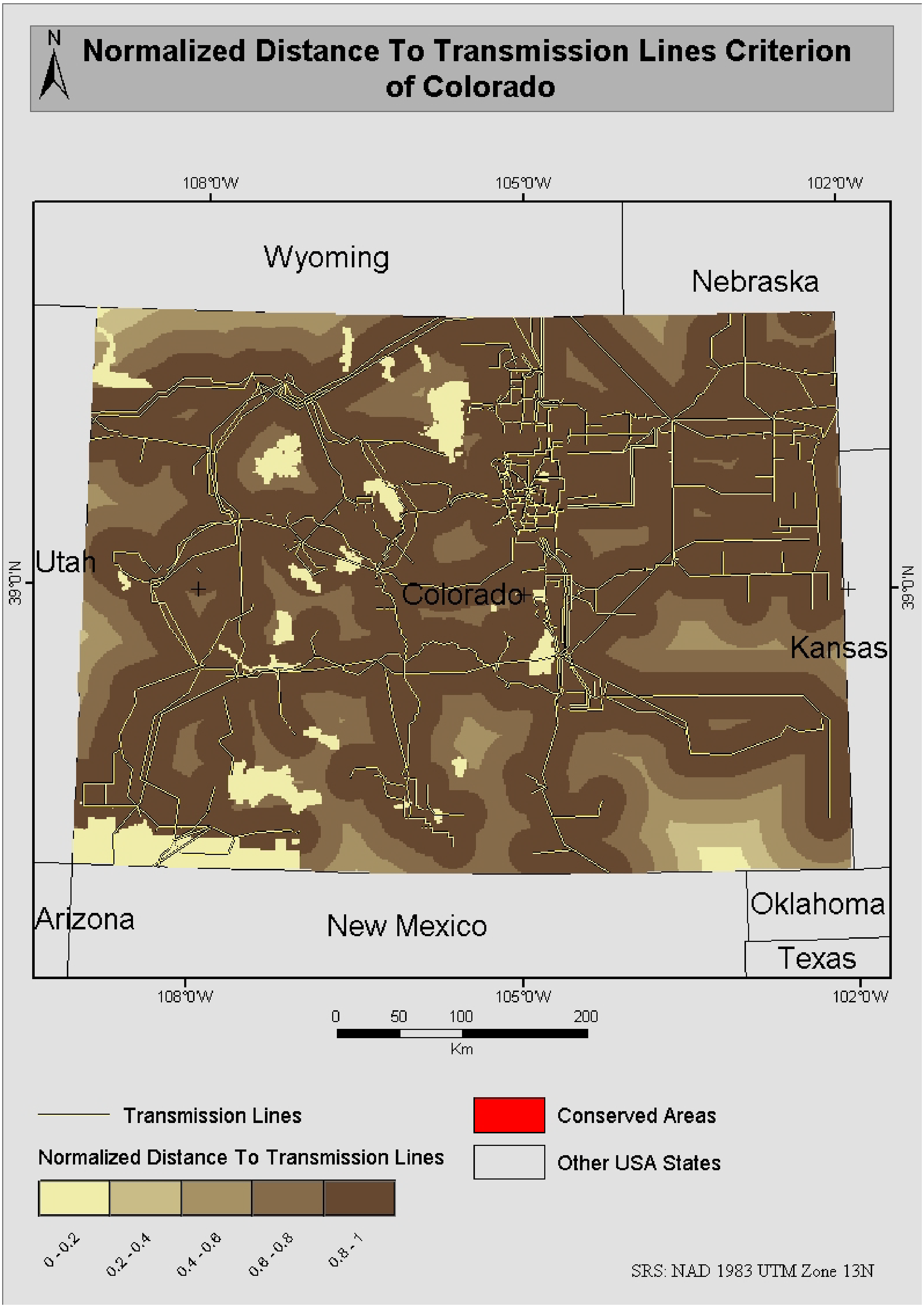


Fig. 5.7: Normalized distance to transmission lines criterion in Colorado.

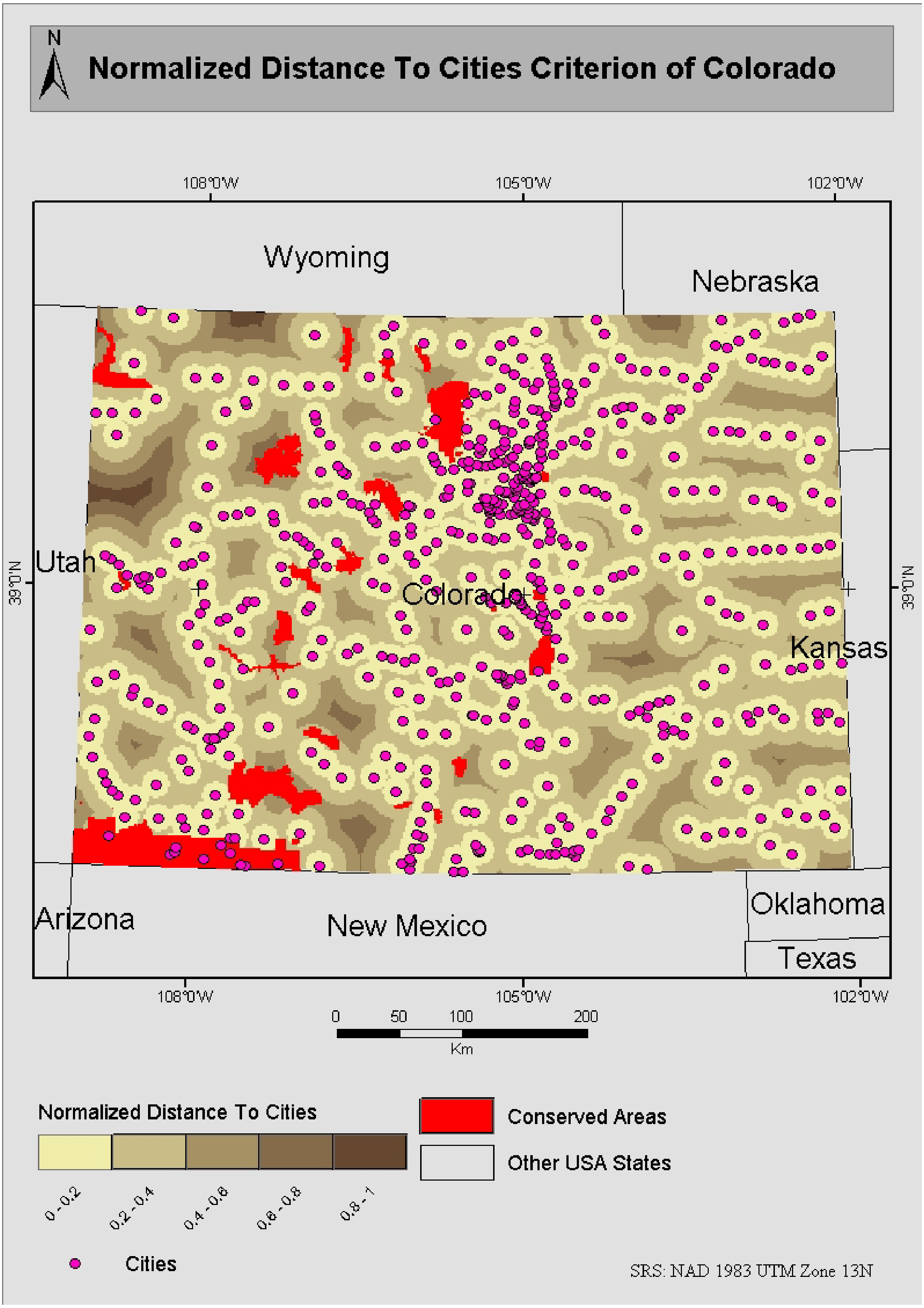


Fig. 5.8: Normalized distance to cities criterion in Colorado.

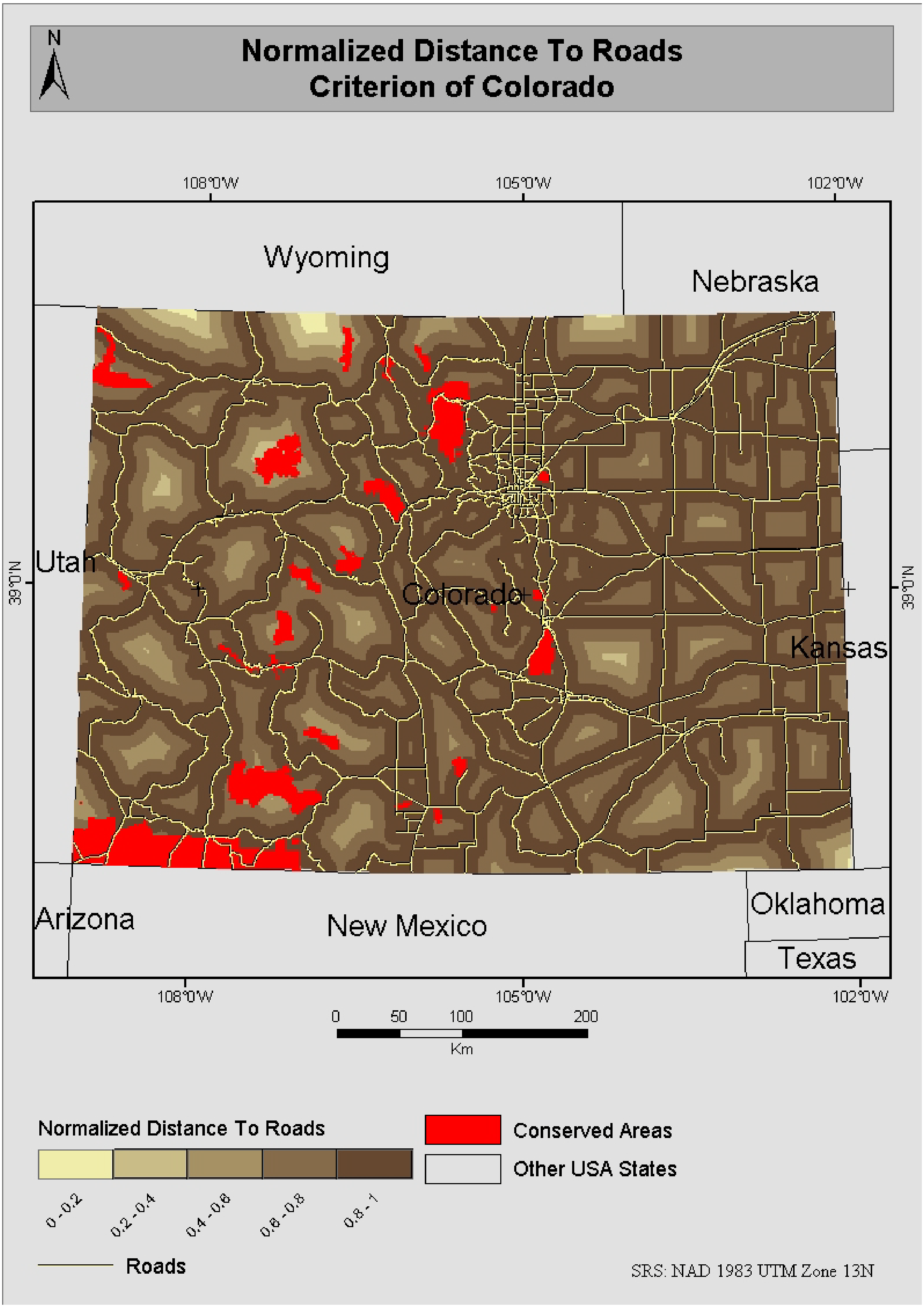


Fig. 5.9: Normalized distance to roads criterion in Colorado.

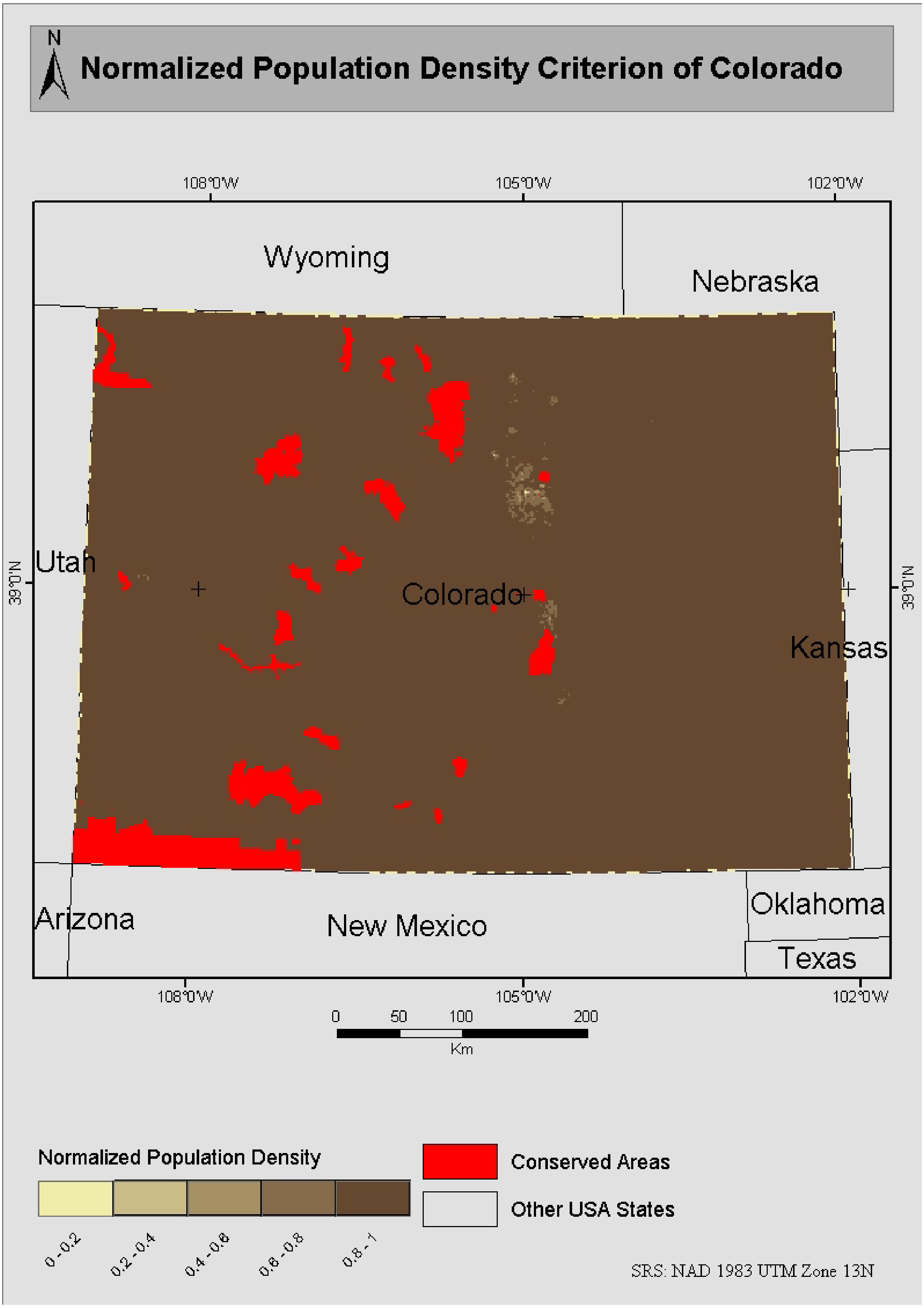


Fig. 5.10: Normalized population criterion in Colorado.

5.2.5 Weights Definition

Weights to be used in the study were taken from Janke (2010). Table 5.1 lists the criteria considered, the order weights assigned by Janke (2010), the priorities suggested using Malczewski (2000) methodology, and the derived weights.

To examine how the final suitability map varies according to the defined weights, two further weights alterations have been applied. Both alterations assigned similar priorities of 30 to all the criteria except for the criteria of distance to transmission lines and distance to roads that were assigned the highest weight of 100 in the second and the third weights alterations respectively. Any other layers could have been used for the alterations. However, the motivation behind choosing transmission lines and roads that they had a noticeable pattern that could be easily seen by the viewer. Thus, there were three weights alterations used in the study. Tables 5.1-3 show the weights for the three alterations. The weighting operation was a simple multiplication between a constant (the weight) and cell values in the criteria rasters.

It is important to draw the attention that in applying WLC model, land cover criterion that had been used in Janke (2010) was dropped in the example here because the downloaded dataset of land cover had classes that did not match with what Janke (2010) used in his study. It is important to again mention here that this case study is only an example to demonstrate MC-Analyst and not intended as real life case on which a decision is made. This actually does not affect the performance or the liability of MC-Analyst because MC-Analyst abstracts the criteria layers of the context. It processes the layers regardless of what phenomenon each layer represents on the field.

Janke (2010) assigned a weight of 1 to land cover which did not have strong impact on the final results given that DNI weight was 3, distance to transmission lines was 2 and all other criteria's weights were similar, 1. If land cover had been included in the study, it would have been assigned the derived weight of 0.1 by applying equations in Malczewski (2000) methodology (Eq. 2-4).

Table 5.1: The model's criteria, their ideal condition according to Janke (2010), ordered weights, priorities, and derived weights.

Criteria Raster	**Ideal Condition**	**Order Weight in Janke (2010) Study**	**Suggested Priority Using Malczewski (2000) Approach**	**Derived Weight**
DNI	Highest solar DNI	3	100	0.4
Distance To Transmission Lines	Closer to transmission lines	2	60	0.24
Distance To Cities	Far from cities	1	30	0.12
Distance To Roads	Far from high population areas	1	30	0.12
Roads	Closer to roads	1	30	0.12

Table 5.2: Second weights alteration.

Criterion	Defined Priorities	Derived Weights
DNI	30	0.13
Distance To Transmission Lines	100	0.45
Distance To Cities	30	0.13
Distance To Roads	30	0.13
Population Areas	30	0.13

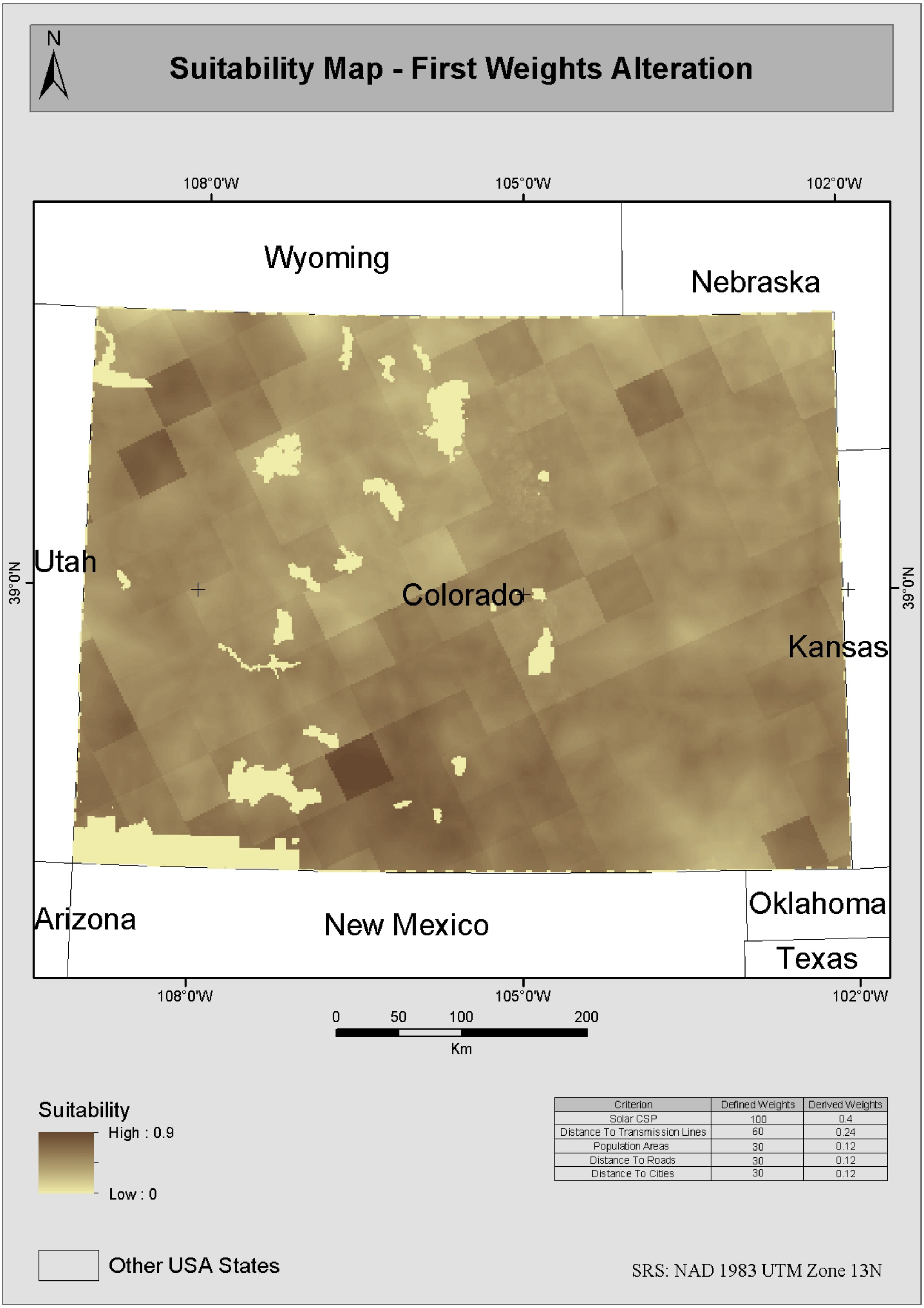
Table 5.3: Third weights alteration.

Criterion	Defined Priorities	Derived Weights
DNI	30	0.13
Distance To Transmission Lines	30	0.13
Distance To Cities	30	0.13
Distance To Roads	100	0.45
Population Areas	30	0.13

5.2.6 Applying the Combination

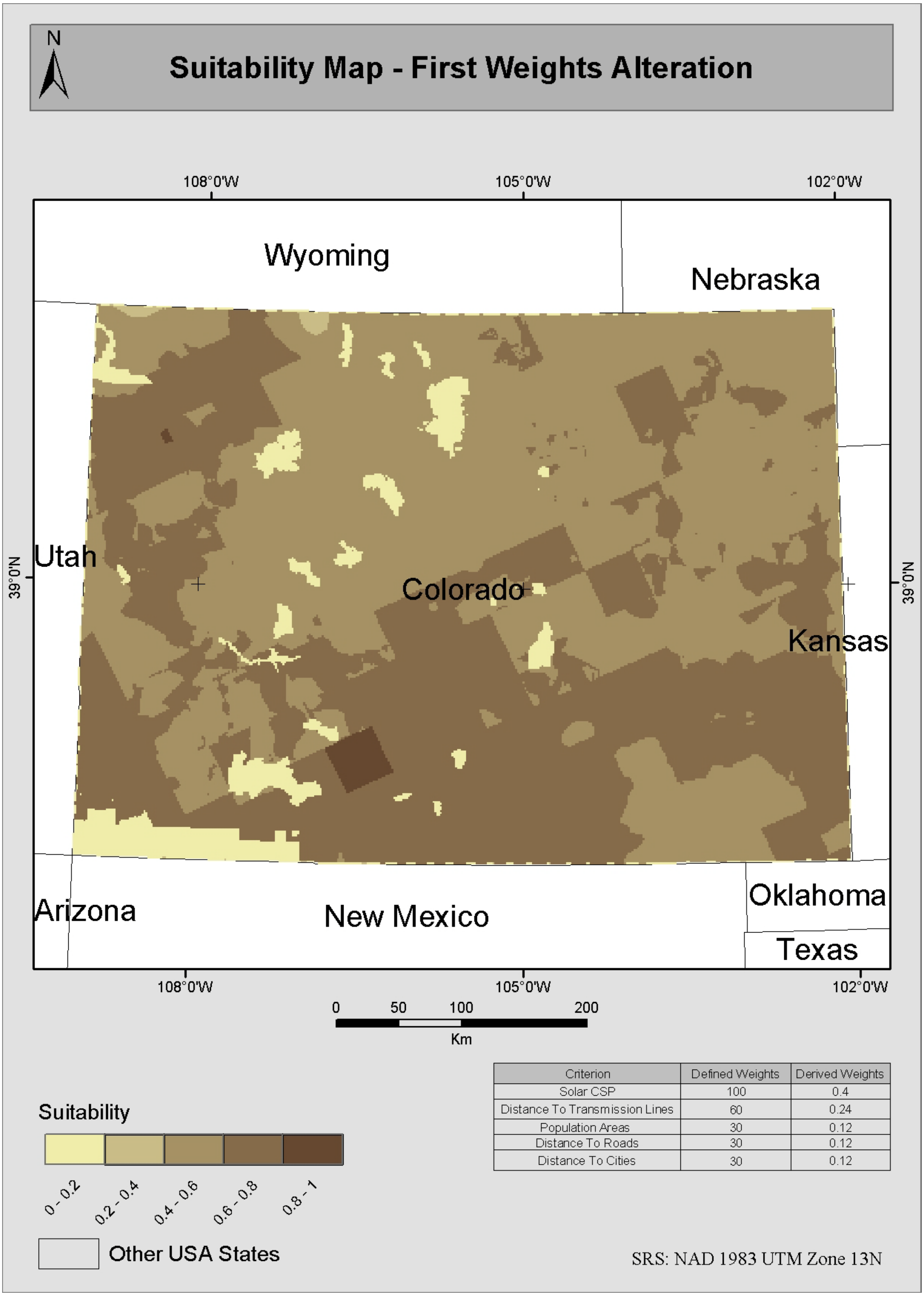
After the criteria layers had been weighted, combination or sum-overlay was applied to these layers in order to produce the final suitability raster. The combination result for the criteria layers that were weighted by the weights defined in Table 5.1 is shown in Fig. 5.11-12.

Suitability maps produced by applying the combination after the second (Table 5.2 - highest weight assigned to the criterion of distance to transmission lines) and the third (Table 5.3 - highest weight assigned to the criterion of distance to roads) weights alterations are shown in Fig. 5.13 and 5.14 respectively.



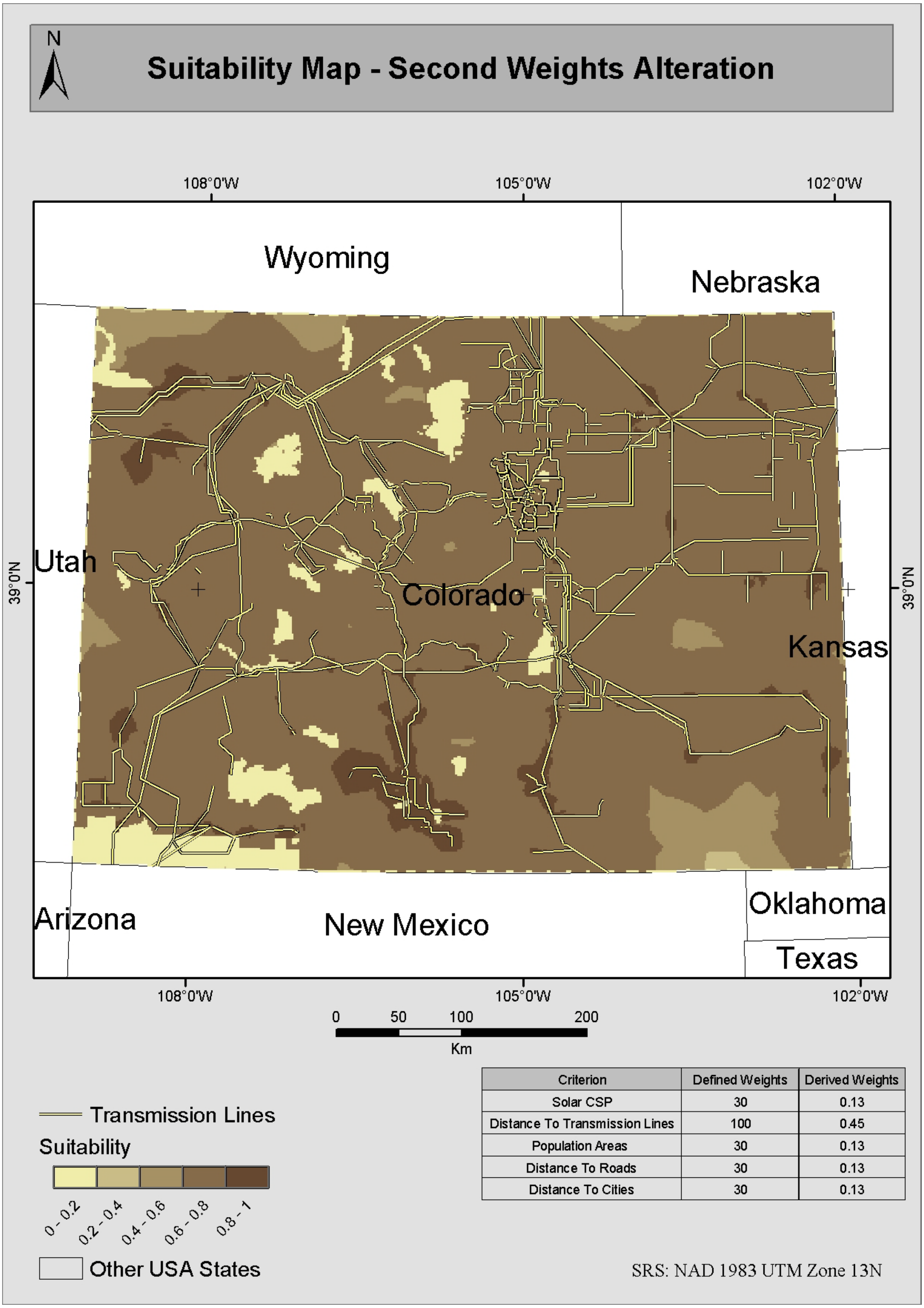
Criterion	Defined Weights	Derived Weights
Solar CSP	100	0.4
Distance To Transmission Lines	60	0.24
Population Areas	30	0.12
Distance To Roads	30	0.12
Distance To Cities	30	0.12

Fig. 5.11: The first weights alteration suitability map produced by the Weighted Linear Combination (WLC) in Colorado.



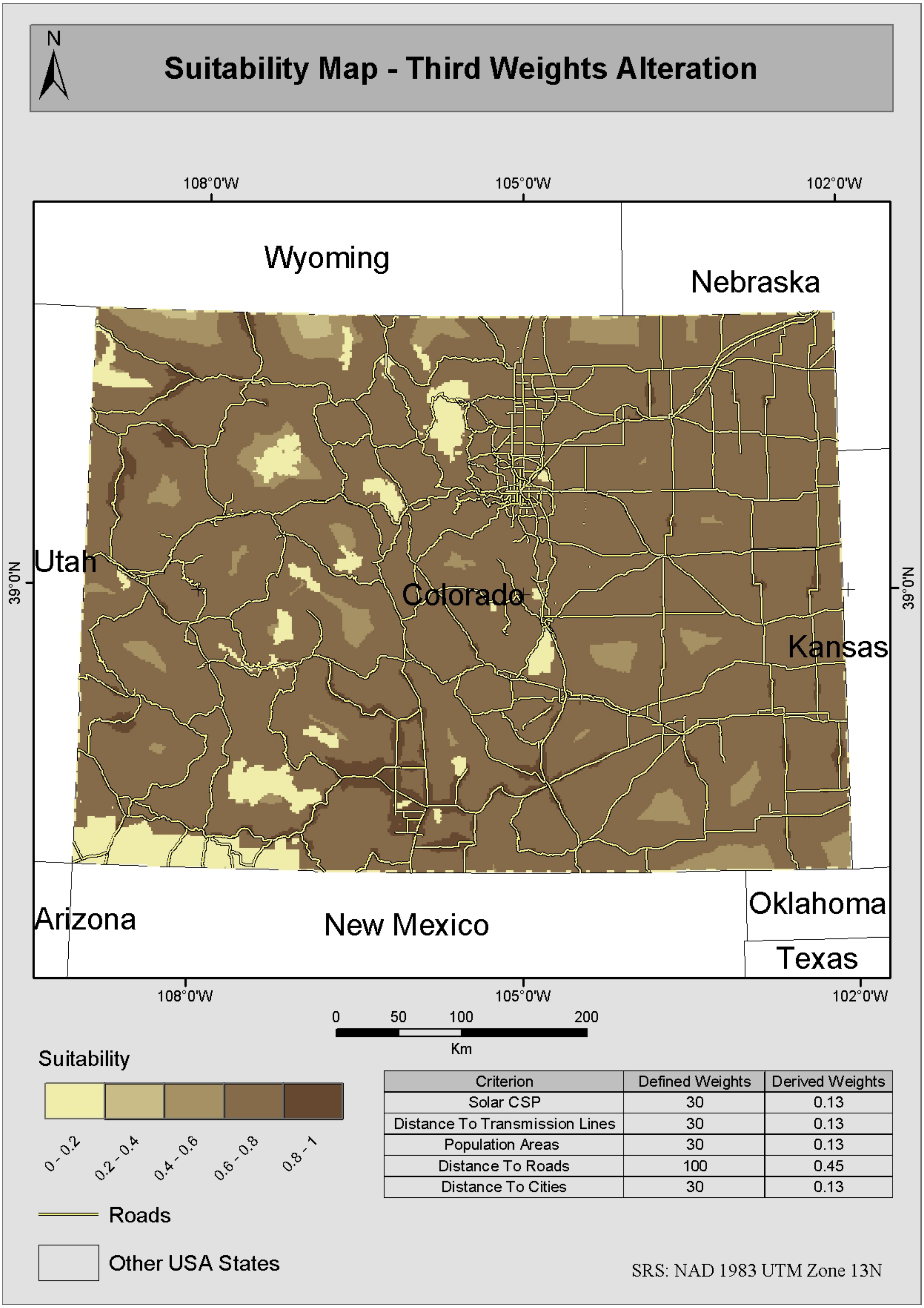
Criterion	Defined Weights	Derived Weights
Solar CSP	100	0.4
Distance To Transmission Lines	60	0.24
Population Areas	30	0.12
Distance To Roads	30	0.12
Distance To Cities	30	0.12

Fig. 5.12: The classified first weights alteration suitability map produced by the Weighted Linear Combination (WLC) in Colorado.



Criterion	Defined Weights	Derived Weights
Solar CSP	30	0.13
Distance To Transmission Lines	100	0.45
Population Areas	30	0.13
Distance To Roads	30	0.13
Distance To Cities	30	0.13

Fig. 5.13: Second weights alteration suitability map produced by the Weighted Linear Combination (WLC) in Colorado.



Criterion	Defined Weights	Derived Weights
Solar CSP	30	0.13
Distance To Transmission Lines	30	0.13
Population Areas	30	0.13
Distance To Roads	100	0.45
Distance To Cities	30	0.13

Fig. 5.14: Third weights alteration suitability map produced by the Weighted Linear Combination (WLC) in Colorado.

6 Results

The main purpose of this analysis is to examine the output of MC-Analyst and to discuss whether or not it performed as expected.

A visual analysis of the suitability result was performed over three areas: area 1, 2, and 3 (Fig. 6.1). The reason for choosing areas 1 and 2 was their visible contrast that had high and low solar CSP values respectively. The reason for choosing area 3 was that it had road and transmission lines patterns that was easily traceable by the viewer.

A statistical analysis was performed to examine the three suitability results in terms of numerical statistics.

6.1 Visual Analysis

When reviewing the first suitability result in Fig. 5.11, the implemented WLC model showed expected results. Since solar DNI was assigned the highest weight (0.4), followed by distance to transmission lines (0.24), then followed by all other criteria (0.12), the first suitability result showed higher suitability in areas where DNI values were higher and transmission lines were closer. Referring to Fig. 5.11 of suitability and Fig. 5.6 of normalized solar CSP, there was spatial correspondence between areas of high suitability with those of relatively high solar CSP. A classified suitability of the first suitability result was showed in Fig. 5.12. It showed that almost 70% of the southern half of Colorado has an intermediate suitability between 0.6 and 0.8. Those areas were concentrated in the middle and eastern south, and in the western north of Colorado (area 1 in Fig. 6.1).

In addition, there was spatial correspondence between areas of low suitability with those of low solar CSP. Those areas were concentrated in the middle north of Colorado (area 2 in Fig. 6.1). As to the criterion of distance to transmission lines, by referring to the first suitability result and to the normalized distance to transmission lines criterion (Fig. 5.11 and 5.7), there was spatial correspondence between areas of high suitability with those that are closer to transmission lines. For example, by examining east of Colorado in Fig. 5.11, there were found darker horizontal light lines which were areas closer to transmission lines (area 3 in Fig. 6.1). Fig. 6.1 shows areas with high suitability originating from the direct effect of the weights of solar CSP (areas 1) and distance to transmission lines (area 3).

Even though other criteria (distance to cities, distance to roads, and high population areas) have been considered, but their existence did not have visually noticeable effect on the final result. This is explainable that those criteria had a weight of only 0.12.

Further testing for MC-Analyst and the WLC was carried out using the weights alterations showed in Tables 5.2 and 5.3. By reviewing the second suitability result shown in Fig. 5.13 where the criterion of distance to transmission lines was assigned the only highest weight (0.45) and all other criteria were assigned similar weights of 0.13, areas that were closer to transmission lines had considerably higher suitability than other areas. Using the same analogy, by reviewing Table 5.3 of weights alteration and Fig. 5.14 of suitability map, areas that were closer to roads had considerably higher suitability than other areas.

6.2 Statistical Analysis

Three smaller areas were created (areas 4, 5, and 6) with 40 km^2 area for each. The fourth area (area 4) was located in a territory that received the highest amount of solar DNI in Colorado (7.42 kWh/m^2/day). The fifth area (area 5) was located in a territory that was on a transmission line, 24 km away from the nearest primary road, and received relatively less solar DNI (5.1 kWh/m^2/day) than area 4. The sixth area (area 6) was located in a territory that was on a primary road, 40 km away from the nearest transmission line, and also received less solar DNI (5.3 kWh/m^2/day) than area 4. Fig. 6.2 shows the created areas (4, 5, and 6) map.

Suitability statistics for the three areas were extracted from the original values, the normalized values rasters of solar CSP, distance to transmission lines, and distance to roads, and the three suitability results. The total is nine extracted small rasters for every area of the areas 4-6. Three for original values, three for normalized values, and three for suitability results.

Referring to Fig. 6.2 and Table 6.3, area 4 had a suitability mean of 0.85, 0.81, and 0.73 for the first, second and third suitability results respectively. Given that the distance to transmission lines and the distance to roads criteria had less weights than the weight of solar CSP criterion in the first suitability weights alteration, area 4 was supposed to show much higher mean of suitability in the first suitability result than those found in the other two results. Nevertheless, area 4 showed high but approximate mean suitability in all of the three results with the mean suitability of the first result as the highest (0.85). This was because areas of high solar CSP in Colorado were also close to some transmission lines and roads. For example, area 4 was 17.8 km and 11.5 km away from the nearest road and transmission lines respectively. In addition, area 4 had mean normalized values of solar CSP, distance to transmission lines, and distance to roads as 1.0, 0.87, and 0.63 respectively (Table. 6.2).

Referring to Fig. 6.2 and Table. 6.3, area 5 showed an above intermediate mean suitability of 0.7 for the second suitability result and less mean suitability of 0.48 and 0.54 for the first and third suitability results respectively. This was because the distance to transmission lines criterion was assigned the highest weight in the second weights alteration, in addition to the spatial characteristics of area 5 described previously. Following the same approach, area 6 had a mean suitability of 0.53, 0.64, and 0.77 for the first, second, and third suitability results respectively. Fig. 6.2 showed the three small created areas 4-6. Tables 6.1-3 show extracted statistics of areas 4-6 in the original and the normalized values of the three weight-altered criteria as well as the three suitability results.

Besides the impact weights have on the model, relatively high normalized criteria values could affect the final suitability results at little impact. For example, referring to middle south of Colorado in Fig. 5.13 and 5.14 of suitability where solar CSP criterion has been assigned the lowest weight of 0.13, high normalized solar CSP values increased the suitability from the range of 0.6-0.8 to 0.8-1.0 in areas closer to transmission lines (Fig. 5.13) and roads (Fig. 5.14). This emphasizes the concept of Malczewski (2000) that removing reserved areas from all criteria layers should be applied before the normalization because of the impact these areas would have on the normalized values of the final suitability result.

Table 6.1: Areas 4-6 statistics of the original values for solar CSP, distance to transmission lines, and distance to roads rasters.

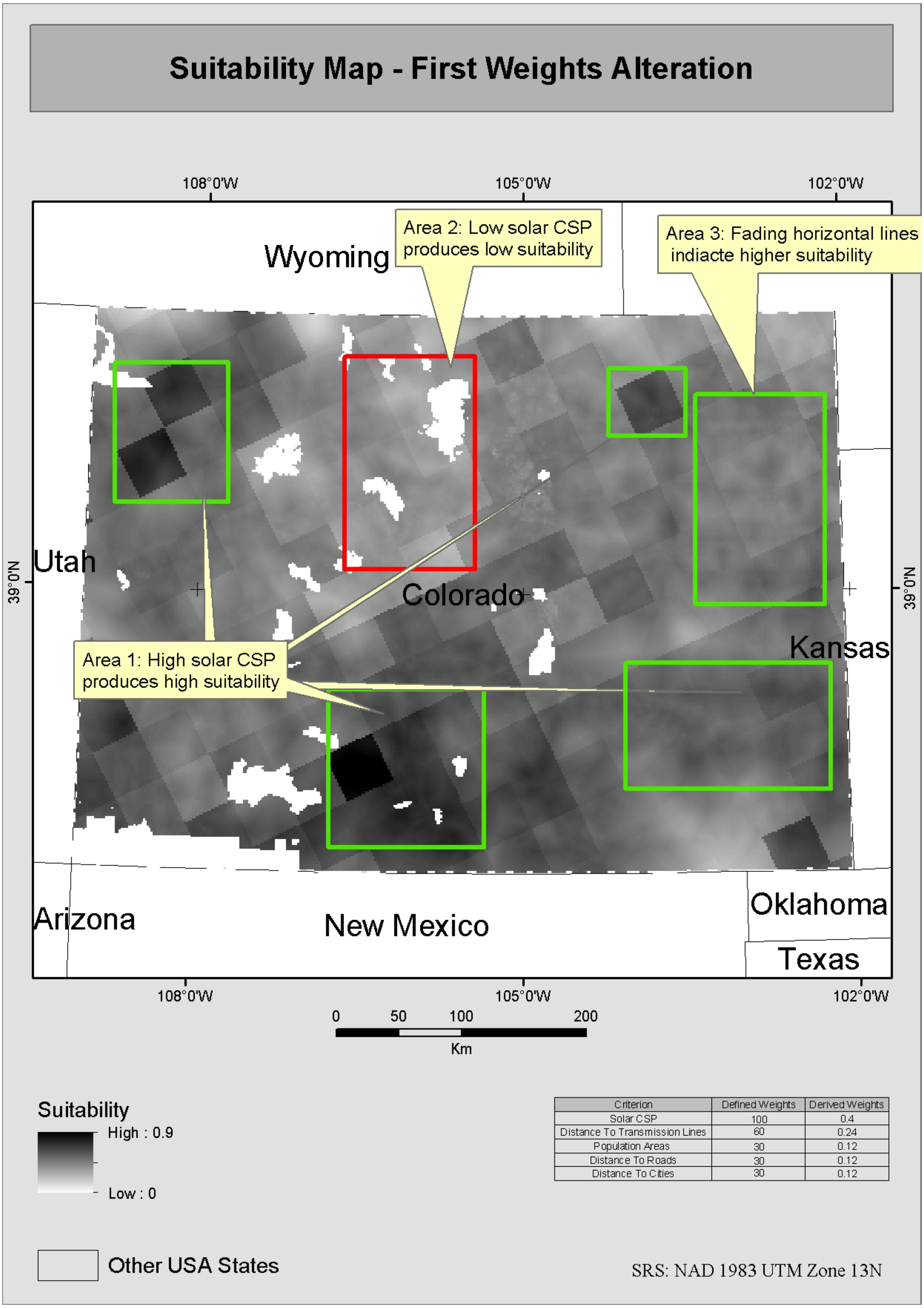
		Original Values		
		Solar CSP ($kWh/m^2/day$)	**Distance to Transmission Lines (km)**	**Distance to Roads (km)**
Area 4	**Min**	7.42	9	14.8
	Max	7.42	14.23	21
	Mean	7.42	11.596	17.822
	standard deviation	0	1.718	1.693
Area 5	**Min**	5.11	0	22.5
	Max	5.15	3.3	27
	Mean	5.14	1.076	24.75
	standard deviation	0.014	1.079	1.732
Area 6	**Min**	5.32	37.1	0
	Max	5.32	44.1	3
	Mean	5.32	40.62	1.8
	standard deviation	0	2.143	1.151

Table 6.2: Areas 4-6 statistics of the normalized values for solar CSP, distance to transmission lines, and distance to roads rasters.

		Normalized Values		
		Normalized Solar CSP	**Normalized Distance to Transmission Lines**	**Normalized Distance to Roads**
Area 4	**Min**	1	0.84	0.57
	Max	1	0.9	0.7
	Mean	1	0.87	0.63
	standard deviation	0	0.018	0.03
Area 5	**Min**	0.1	0.96	0.45
	Max	0.12	1	0.54
	Mean	0.11	0.98	0.5
	standard deviation	0.005	0.012	0.03
Area 6	**Min**	0.18	0.52	0.9
	Max	0.18	0.59	1
	Mean	0.18	0.56	0.96
	standard deviation	0	0.023	0.02

Table 6.3: Areas 4-6 statistics of the suitability results of the three weights alterations.

		Suitability Values		
		First Suitability (solar CPS has the highest weight)	**Second Suitability (distance to transmission lines has the highest weight)**	**Third Suitability (distance to roads has the highest weight)**
Area 4	**Min**	0.84	0.79	0.71
	Max	0.86	0.82	0.76
	Mean	0.85	0.81	0.73
	standard deviation	0.005	0.009	0.01
Area 5	**Min**	0.46	0.67	0.51
	Max	0.49	0.7	0.57
	Mean	0.48	0.7	0.54
	standard deviation	0.008	0.009	0.017
Area 6	**Min**	0.51	0.62	0.75
	Max	0.54	0.65	0.79
	Mean	0.53	0.64	0.77
	standard deviation	0.006	0.01	0.011



Suitability Map - First Weights Alteration

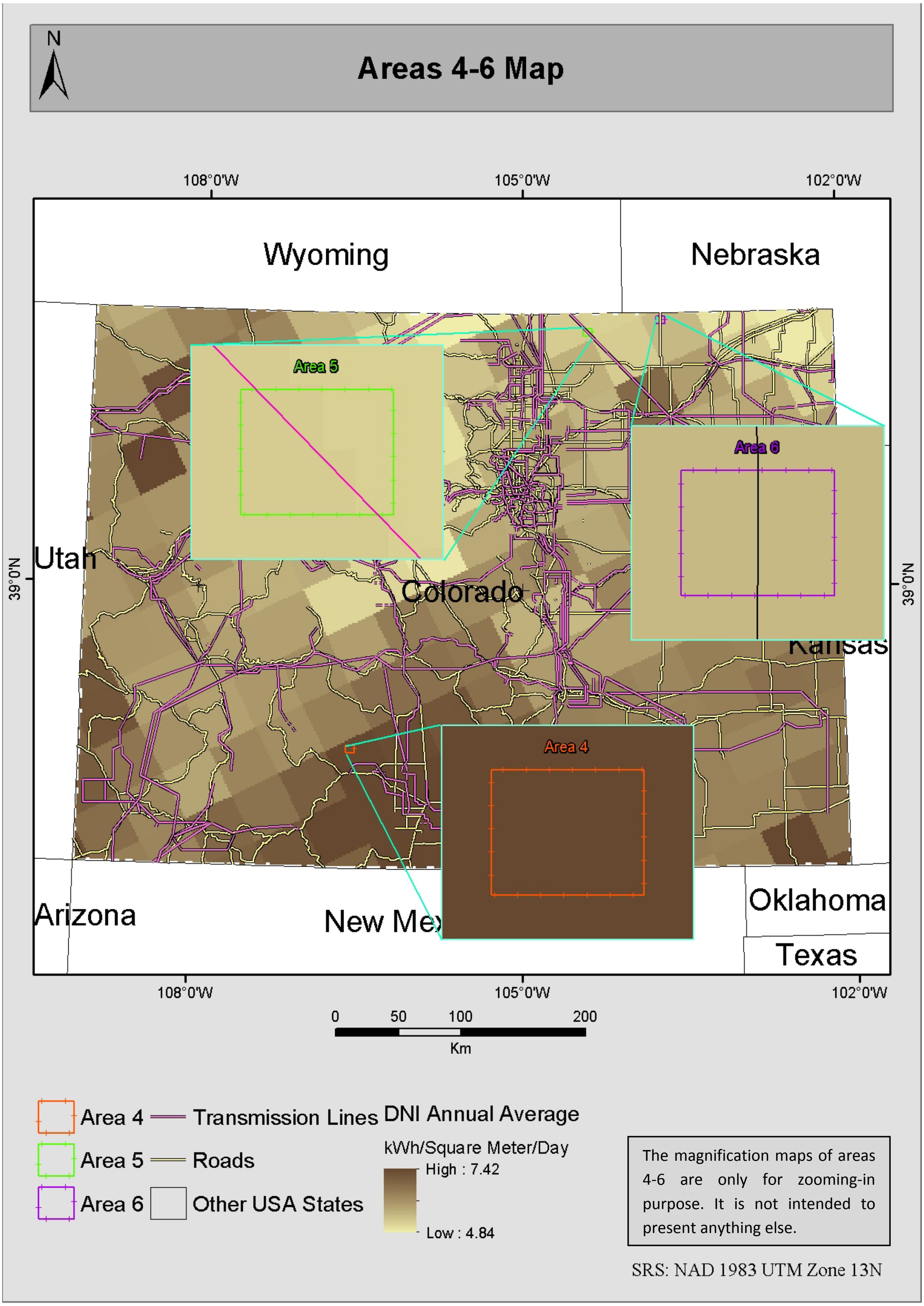
Area 2: Low solar CSP produces low suitability

Area 3: Fading horizontal lines indiacte higher suitability

Area 1: High solar CSP produces high suitability

Criterion	Defined Weights	Derived Weights
Solar CSP	100	0.4
Distance To Transmission Lines	60	0.24
Population Areas	30	0.12
Distance To Roads	30	0.12
Distance To Cities	30	0.12

Fig. 6.1: High suitability and low suitability areas.



The magnification maps of areas 4-6 are only for zooming-in purpose. It is not intended to present anything else.

Fig. 6.2: Areas 4-6 map.

6.3 GeoTools Application Programming Interface (API)

GeoTools provided basic functionality through its set of classes (refer to appendix A). The simplicity of the WLC model allowed its implementation using GeoTools API with little coding. Basic raster operations of the WLC model such as resampling, multiplication, adding, and interpolation were available in GeoTools API through its implementation of Operations.DEFAULT class.

Since GeoTools was implemented in Java SDK, a great dependency of its model is taken from Java raster model. While Java package of java.awt.image implemented the abstract raster model without referencing it to any geographical reference, GeoTools added the geographical aspect of the raster and produced a new model named Coverage which is a raster of Java that has spatial dimension through an extent (Envelope class of GeoTools) and a spatial reference system (Coordinate Reference System class of GeoTools).

Even though MC-Analyst used GeoTIFF format for its raster input and output as it was recommended by GeoTools image tutorial (docs.geotools.org) that GeoTIFF had fast performance, GeoTools was capable of handling other raster formats such as ArcGRID and geo-referenced JPEG files.

As it is the case with raster data format, shape files were strongly supported in GeoTools. Different classes that provided the basic functionality such as reading, writing, schema building, and features creating were implemented in GeoTools. Just as the java raster model was utilized in GridCoverage class of GeoTools, JTS was utilized in GeoTools as its geometry model implementation.

A key point of open source is that developers have a great opportunity of knowledge, experience, efforts, and code sharing. The feature rasterization functionality that was used in the auxiliary wizard of Multi-Buffering and Rasterization (4.3.2 Multi-Buffering and Rasterization) was not built from scratch nor implemented in GeoTools. Instead, it was a contribution by other users (refer to FeatureRasterizer in Appendix A.1) of GeoTools who made it available for the public to use it as desired which saved time and efforts during the development of MC-Analyst.

7 Discussion

MCEA is a practical and easily applicable tool for solving multi-criteria decision problems. The WLC method showed expected results using basic GIS functions. The strength of the WLC method came from its simple application by employing simple GIS operations such as raster reclassification, multiplication, and sum-overlay. Through the different phases of the WLC method defined by Malczewski (2000), incorporating different spatial criteria of the spatial problem (locating sites suitable for solar farms) was applicable. The WLC method evaluated the sites suitability of the whole study area. Since the criteria represented the structure of the problem, before proceeding in applying the methodology, it was important to carefully select the criteria involved in the problem at hand as it will affect the final result. The selected weights reflect how important or influent a criterion for the problem is. Thus, before selecting the criteria and the weights, knowledge about the problem being questioned is necessary to describe the problem, define the goal, and define the factors affecting the goal achievement.

The selected criteria should be measureable, complete, non-redundant and minimal (Malczewski, 2000). This was held in the case study where every selected criterion by itself was unique, independent from other criteria, numerically measured, and continuous across the example case area.

Normalization of the values of the spatial criteria was a requirement step without which incorporating the criteria into the model would have been impossible. The normalization process should be performed after removing reserved areas from criteria layers; otherwise, the results would be less liable and the normalization could include unneeded values. This is why MC-Analyst prompted the user to enter the constraints raster layer first before choosing the criteria raster layers. The normalization was achieved using the Score Range Procedure (SRP) and its opposite that converted the criteria values into the range of [0-1].

Weighting the criteria had the most effective impact on the results after the normalization process. Altering the weights among different criteria of solar CSP, distance to transmission lines, and distance to roads demonstrated how switching the weights would change the result in WLC method. MC-Analyst validation for weights range (validation that weights must be between 0-1) added liability that the user would not be allowed to use invalid weights.

Employing GeoTools to develop customized GIS applications was possible when a developer is familiar with the library which actually was the major concentration of this study along with the implementation of the WLC model.

GeoTools provided many core GIS functionality such as reading and writing shape files and raster data, topology predicates, map algebra, and many other operations. As a matter of fact, to master the usage of GeoTools it would take much more time than had been assigned to prepare this study (about 3 months of coding).

However, as it was the case with any FOSS software, GeoTools had some lack of documentation. Its documentation is still in need of further development at the level of API and tutorials. The documentation did not provide usage examples for all the implemented functionality. For example, there have been many attempts during the study to write metadata to the output .TIF raster files that were created along different phases of the WLC but unfortunately none of them worked. Even though the documentation of GeoTools API included a class named GeoTiffWriter that writes geo-referenced .TIF files and had methods related to the creation of metadata, no examples were provided on how to use this class. Many attempts to work around the problem failed during the coding process. Blogs and forums have been checked out for this but there was no clear usage for the class at the time of development. Even a question was posted concerning this problem on stackoverflow.com without any answer (http://gis.stackexchange.com/questions/77905/missing-srs-and-nodata-when-writing-tiff-rasters-in-geotools).

An example on how proprietary software provides some features that are not available in FOSS software was the resampling techniques used in GeoTools. At the time of development there have been nine interpolation methods available (while ArcGIS Desktop 10.0 had 4 resampling techniques in its Resample toolbox) for resampling in GeoTools but there was still one important method missing which was the majority method (implemented in ArcGIS Desktop 10.0).

Among the remarkable issues faced during the development was the reading of SRS of rasters. Reading the right SRS of rasters was not as easy as mentioned in the image tutorial found in GeoTools website. Actually, reading the right SRS was only possible when the metadata citation format was already known to the developer or by setting the DefaultCRS hint when reading rasters.

In spite of the aforementioned issues, using GeoTools to develop a custom desktop application helped create a business oriented application with small costs and efforts. Had MC-Analyst been developed using a proprietary framework such as ArcGIS, there would have been costs that might not be affordable in all situations. In addition to the work provided by GeoTools team, source code contributions from

GeoTools developer users could add to the advantages of FOSS in general and GeoTools in specific.

According to the type of integration between GIS and MCDA systems defined by Malczewski (2010), MC-Analyst falls into the loosely-coupled integration of MC-SDSS. MC-Analyst received the criterion raster files from a GIS system (ArcMap), then, from within MC-Analyst, the analysis starts as an independent process. In order to make a further step towards the full integration, the WLC model developed in GeoTools could be integrated in one of FOSS GIS desktop applications such as uDig. This would save huge efforts especially at the level of Graphical User Interface (GUI) and the basic functional GIS tasks since other functionality provided by the hosting application such as data conversion, reporting, and map displaying could be used directly for the benefit of the WLC model. This also would allow the focus to be only on implementing the WLC.

8 Conclusions

GIS-MCEA could be applied in any discipline that seeks spatial decision-making analysis. As long as the criteria involved in the problem had spatial aspects and conformed to the specifications recommended by the selected MCEA model, incorporating these criteria into that MCEA model using GIS was possible. Thus, while MCEA model structured the problem, GIS provided the tools to run the model spatially.

MCEA models were created by sequential phases where the output of every phase was the input of the successor one. Even though different MCEA models could differ in the methodology approached for weights definition, any MCEA problem required the representation of experts, researchers, or stakeholders preferences in the questioned discipline through weights definition. An important point was that those individuals should have full understanding of the problem and are able to evaluate how one suggested solution of the problem outweighed another one in order to apply the WLC model properly.

Using an open source GIS software library was a sufficient core tool to build GIS models that produce liable results. Deeper investigation of these libraries would bring better experience on how to use them in a productive and sufficient manner. Before choosing a GIS software library to start with, it would be better for the developer to get her/himself acquainted enough with other GIS software libraries in order to incorporate them into the work. Another reason for this is that no software library would provide all the needed functionality in a fully productive application. Reading, writing, and processing GIS data, further presentation, analysis, and reporting modules would be needed in a fully productive application.

Contacting developers with experience in open source GIS development either directly, through mailing lists, blogs, or forums, could help the developer narrow his choice of the open source GIS software libraries to use and to solve the problems faced during the development.

Unlike proprietary software that usually provides complete and full documentation for the development libraries, FOSS GIS libraries might suffer from lack of documentation and support. A GIS software developer should account for this before proceeding to work with any library.

Future development of MC-Analyst could be added in order for MC-Analyst to be used efficiently in a production environment. For example, more MCEA models such as AHP and the Ideal/Reference Point could be implemented besides the WLC model. In addition, a richer interface along with a reporting module could be developed. Using a development kit of a developed FOSS GIS desktop application such as uDig or gvSig could be a good approach to incorporate the MCEA developed model in the developed work of those applications. This approach would also allow the use of the core functionality developed by those desktop applications for data geo-processing and conversion. It also would increase the coupling of GIS with MCDA model to produce a fully-integrated MC-SDSS.

References

Colorado Energy, Colorado Renewable Resource Database. Available from <http://www.coloradoenergy.org/corrd/default.asp>. [1 March 2014].

Eldrandaly K., 2011. Developing a GIS-Based MCE Site Selection Tool in ArcGIS Using COM Technology. The International Arab Journal of Information Technology No.3, May 2013.

GeoTools, GIS FOSS software library. Available From <www.geotools.org>. [1 March 2014].

Greene R., Devillers R., Luther J. E., and Eddy B. G., 2011. GIS-Based Multiple-Criteria Decision Analysis. Geography Compass, (2011, 5/6): 412–432.

Hall G. B. and Leahy M. G., 2008, Advances in Geographic Information Science: Open Source Approaches in Spatial Data Handling. Springer-Verlag Berlin Heidelberg, ISBN: 978-3-540-74830-4.

Hepner G. F., 1984. Using of value functions as a possible suitability scaling procedure in automated composition mapping. Professional Geographer 36, 468-472.

Janke J. R., 2010. Multi-criteria GIS modelling of wind and solar farms in Colorado. Renewable Energy, Volume 35, Issue 10, October 2010, Pages 2228-2234.

Jankovic D. and Milidragovic R., 2013. Integration open source GIS software for improving decision-making in local communities. ACTA TECHNICA CORVINIENSIS – Bulletin of Engineering, ISSN 2067-3809, Tome VI (Year 13), Fascicule 4 [October–December], Pages: 73-76.

Java Platform, Java Standard Edition. Available from <http://www.oracle.com/technetwork/java/javase/overview/index.html>. [1 March 2014].

Malczewski J., 1999. GIS and Multicriteria Decision Analysis. John Wiley & Sons, New York, 1999. ISBN: 0-471-32944-4.

Malczewski J., 2000. Review Article on the Use of Weighted Linear Combination Method in GIS: Common and Best Practice Approaches. Transaction in GIS 4(1), 5-22.

Malczewski J., 2006. A GIS-based multi-criteria decision analysis: A survey of the literature. International Journal of Geographical Information Science Vol. 20, No. 7, 703–726.

Malczewski J., 2010. Multiple Criteria Decision Analysis and Geographic Information Systems. Trends in Multiple Criteria Decision Analysis, International Series in Operations Research and Management Science, 2010, VOL: 142, 369-395.

Nagy, D., Yassin, A. M., & Bhattacherjee, A., 2010. Organizational adoption of open source software. Communications of the ACM, 53(3), 148–151.

Open Source Geospatial Foundation (OSGeo), 2013, http://www.osgeo.org/. [1 March 2014]

Open Source Initiative Homepage. Available from <http://opensource.org/>. [1 March 2014]

Oracle. Available from <http://www.oracle.com/us/technologies/java/overview/index.html>. [1 March 2014]

Ramsey P, 2007. The State of open Source GIS. Refractions Research Inc. Available from <http://www.refractions.net/expertise/whitepapers/opensourcesurvey/survey-open-source-2007-12.pdf>. [1 March 2014]

Rinner C. and Voss S., 2013. MCDA4ArcMap – An Open-Source Multi-Criteria Decision Analysis and Geovisualization Tool for ArcGIS 10. Feature Article, Cartouche, Newsletter of the Canadian Cartographic Association, Number 86, Winter/Spring 2013, pp. 12-13.

Solar and PV Data. Available from <http://solargis.info/doc/103>. [1 March 2014].

State of Colorado, Department of Local Affairs (DLA), 2010. Available from <http://dola.colorado.gov/gis-cms/content/census-resources>. [1 March 2014].

Steiniger S. and Bocher E., 2009. An overview on current free and open source desktop GIS developments. International Journal of Geographical Information Science, Vol. 23, No. 10, October 2009, 1345–1370.

Steiniger S. and Hay G. J., 2009. Free and open source geographic information tools for landscape ecology. Ecological Informatics 4 (2009) 183–195.

Steiniger S. and Hunter A. J. S., 2012a. Chapter 15: Free and Open Source GIS Software for Building a Spatial Data Infrastructure. Geospatial Free and Open Source Software in the 21st Century, Lecture Notes in Geoinformation and Cartography. Springer-Verlag Berlin Heidelberg 2012, DOI 10.1007/978-3-642-10595-1_15, 247-261.

Steiniger S. and Hunter A. J. S., 2012b. The 2012 free and open source GIS software map - A guide to facilitate research, development, and adoption. Computers, Environment and Urban Systems 39, 2013, 136–150.

Turton I., 2008. Chapter 8: GeoTools. Open Source Approaches in Spatial Data Handling. Advances in Geographic Information Science 2. Springer-Verlag Berlin Heidelberg 2008, Pages: 153-169.

United States Census Bureau, US Department of Commerce, Topologically Integrated Geographic Encoding and Referencing (TIGER Products), 2013. Available from <http://www.census.gov/geo/maps-data/data/tiger.html>. [1 March 2014].

Appendix A

Listed below in A.1 and A.2 are some GeoTools classes and interfaces that were mostly used in MC-Analyst. The description of every class or interface is based on its usage in MC-Analyst and the documentation of GeoTools API found in http://docs.geotools.org/latest/javadocs/.

A complete NetBeans project of MC-Analyst was associated with the thesis.

A.1 GeoTools Raster Data support

1. AbstractGridCoverage2DReader: a class that is used to access the actual GridCoverage2D object that reads the underneath raster cell data.
2. AbstractGridFormat: a class that represents the upper level handler for raster (grid) data.
3. Envelope2D: a spatial extent object representation in GeoTools.
4. GridCoverage2D: a class that represents the raster object with its spatial coordinate system. It utilizes the java.awt.image.Raster class as the object holding its actual cell values and binds it to a spatial extent through an Envelope2D object.
5. GridCoverageFactory: a class that creates a GridCoverage2D.
6. GridFormatFinder: the basic class that finds the format of a raster file given its File object. It returns an AbstractGridFormat object by calling its method findFormat();
7. Operations.DEFAULT: a class for raster processing in GeoTools. This class implements most of the important geo-processing tasks for rasters such as cropping, adding, multiplying, and resampling.
8. Package java.awt.image.*: a package of Java SE SDK. It provides essential classes and interfaces that are important for the work of GeoTools image model. It includes classes such as Raster, WritableRaster, RenderedImage, DataBuffer, DataBufferDouble, PixelInterleavedSampleModel, and SampleModel which are used to implement the raster grid implementation.

A.2 GeoTools Vector data support

1. FeatureRasterizer: this class was developed for GeoTools by two of its users, Steve Ansari and Ferdinando Villa, who work for NOAA's National Climatic Data Centre. The source code was published to the open source community. The class converts vector features into raster (http://mvn.idelab.uva.es/gwc-modules/xref/org/geotools/process/feature/FeatureRasterizer.html).

2. FeatureReader: a Java interface that is used to sequentially read the features from a data source.
3. Geometry: a class of the lead topology suite, JTS. It represents the geometry description. One of its usage is as the type of "the_geom" attribute of features. This attribute represents the spatial aspect of features in a shape file. Geometry class implements different methods that are used for spatial predicates such as contains() and distance(). It also implements the buffer() method which buffers a feature by a passed value using the unit of its coordinate reference system.
4. SimpleFeature: a class that represents a single feature. For example, a city feature of the city features stored in a shape file.
5. SimpleFeatureBuilder: a class that is used to build features out of a SimpleFeatureType. It builds a feature by calling its method of buildFeature();
6. SimpleFeatureType: a Java interface that represents a description of the schema of feature class in a shape file.
7. SimpleFeatureTypeBuilder: a class responsible for creating a feature class via defining its geometry type, spatial reference system and attributes. After setting all such properties, the feature type is created by calling buildFeatureType() method. This method returns an object of SimpleFeatureType.

A.3 Overview of MC-Analyst WLC Java Classes

Different classes of MC-Analyst implemented different functionality. The most important code portions of MC-Analyst were implemented by the following classes:

1. mceawlc.gui.JPBuffer: is the panel included in the dialogue of Multi-Buffering and Rasterization. The core work of this dialogue was implemented by method bttnBufferActionPerformed.
2. mceawlc.gui.JPReclass: is the panel included in the dialogue of Reclassification. The core work of this dialogue was implemented by method reclass().
3. mceawlc.gui.JPWeights: is the panel included in the WLC wizard in the step where the weights are defined, validated (method validateControls) and set to the model layers (method writeToModel).
4. mceawlc.wlcmodel.WLCEngine: is the engine class where the global work of WLC is implemented. Method applyConstraintsRaster sets the global variables, sets NoData value of model layers, and multiplies the model layers by the constraints layer. Method applyWeightsAndSum multiplies each model layer with its assigned weight and finally combines the weighted layers.

5. mceawlc.processing.Geoprocessoer: is the most important class of all the classes where all the core work of the application was implemented. Core works such as rasters reading, reclassification, and writing were implemented in this class.

Series from Lund University

Department of Physical Geography and Ecosystem Science

Master Thesis in Geographical Information Science (LUMA-GIS)

1. *Anthony Lawther:* The application of GIS-based binary logistic regression for slope failure susceptibility mapping in the Western Grampian Mountains, Scotland. (2008).
2. *Rickard Hansen:* Daily mobility in Grenoble Metropolitan Region, France. Applied GIS methods in time geographical research. (2008).
3. *Emil Bayramov:* Environmental monitoring of bio-restoration activities using GIS and Remote Sensing. (2009).
4. *Rafael Villarreal Pacheco:* Applications of Geographic Information Systems as an analytical and visualization tool for mass real estate valuation: a case study of Fontibon District, Bogota, Columbia. (2009).
5. *Siri Oestreich Waage:* a case study of route solving for oversized transport: The use of GIS functionalities in transport of transformers, as part of maintaining a reliable power infrastructure (2010).
6. *Edgar Pimiento:* Shallow landslide susceptibility – Modelling and validation (2010).
7. *Martina Schäfer:* Near real-time mapping of floodwater mosquito breeding sites using aerial photographs (2010)
8. *August Pieter van Waarden-Nagel:* Land use evaluation to assess the outcome of the programme of rehabilitation measures for the river Rhine in the Netherlands (2010)
9. *Samira Muhammad:* Development and implementation of air quality data mart for Ontario, Canada: A case study of air quality in Ontario using OLAP tool. (2010)
10. *Fredros Oketch Okumu*: Using remotely sensed data to explore spatial and temporal relationships between photosynthetic productivity of vegetation and malaria transmission intensities in selected parts of Africa (2011)
11. *Svajunas Plunge:* Advanced decision support methods for solving diffuse water pollution problems (2011)
12. *Jonathan Higgins:* Monitoring urban growth in greater Lagos: A case study using GIS to monitor the urban growth of Lagos 1990 - 2008 and produce future growth prospects for the city (2011).
13. *Mårten Karlberg:* Mobile Map Client API: Design and Implementation for Android (2011).
14. *Jeanette McBride:* Mapping Chicago area urban tree canopy using color infrared imagery (2011)

15. *Andrew Farina:* Exploring the relationship between land surface temperature and vegetation abundance for urban heat island mitigation in Seville, Spain (2011)

16. *David Kanyari*: Nairobi City Journey Planner An online and a Mobile Application (2011)

17 *Laura V. Drews:* Multi-criteria GIS analysis for siting of small wind power plants - A case study from Berlin (2012)

18 *Qaisar Nadeem:* Best living neighborhood in the city - A GIS based multi criteria evaluation of ArRiyadh City (2012)

19 *Ahmed Mohamed El Saeid Mustafa:* Development of a photo voltaic building rooftop integration analysis tool for GIS for Dokki District, Cairo, Egypt (2012)

20 *Daniel Patrick Taylor*: Eastern Oyster Aquaculture: Estuarine Remediation via Site Suitability and Spatially Explicit Carrying Capacity Modeling in Virginia's Chesapeake Bay (2013)

21 *Angeleta Oveta Wilson:* A Participatory GIS approach to *unearthing* Manchester's Cultural Heritage '*gold mine'* (2013)

22 *Ola Svensson:* Visibility and Tholos Tombs in the Messenian Landscape: A Comparative Case Study of the Pylian Hinterlands and the Soulima Valley (2013)

23 *Monika Ogden:* Land use impact on water quality in two river systems in South Africa (2013)

24 *Stefan Rova:* A GIS based approach assessing phosphorus load impact on Lake Flaten in Salem, Sweden (2013)

25 *Yann Buhot:* Analysis of the history of landscape changes over a period of 200 years. How can we predict past landscape pattern scenario and the impact on habitat diversity? (2013)

26 *Christina Fotiou:* Evaluating habitat suitability and spectral heterogeneity models to predict weed species presence (2014)

27 *Inese Linuza:* Accuracy Assessment in Glacier Change Analysis (2014)

28 *Agnieszka Griffin:* Domestic energy consumption and social living standards: a GIS analysis within the Greater London Authority area (2014)

29 *Brynja Guðmundsdóttir:* Detection of potential arable land with remote sensing and GIS - A Case Study for Kjósarhreppur (2014)

30 *Oleksandr Nekrasov:* Processing of MODIS Vegetation Indices for analysis of agricultural droughts in the southern Ukraine between the years 2000-2012 (2014)

31 *Sarah Tressel*: Recommendations for a polar Earth science portal

in the context of Arctic Spatial Data Infrastructure (2014)

32 *Caroline Gevaert:* Combining Hyperspectral UAV and Multispectral Formosat-2 Imagery for Precision Agriculture Applications (2014).

33 *Salem Jamal-Uddeen:* Using GeoTools to implement the multi-criteria evaluation analysis - weighted linear combination model (2014)