HELSINGIN YLIOPISTO
HELSINGFORS UNIVERSITET
UNIVERSITY OF HELSINKI

Master's thesis
Geography
Geoinformatics

DEVELOPING A GLOBAL LOCATION OPTIMIZATION MODEL FOR UTILITY-SCALE SOLAR POWER PLANTS

Laura Kauria

2016

Instructor:
Tuuli Toivonen

UNIVERSITY OF HELSINKI
FACULTY OF SCIENCE
DEPARTMENT OF GEOSCIENCES AND GEOGRAPHY
DIVISION OF GEOGRAPHY

PL 64 (Gustaf Hällströmin katu 2)
00014 University of Helsinki

HELSINGIN YLIOPISTO
HELSINGFORS UNIVERSITET
UNIVERSITY OF HELSINKI

Faculty	Department
Faculty of Science	Department of Geosciences and Geography

Author
Laura Kauria

Title
Developing a global location optimization model for utility-scale solar power plants

Subject
Geography (Geoinformatics)

Level	Month and year	Number of pages
Master's thesis	April 2016	92

Abstract

The purpose of this Master's thesis was to create a new model for screening possible optimal locations for utility-scale solar power plants (i.e. solar parks, solar power stations and solar farms) in larger city areas. The model can be used as a part of a decision making when examining site potentiality in a particular city of interest. The model includes forecasts for the year 2040. The main questions of the thesis are as follows: 1) What are the main criteria for a good location for a utility-scale solar power plant and 2) how to build a geographic information system (GIS) model for solar power plant location optimization? Solar power plants provide an alternative to producing renewable energy due to the enormous distribution potential of solar energy. A disadvantage of utility-scale solar energy production is the fact that it requires larger areas of land than the more traditional power plants. Converting land to solar farms might threaten both rich biodiversity and food production, which is why these factors are included in the model.

In this study, methods from the field of geographic information science were applied to quantitative location optimization. Spatial analytics and geostatistics, which are effective tools to narrow down optimal geographical areas, were applied for finding optimal locations for solar power plants, especially in larger city regions. The model was developed by an iterative approach. The resulting model was tested in Harare (Zimbabwe), Denver (United States) and Helsinki (Finland).

The optimization model is based on three raster datasets that are integrated through overlay analysis. The first one contains spatial solar radiation estimates for each month separately and is derived from a digital elevation model and monthly cloud cover estimates. The resulting radiation estimates are the core factor in estimating energy production. The second and the third dataset are two separate global datasets, which were used to deal with land use pressure issues. The first of these is a hierarchically classified land systems model based on land cover and intensiveness of agriculture and livestock, while the second is a nature conservation prioritization dataset, which shows the most important areas for conserving threatened vertebrate species. The integration of these datasets aims to facilitate smart and responsible land use planning and sustainability while providing information to support profitable investments.

The model is based on tools implemented in the ArcGIS 10 software. The Area solar radiation tool was used for calculating the global and direct radiation for each month separately on clear sky conditions. An estimate of the monthly cloud coverage was calculated from 30 years' empirical cloud data using a probability mapping technique. To produce the actual radiation estimates, the clear sky radiation estimates were improved using the cloud coverage estimates. Reclassifying the values from land use datasets enabled the exclusion of unsuitable areas from the output maps. Eventually, the integration and visualization of the datasets result in output maps for each month within a year. The maps are the end product of the model and they can be used to focus decision making on the most suitable areas for utility-scale solar power plants.

The model showed that the proportion of possible suitable areas was 40 % in Harare (original study area 40 000 km^2), 55 % in Denver (90 000 km^2) and 30 % in Helsinki (10 000 km^2). This model did not exclude areas with low solar radiation potential. In Harare, the yearly variation in maximum radiation was low (100 kWh/m^2/month), whereas in Denver it was 2.5-fold and in Helsinki 1.5-fold. The solar radiation variations within a single city were notable in Denver and Harare, but not in Helsinki. It is important to calculate radiation estimates using a digital elevation model and cloud coverage estimates rather than estimating the level of radiation in the atmosphere. This spatial information can be used for directing further investigations on potential sites for solar power plants. These further investigations could include land ownership, public policies and investment attractiveness.

Keywords
Solar energy, utility-scale solar power plant, location optimization, GIS modelling, land use pressure

Where deposited
E-thesis (University of Helsinki)

Additional information

HELSINGIN YLIOPISTO
HELSINGFORS UNIVERSITET
UNIVERSITY OF HELSINKI

Tiedekunta		Laitos
Matemaattis-luonnontieteellinen tiedekunta		Geotieteiden ja maantieteen laitos
Tekijä		
Laura Kauria		
Työn nimi		
Developing a global location optimization model for utility-scale solar power plants		
Oppiaine		
Maantiede (geoinformatiikka)		
Työn laji	Aika	Sivumäärä
Pro gradu -tutkielma	Huhtikuu 2016	92

Tiivistelmä

Tässä Pro gradussa luodaan uusi malli, joka tuottaa aurinkovoimaloille (myös aurinkopuisto ja aurinkovoimalaitos) mahdollisia optimaalisia sijainteja erityisesti suurkaupunkialueilla. Mallia voidaan käyttää osana päätöksentekoinfrastruktuuria arvioitaessa tietyn kaupungin sijaintien soveltuvuutta aurinkovoimalalle. Mallissa käytetään vuoden 2040 ennusteita. Tutkimuskysymykset ovat: 1) Mitkä ovat aurinkovoimalalle hyvän sijainnin pääkriteerit ja 2) miten rakennetaan GIS-malli aurinkovoimalan sijainnin optimointiin? Aurinkovoimalat tarjoavat vaihtoehdon uusituvan energian tuottamiseen, koska auringonsäteilyn hyödyntämispotentiaali on merkittävää suuressa osassa maailmaa. Aurinkoenergiatuotannon negatiivisena puolena on ainakin se, että voimalat vaativat isoja maa-aloja verrattuna perinteisimpiin energiavoimaloihin. Tulevaisuuden aurinkovoimaloiden maa-alatarve voi heikentää rikasta biodiversiteettiä sekä ruoantuotantoa, minkä vuoksi nämä tekijät ovat sisällytettynä malliin.

Tässä tutkimuksessa käytetään metodeja geoinformatiikan tieteenalalta, joita sovelletaan kvantitatiiviseen sijainninoptimointiin. Geospatiaalinen analytiikka ja geostatistiikka ovat tehokkaita työkaluja optimaalisten sijaintien kohdentamiseen. Tässä tutkimuksessa näitä työkaluja on sovellettu aurinkovoimaloiden optimaalisten sijaintien löytämiseen laajennetuilla kaupunkialueilla. Mallin luomisprosessin tukena on käytetty iteratiivista lähetymistä. Mallin toimivuutta on testattu Hararessa (Zimbabwe), Denverissä (Yhdysvallat) ja Helsingissä (Suomi).

Malli perustuu kolmeen paikkatietoa sisältävään rasteriaineistoon, jotka ovat integroitu päällekkäisanalyysia käyttäen. Ensimmäinen aineistokokonaisuus sisältää säteilyestimaatit jokaiselle 12 kuukaudelle erikseen. Nämä aineistot ovat johdettu käyttäen globaalia korkeusmalliaineistoa ja kuukausittaisia pilvisyysestimaatteja. Nämä mallista johdetut auringonsäteilyestimaatit ovat ydintietoa aurinkoenergiatuotantoa arvioitaessa. Maakäytön paine sisällytetään malliin kahden eri aineiston avulla. Ensimmäinen on globaali maankäytön intensiivisyyttä kuvaava ennustemalli vuodelle 2040. Aineisto perustuu maanviljelyn ja karjankasvatuksen intensiteettiin sekä maanpeiteluokkaan. Toinen globaali aineisto priorisoi luonnonsuojelualueverkostoihin lisättäviä alueita siten, että mahdollisimman moni uhanalainen selkärankainen säästyisi. Näiden kolmen aineiston yhteenliittymän tarkoitus on helpottaa viisasta ja vastuullista sijaintisuunnittelua aurinkovoimaloille ja tuottaa informaatiota kannattavien investointien tekemiseen.

Malli pohjatuu työkaluihin, jotka ovat käytettävissä ArcGIS 10 -ohjelmistossa. *Area solar radiation* –työkalulla lasketaan säteilyestimaatti jokaiselle kuukaudelle pilvettömän taivaan olosuhteilla. Pilvisyysestimaatti lasketaan 30 vuoden tilastoaineistosta käyttäen. Ennustelukuna käytetään yksisuuntaisen luottamusvälin ylärajaa pilvisyyden keskiarvolle todennäköisyydellä 97,5 %. Säteily- ja pilviestimaateista johdetaan lopullinen globaali säteilyestimaatti. Muut aineistot uudelleenluokitellaan sopiviksi ja epäsopiviksi alueiksi kirjallisuuteen perustuen, jonka jälkeen aineistojen integrointi ja tuloskarttojen visualisointi tehdään. Tuloskarttoja voi hyödyntää päätöksenteon tukena siten, että aurinkoenergiavoimalan sijoittamisen tutkimusvarat kohdennetaan potentiaalisimmille alueille.

Mallin tuloksena potentiaalisten alueiden pinta-alat tippuivat Hararessa 40 % (tutkimusalue 40 000 km^2), Denverissä 55 % (90 000 km^2) ja Helsingissä 30 % (10 000 km^2). Hararen vuosittainen säteilyvaihtelu oli suhteellisen vähäistä (100 kWh/m^2/kk). Denverissä erot olivat 2,5-kertaiset ja Helsingissäkin 1,5-kertaiset verrattuna Harareen. Tutkimusalueiden sisäiset vaihtelut olivat merkittäviä Denverissä ja Hararessa johtuen maanpinnan korkeuseroista. Koska temporaalista ja spatiaalista vaihtelua säteilyssä tulosten perusteella on, on säteilyn tarkka estimointi korkeusmallia käyttäen tarpeellista. Mallia on syytä kehittää, jotta sillä pystyttäisiin kohdentamaan tarkemmin potentiaaliset voimalasijainnit. Tämä vaatii paikallistason paikkatietoja, muun muassa maanomistajuuksista, ja maakohtaista tietoa, esimerkiksi poliittisista päätöksistä ja investointien houkuttelevuudesta kohdemaassa.

Avainsanat
Aurinkoenergia, aurinkovoimala, sijainnin optimointi, GIS-mallintaminen, maankäytön paine

Säilytyspaikka
E-thesis (Helsingin yliopisto)

Muita tietoja

Acknowledgements

First of all, thank you Tuuli for being a supportive supervisor and giving constructive feedback whenever needed. The subject took its better shape and starting point after you emphasized the importance of land use issues alongside the magnitude of solar radiation. I want to thank Fortum foundation for the personal scholarship which enabled full concentration to the thesis.

Special thanks to my big sister Elisa who forced me to write on troubled days, wisely challenged my deduction and reminded me hard-headedly the realistic scale of this thesis if I was digging way too deep. The progress of the project was also secured by an amazing “rocket team”, social network, consisting of Master’s thesis writers and eventually close friends Sofia, Iiris and Anniina. Other helpers were my other family members. Thanks to Anni, my twin sister, for explaining the logic of the electricity market and getting excited about the topic. Thanks to Minna, my big sister, for an excellent brainstorm concerning the statistics of the cloud model and too few coffee sessions in Kumpula. Also a big thanks to my parents.

The academic credibility of the thesis increased thanks to Samuli for your straightforward comments regarding the research questions and help with assimilating what really is methodology. I want to thank my fellow student Johannes who as my peer gave concrete development ideas to the methodology in the early phase. Thank you Vuokko for hardcore help with GIS and getting tanned. Thanks to Laura for instructing me on thin film technology applications from the field of solar energy technology. It is not included in my thesis, but it energized me knowing what is happening in the field and what is impacting the implementations of solar energy. Let’s continue jogging. Natalia, Lucas, Mirka and Mert, thank you for your language revision and priceless help making the text readable. I express my deepest gratitude to my spouse Mert who has supported me unyieldingly through the swamps of writing and pondering sessions. As a geographer and energy-enthusiast yourself you made great discussions possible. Thank you for your time.

Helsinki, 20th of April in 2016

Laura Kauria

CONTENTS

CONCEPTS

Power	Amount of energy per time, unit e.g. kilowatt-hour (kWh)
MW (unit of power)	Power (unit) in the utility industry is defined in megawatts i.e. billion watts
Solar power	Energy from the sun that is converted into electrical energy
Solar power plant	Conversion of sunlight into electricity by technologies such as photovoltaics (PV), concentrator photovoltaics (CPV) or concentrated solar thermal (CSP)
Photovoltaics (PV)	Method to convert solar energy into electricity by semiconducting materials
Electricity	Electrical energy, form of energy used to power electrical devices
Electric utility	Supplies electricity generally in a regulated market, utility is regulated by local or national authorities and is often public utility
Energy	Total amount of work done, capacity for activity (unit 1 joule or 1 watt)
Solar energy	Energy from the sun that can be harnessed e.g. by technology (active)
Solar radiation	Radiant energy emitted by the sun, electromagnetic energy
Global radiation	Total radiation that touches the surface of the ground, the sum of direct, diffuse and reflected radiation
Direct radiation	Also known as beam radiation, direct radiation that goes directly through the atmosphere from the sun to the surface of the earth; solar technologies can harness this
Irradiance (unit)	Power per area (unit e.g. W/m^2), radiation energy that touches the ground
Insolation	The amount of electromagnetic energy within a period of time (kWh/m^2)
Grid (electric)	Transmission network for delivering electricity from suppliers to consumers
Off-grid	Separated electric network from e.g. national electricity grid
Grid-parity	Occurs when an alternative source of energy generates power at a levelized cost of electricity (LCOE)
Extensive agriculture	System of crop cultivation that uses small amounts of labor, fertilizers and capital in relation to area cultivated; opposed to intensive agriculture, livestock production systems that have similar definition
Parochialism	(syn. Provincialism or localism), instead of considering the wider context, e.g. threatened vertebrates in global nature protection, local consideration is done e.g. species locally
GIS	Geographic information systems is e.g. a system that enables to store, manipulate, analyze and visualize spatial data

Reading tips:

There are concepts marked by *italics* within the text that will not get necessarily space in the text. Please check the meaning of the word from the glossary. These concepts focus mainly on the field of energy.

1 INTRODUCTION

1.1 Insights for solar production

The purpose of this Master's thesis is to create a model which can be used to find optimal locations for solar energy production for *utility-scale* solar power stations. This chapter gives background knowledge of solar energy production and its characteristics. Later on, the research problems and objectives as well as restrictions and assumptions are presented.

Electricity produced by solar energy is in a turning point. On the 2010s solar energy *grid-parity* events will occur reaching 75 - 90 % of total global electricity market (Breyer & Gerlach 2013). Grid-parity means that it is profitable to produce electricity from an alternative source of energy compared to buying the electricity from the electricity grid mainly produced by utility companies. Today solar energy production could be efficient enough in certain regions to compete with conventional ways of producing energy such as burning fossil fuels, and this is crucial for the diffusion of utility-scale solar power. A potential for utilizing solar energy is at another level because of intensive *irradiation* covers a great part of the world when compered potentiality to other source of renewable energy such as biomass or wind (Pogson et al. 2013). Also hydroelectric power has strong regional limitations by physical geography (Kemenes et al. 2007).

Nevertheless, solar *radiation* reaches almost every corner of the world the first limits of solar energy production are in the radiation magnitude and crave for large areas, as seen in Apple's case later. In Asselen's and Verburg's paper (2012) food production and land cover are conducted into a dataset called Land systems (LS). Their global dataset gives one tool to rationalize and regionally limit optimal solar energy sites from e.g. essential food production. Another global issue among competing land use forms is the need, and will, to protect biodiversity. This has been assessed by

Montesino Pouzols et al. (2014). These two different studies dealing with land use pressure are taken into account when examining optimal areas for utility-scale solar energy power plants.

In the energy part of this study, solar energy is the main focus, but when talking about energy dismissing *electricity* is difficult. The most typical form of production from solar energy is electricity in addition to heat. As seen in Figure 1 both photovoltaics and solar thermal technology are suited for electricity production. A concerned unit in this thesis around solar radiation is insolation per area (kWh/m2), but when motivating the subject in this Introduction (Section 1) electricity (kWh) is often mentioned as end-product.

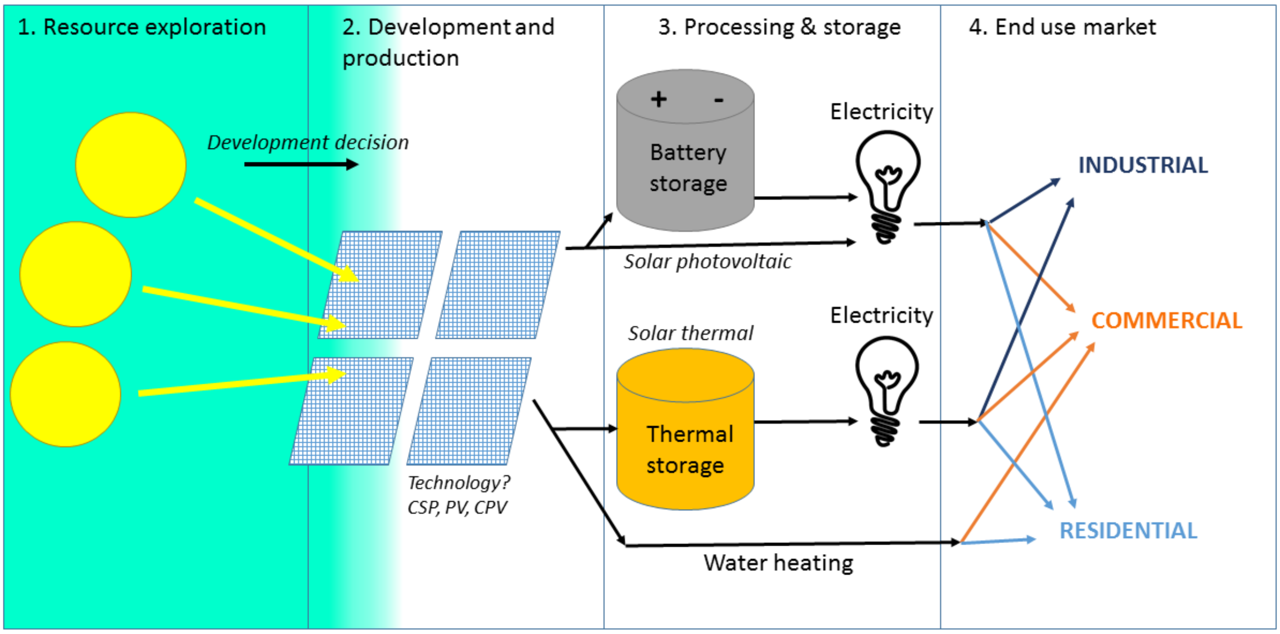


Figure 1. Phases of utilizing solar energy (A bulb icon by Kreativkolors / Freepik). The idea of the drawing is based on Australian Energy Resource Assessment: Chapter 10 Solar Energy (Carson 2014).

This thesis emphasizes on the first theme of the Figure 1, resource exploration. The latter phases (2-4) are scratched. There are different ways for *electric utility* to produce solar energy. *Photovoltaic* (later PV) cells convert solar radiation into electricity and this is the most common but not the most efficient (Green et al. 2015) way of producing electricity from the sun. More efficient concentrated solar power (CSP) production is taking giant leaps in the utility world. Apple turned eyes towards solar energy production by investing in 850 million US dollars for CSP. This 25-year deal with a technology company First Solar is only one project operating at this scale. Capacity of the solar power station (also called as solar park, solar farm, solar power plant) is 130 megawatts (later MW), which is enough to power 60 000 homes in California. The footprint of the project is 11.7 square kilometers. (Bloomberg 2015) As the name suggests concentrated solar power (CSP) systems generate solar power by concentrating solar thermal energy with lenses or mirrors onto a certain spot in order to boil liquid that powers a generator. Concentrating photovoltaic power plant (CPV) is the third technology to be mentioned.

On the background of the Figure 2 there is a solar thermal complex with high towers that are heated by radiation reflected from the mirrors. In the foreground there are stations that use concentrated solar technology: parabolic shaped collectors catch solar radiation. Apparently the land area would be excellent for food production as well.

Figure 2. Solvana Solar Power Station (CSP) near Sevilla, Spain. (https://commons.wikimedia.org /wiki/File%3AFoto_a%C3%A9re_de_solnovas_y_torre_junio_2010.jpg) CC-licensed picture.

Solar energy has been a hot topic in this decade. Electricity costs have rapidly fallen because of solar cell prices have fallen sharply. While industry scale price for one *Watt* (power unit) was 6 US dollars in the year 2010 the reference value in the 2014 was 2 to 3 US dollars (SEIA 2015). Yet solar energy covers only 1 % of global electricity generation and without a substantial price put on carbon dioxide emissions the future growth of solar energy utility is not certain (Schmalensee & Bulovic 2015). Solar energy is seen as one of the components among other renewable energy sources that is strategizing to mitigate climate change (Schmalensee & Bulovic 2015; Santangeli et al. 2015).

In general, increasing electricity production especially in developing countries is important because of lack of electricity which is a barrier for economic growth (The Economist 2015). Scarcity of electricity increases the price and makes it unaffordable for some consumers. Realistically solar energy, both residential-scale and utility-scale, could answer to this issue in some areas of the world. Solar energy has many downsides or at least concerns in utility-scale production. It is unpredictable and there are not yet efficient ways to store electricity, which may change in result of Tesla's five billion US dollars investment (Hayers 2015) in battery technologies. There are competitors but Tesla dominates the news. The governmental subsidies towards renewable energy confuse the conventional energy investors which are still for a long time mandatory for the stable power production. Solar station construction sites are increasing despite the downsides.

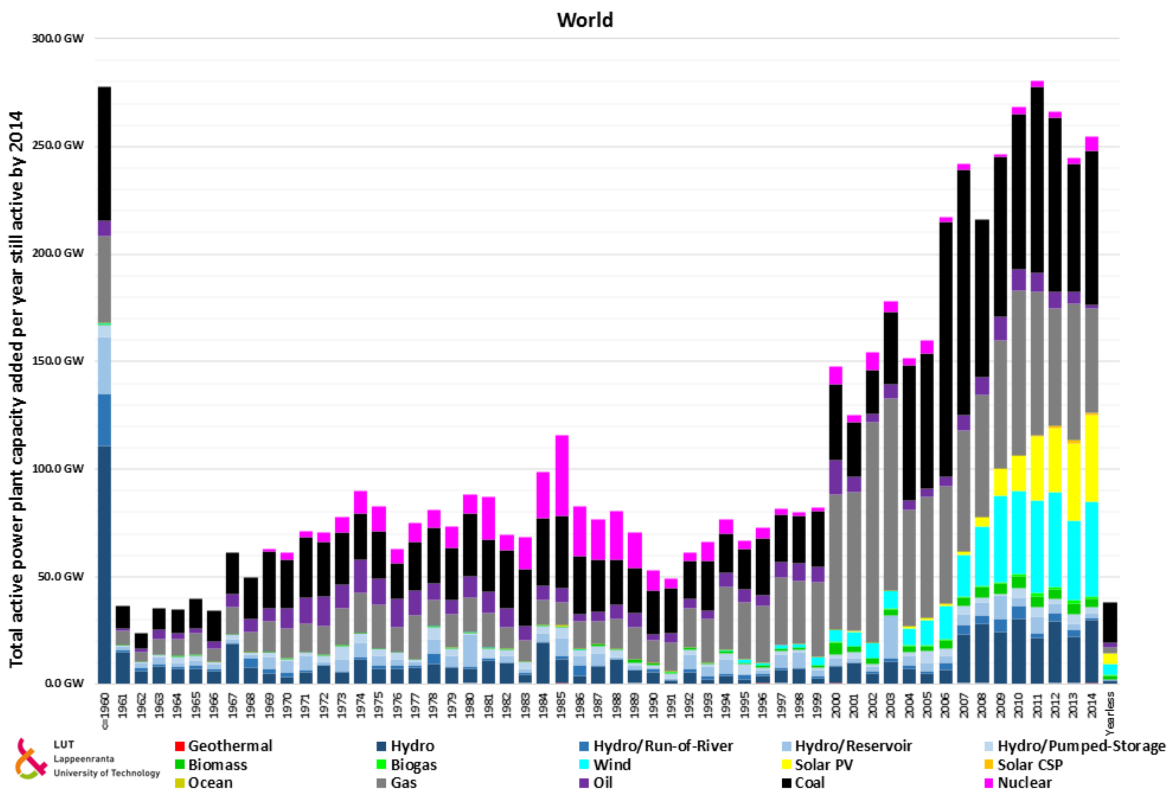


Figure 3. The total new power capacity built in the world. Figures are from the paper in preparation (Farfan & Breyer 2016)

A recent LUT study (Figure 3) indicates that 16 % (40,5 GW) of the total new power capacity built in the world in 2014 was solar photovoltaic power. New solar power capacity has remarkably increased every year after 2008. Figures are from the paper in preparation. (Farfan & Breyer 2016)

1.2 Solar energy steps from 1950s' to 2010s' – why it is current now?

Solar energy production has been under huge transition in the 21st century due to solar technology development. Efficient technology and political intervention programs in the form of subsidies have made solar energy production boom (Breyer & Gerlach 2013). To explain the boom of production it is reasonable to discuss the history of a certain technology, a solar photovoltaic cell (PV), as means of generating electricity.

Figure 4 gives a view on the four phases of PV use since 1955. We are now moving from the third to the rising fourth phase of solar PV era. In the year 1954 silicon-based solar cell was invented by Darryl Chapin, Calvin Fuller and Gerald Pearson. This innovation was crucial for supplying a spacecraft with power and was also the most cost-efficient solution at that time. The second phase of solar photovoltaic cell diffusion came during the oil crisis of the 1970s' because it could offer a cost-efficient way to produce electricity in *off-grid* areas. (Breyer & Gerlach 2013) The term off-grid indicates that the area or house in question is not connected to a broader or national electricity grid which delivers electricity from utility suppliers to consumers. For instance, on a faraway island there are autonomous homes which are off-grid that create energy as well as other utility services on their own.

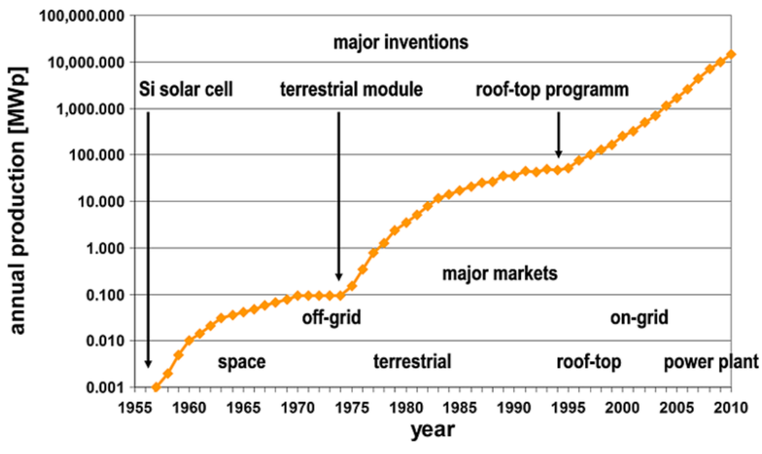


Figure 4. Waves of PV production (Breyer & Gerlach 2013).

As stated, we are now in the third phase. This phase is caused by political interventions such as rooftop programs and feed-in-tariffs, but also due to new technological innovations which bring higher conversion efficiency figures and the crash of solar photovoltaic cell technology prices. The conversion efficiency is a coefficient that solar radiation energy is enabled to convert into electricity. This time instead of off-grid the focus is on the grid areas since *grid-parity* by solar power has been reached in some areas that are connected to grid (electrical network). Grid-parity means that it is cheaper for e.g. households to create electricity by themselves rather than buying it from the main or national grid. In their paper Christian Breyer and Alexander Gerlach (2013) argue that grid-parity is a very important milestone for further diffusion of solar energy production. When the third phase is concentrated on houses of consumers or roofs of commercial shopping centers, a fourth phase is starting at the same time. Utility-scale companies are becoming more interested in photovoltaic power plants. (Breyer & Gerlach 2013) This thesis addresses this new phase of solar development. Utility-companies need decision-making tools for location optimization for utility-scale solar power plants, thus, location planners and decision makers are the target users of the model.

To illustrate the situation of solar energy production, Figure 6 shows the utility-scale solar power plants operating or under development in United States. Red points are power plants in action and yellow ones under development. Both are using different kind of PV technology and the blue ones are operating station using solar thermal technology. Turquoise points are power plants under development. The one popped up is a 55 MW PV solar power plant.

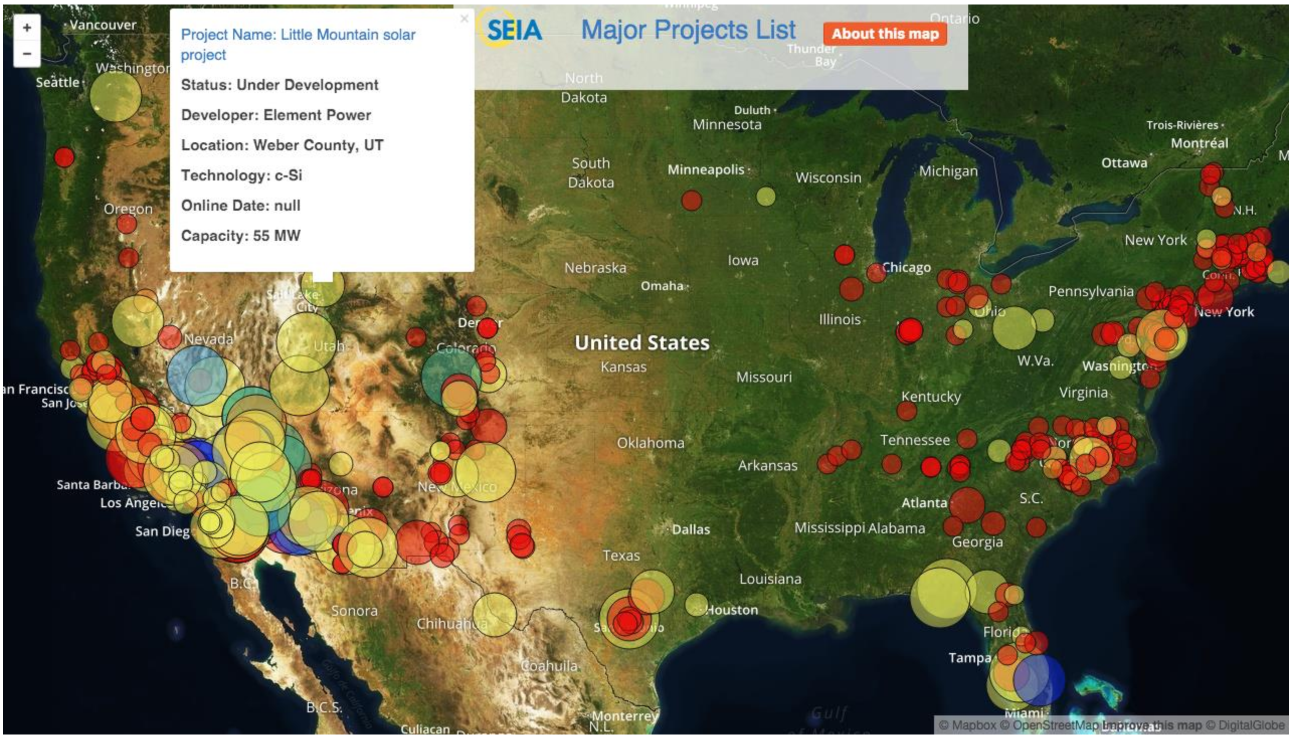


Figure 5. A web map that elucidates utility-scale solar power plant projects in the United States. (Solar Energy Industries Association 2015)

1.3 Research problems and objectives

The main focus of this study is to create a predictive and prescriptive model for optimizing solar station location anywhere in the world. The factors included in the model are solar radiation, globally classified land use pressure and suggested nature protection areas. Integrating these three factors aims to support smart land use planning done in a sustainable way. The information produced by this model can be a vital piece in the decision making infrastructure of utility-scale solar power plant planning.

In the light of present knowledge thousands of megawatts (thousands of hectares) of utility-scale solar power plants are going to be built in the coming decade. However, the primary problem is the increasing competition of land use. Land use conflicts since new land acquisitions can be soften by thorough decision-making support, where spatial analysis play a great role in creating information. The key role of this study is to optimize location for utility-scale solar stations. More specifically the research questions are:

Question 1: **What are the main criteria for a good location for utility-scale solar power plant?**
Question 2: **How to build a GIS model for solar power plant optimization?**

These questions are going to be answered through a comprehensive literature review including methodologies from geographical information studies, solar radiation potential, digital elevation model's (DEMs) role in in radiation estimating, site planning rom profitability and sustainability perspectives. To read more about DEM, see Section 3.2.1.1 Digital elevation model. The site planning literature and literature on solar energy are a base for most of the components of criteria screening. The actual model is built and its validation is done in three different greater city areas: Harare in Zimbabwe (20th parallel South), Denver in United States (40th parallel North) and Helsinki in Finland (60th parallel North). The research is focused in cities where the demand exists and the target cities in the sample were chosen because of their semi-large size, one and a half to 3 million people, and their latitudinal distribution. The model can be applied anywhere in the world with certain limitations. Its main targets are the greater city areas. *GIS* model, as it is used here, is a model that applies geospatial datasets and tools that handles the datasets. The model is demonstrated with a graphical flow chart.

1.4 Restrictions and assumptions

The model presented in the thesis is subject to certain limitations (Figure 6). The model can be implemented as a part of a larger model into a decision-making infrastructure of planning sites for solar stations.

The model does not take into account country-specific questions. These include domestic policies, attractiveness of investment, labor force situation and many other aspects that affects the decision making when placing solar energy production. A SWOT (strengths, weaknesses, opportunities and threats) analysis can be one tool for investigating and screening targeted countries for further investigations.

Factors taken into account

- **Radiation (12 months):**
  - global radiation (dir + dif + ref)
  - cloud cover
- **Land use pressure for long term (year 2040):**
  - Nature protection suggestions (year 2040)
  - Land systems (year 2040)
    - Food production intensiveness
    - Land cover

Factors not taken into account

- Country or city specific questions such as
  - Politics
  - Attractiveness of investments
  - Planning and zoning regulations
  - Land owning
- Demand for electricity
- Transmission network
- Soil type
- Particular size of a land; depends on station size and technology
- Flatness → topography
- Risks for nature hazards
- Agglomeration advantage
- Aesthetic value, landscape architecture

Figure 6. Factors included and excluded in the model.

Connectivity and accessibility of transmission network are key aspects for energy supply and these are one of factors that needs to be taken into account when the model is developed further. Connectivity as in topological space is closely related to accessibility concept and examining connectivity of the infrastructure between energy utility and consumers is crucial for total cost.

Demand for electricity and its location is excluded from this study. It is said that enormous subsidies towards renewable energy source disturb the energy market. This study will not cover market conditions and does not inspect the hypothesis that unstable production of renewable energy and renewable energy subsidies may eventually raise the prices by paralyzing traditional and reliable ways of producing energy. For example, in autumn of 2015 in Finland, it was hard to find investors in traditional energy venture to invest in a nuclear power plant. A state-owned company took on the project likely because a traditional energy project was not attractive enough to private investors.

One of the experimental validation areas for the test analysis is chosen by knowledge that in Harare, Zimbabwe, development is accelerating and demand for electricity is growing rapidly because of the growth in the standard of living and industry (Potts 2015). The World Bank is estimating that 4 % of economic growth across sub-Saharan Africa is being suppressed because of lack of electricity (The Economist 2015). The World Bank has collected data concerning the access to electricity by country. 40,5 % of population of Zimbabwe could access electricity in the year 2012. Zimbabwe ranked better than 30 countries in the whole world. For instance, a figure of Niger was 14,2 % and Kenya 23,0 %. (The World Bank webpage) This base information helps identify countries where demand for electricity exceeds supply.

Power production concentrates more around cities (Deshpande 2009) and therefore the model is built especially for the examination of the city areas. Validations of the model happens in Harare, Denver and Helsinki. Agglomeration and ready-made infrastructure that facilitates the geographical transfer of profits makes cities more attractive for investments as described by Pacione (2001) Harvey's (1985) idea of "circulation of capital" (cited in Pacione 2001).

The analysis has been made by using monthly frequency. The radiation has been calculated on the 15^{th} day of each month and the measuring frequency is half an hour. This gives adequate estimates of solar radiation hitting the ground and no daily frequency is needed. This is explained in more detail in the methodology chapter.

For the cloud cover estimate concluded in the study, monthly cloud cover data from 1980 to 2009 has been used. The estimate for each month is concluded using geostatistical analysis. In some areas the cloud coverage depends heavily on the time of the year.

The source code of the tool used for the solar radiation analysis is closed. Thus, it is not possible to explore its algorithms as a detailed code. The tools are justified through papers in the field of solar radiation research, not examining the source code and meteorological theories, which is beyond the scope of this analysis.

The actual model is executed by ArcGIS 10.2.1 software. Here are the specs of the computer environment to compare how long runs might take at certain environments: Windows 7, 8 GB of RAM, Intel i5, 3,3 GHz processor and graphical processor with memory of 2GB. Since ArcGIS 10.2.1 is not optimized for parallel runs it does not use all four CPU processors which would be effective. This is why the amount of processors does not matter, but speed of each processor definitely does. For larger study area parallel optimization is compulsory to save time. It took 20 hours for Area solar radiation tool to estimate global radiation for the study area of Harare (around 6 000 x 6 000 cells). Since Denver is far from the equator and the resolution is in degrees, the amount of cells increased dramatically. Because radiation model is a global model and takes into consideration all the cells in the study area for each individual cell value calculation, it took 100 hours to calculate Denver's study area with around 10 000 x 10 000 cells. The study area of Helsinki took 20 hours with higher resolution (10m x 10m) due to a national DEM and much smaller city area (40km * 100km) because the proximity to the coast. No other phases in the model is computing intensive.

1.5 Earlier studies

The site search modeling has challenged researchers such as geoinformaticians and other planners for a long time. A landscape architect, Ian McHarg (1920-2001) has been contributing enormously to the birth and development of GIS, especially from the planning approach. The McHarg's book "Design with nature" from the year 1969 pioneered the concept of ecological planning by overlaying ecological inventories using "layers", which later led to development of geographical information systems. Esri's founder Jack Dangermond is another contributor to environmental GIS whose background is as well on landscape architecture, and furthermore, computing. In his early career he studied integration possibilities of an overlay analysis. Especially, he extended McHarg's overlay method by integrating environmental and engineering planning issues. (Dangermond 1979, cited in Harvey 1997). On a more concrete level the next paragraphs tell about three types of earlier studies: solar potential, site optimization and ecological observations in site planning.

Solar energy potential (1) has been a hot topic in 21th century and it is one reason why numerous amounts of research of solar radiation by different methods have been done (e.g. Schmalensee & Bulovic 2015; Shahzad et al. 2014, Tovar-Pescador 2006; Ruiz-arias et al. 2009; Fu & Rich 1999; Watson & Hudson 2015; Brewer et al. 2015; Mei et al. 2015; Breyer et al. 2010; Pogson et al. 2013). The optimization application for sites (2) have also been of interests for planning departments and other utility companies and it has been studied with great variety (e.g. DeMers 2002b; Cova & Church 2000; Hastings et al. 2009; Williams 1985). Many studies include ecological and land use pressure aspects (3) to help site planning instead of factors regarding investment attractiveness (e.g. Brewer et al. 2015; Watson & Hudson 2015; Tsoutsos et al. 2005; DeFries & Rosenzweig 2010; van Asselen & Verburg 2012; Santangeli et al. 2015). In the discussion section in the end the success of the thesis in adding to earlier studies is reassessed.

Solar energy potential research has been distributed widely on different research fields as it is a current topic. Companies compete with each other to see who creates the best, most efficient and profitable product to conduct electricity efficiently or technology that retains energy: thermal storages and batteries. The side of technological development is excluded from this thesis and the focus of earlier studies will be in the resource exploration as seen in Figure 1.

The calculation of solar potential is a methodological question. The earlier studies are comprehensively compared and studied in the methodology section, Section 3.2.1.1 Digital elevation model, and in the theory section, Section 2.2.2 Measuring solar radiation. To sum up here the earlier studies regarding solar energy potential, solar radiation estimates based on the digital elevation model (DEM) are the most accurate and useful to optimize location for solar stations. DEM-based is needed especially in rugged topography areas. Other techniques such as interpolating the radiation values from weather stations or satellite image interpretation cannot compete in accuracy with DEM-based calculations in complex terrain areas.

Nowadays many free solar radiation web map service providers exist, e.g. City of Minneapolis and City of Espoo on their own regions, and Google's Project Sunroof (demo) has produced information of solar radiation potentiality. Google starts from the area of California. These applications show on a house-scale where solar panels should be installed and what is the magnitude of insolation on each spot. These could be used for spotting possible areas of utility companies as well but it takes some time to include global land use models in the services.

DeMers (2002) describes site planning as a prescriptive model since the site planning starts from the premise that the "cure" for land need is going to be found i.e. the best place is found. Cova and Church (2000) are listing criteria for site planning in their paper that handles continuity problems around a single-area site searching. Their geographical criteria for site are area suitability, cost, shape and proximity to surrounding geographical features. Otherwise, their approach is concentrated on contiguous problems which are not assessed in this study. Preferably suited to the subject, Land Evaluation and Site Assessment system, LESA, is created to support decision-making in The U.S. Department of Agriculture. The model describes and integrates geographical data that is in the interest of regional and local planners. When speaking of agricultural site assessment, the factors included are agricultural land use, agricultural economic viability, land use regulations, alternative locations, compatibility with master plans and proposed use and infrastructure. This site analysis was implemented by Geographical Information System. (Williams 1985)

DeMers (2002b) conceptualize the land planning scenario with seven premises, which are simplified here:

1. Land has some level of demand.
2. Someone already owns the land.
3. The land has a particular size.
4. Current use tend to continue, since changing the land use function needs more energy.
5. Aesthetic value (whatever considered aesthetic, e.g. cropland)
6. Zoning regulations, legal restrictions, zoned for particular use.
7. Adjacency e.g. distance to other sites.

The list above contains many aspects that are added after the actual model is constructed in this thesis. These detailed or local level aspects are dealt with after the unsuitable areas are excluded by the model of this thesis. The model will be roughly screening the unsuitable areas off and the resources for finding optimal locations can be targeted to investigate probable suitable areas e.g. with DeMers' list above.

To handle site planning from the environmental aspect, we may counter many contradictory claims. Brook & Bradshaw (2015) examined nuclear role in global biodiversity conservation and said that renewable energy is typically more land-use intensive than other energy sources. They suggest that nuclear be used along renewable energy forms in the future to avoid dependency on coal, oil and gas. The study has been done by using multi-criteria evaluation for seven electricity-generation sources. Multi-criteria evaluation used in their work is a common way of studying suitability. This manner of studying is being utilized in this thesis though the term is not mentioned. Brook's & Bradshaw's (2015) criteria for the best energy sources are lowest environmental or economic impact, emissions, levelized cost of electricity, capacity of electricity delivery, fuel mining footprint, deaths from accidents and volume of radiotoxic waste stream. This criteria is weighted by knowledge of experts. (Brook & Bradshaw 2015) Nevertheless, their study focus is on comparison between the different energy sources the valued criteria are interesting from this thesis' aspect as well.

Solar energy stations impact on ecology has been assessed by Tsoutsos et al. (2005) in the paper Environmental impacts from the solar energy technologies. It was one of the first studies to assess the negative sides of solar energy production from the ecological perspective (Watson & Hudson 2015). However, Tsoutsos and others argued that environmental impacts could be mitigated by optimizing the locations on the roofs and the environments that are already human-used. In the year 2005 the utility-scale solar stations were not yet popular (see Figure 4). Nevertheless, in the conclusions of their paper they say that central solar systems assessment should be done. (Tsoutsos et al. 2005) The research of Chiabrando et al. (2009) concludes that both the viability of the PV products and the large area required are both problematic. Both of these have improved by the technological development e.g. in the efficiency of the PV cells.

A recent study (Watson & Hudson 2015) about solar and wind opportunities in Southern England were assessed by multi-criteria evaluation. The three stage approach using GIS tools was used. On the other hand, suitable areas were categorized hierarchically, and wind and solar competed on the same areas while solar often "won". Unsuitable areas are a result of a binary model which is explained in the theory section (see Section 2.1 GIS modeling concepts). Multi-criteria decision making (MCDM) and GIS are the two essential tools for creating maps for decision-makers. In their paper (Watson & Hudson 2015) the criteria for location optimization are agricultural land classification, historically important areas, landscape designations, residential areas, wildlife designations, aspect for solar radiation and slope. An analytical hierarchy process (AHP) is used to weigh the variables. Weighing is decided by the knowledge of experts on the field of renewable energy site orientation. (Watson & Hudson 2015) This same weighing method is used in van Asselen's and Verburg's (2012) paper; they globally and hierarchically categorizes land use by the land cover and intensiveness of land use. The importance of their work lies in the global model that tells the human use intensiveness i.e. land use pressure. This work is one of a kind since it is globally workable and includes the intensiveness of human use of land as well as the land cover. The protection of intensive food production areas is especially crucial as rising food prices cause people in the developed countries to suffer from lack of food and the need for food aid increases e.g. in

Great Britain (Lambie-Mumford & Dowler 2014). The situation is worse in developing countries and food production must be secured.

In the study of Santangeli et al. (2015) the global comparison between different renewable energy resources is made assessing how these energy production forms (biomass, solar and wind) conflict with biodiversity protection area suggestions made by Montesino Pouzols et al. (2014). The biodiversity protection dataset is used in the model of this thesis as well. It is presenting the situation in the year 2040. Santangeli and fellows found that globally there are regions that are interesting from the point of few of wind and solar energy development that do not have that rich biodiversity. An exception is Central America because of its rich biodiversity coverage. They concluded that solar energy is by far the best way of producing renewable energy when environmental costs are taken into account. Utilizing 1 % of the land area that is outside the most critical biodiversity areas could electrify the whole globe. This includes the assumption that energy is not locally produced. The solar potential was originated from Pogson's et al. (2013) work. The solar potential is based on latitude, time of the year, 10 % conversion efficiency and cloud coverage. Pogson's et al (2013) way of calculating the radiation differs from this thesis. They use Miscanfor model for estimating the global solar radiation and this thesis uses Area solar radiation tool from ArcGIS software and NASA's cloud cover model (NASA 2015).

Before now there has not been comprehensive global datasets that support investment decisions from the aspect of land use pressure. More specifically, the intensiveness of human land use and prioritized areas for protection of vertebrates are two crucial datasets that are used in the outcome model to direct the decision-making products environmentally and socio-economically sustainable.

1.6 Outline

This thesis working process follows the form of PPDAC-model (Problem-Plan-Data-Analysis-Conclusions), an iterative process model from the field of analytical methodology (De Smith et al. 2009). The text is structured this way as well.

In the *Introduction (1)* of the thesis the main point is to clarify the need for this study: Why it is important to create a model for investigating potential areas for utility-scale power plants and why land use pressure is taken into account. Also its goal is to give an understanding and background information of solar energy production on a broader scale. It addresses the motivation for solar technology and phases of implementation growth from the 1950s to the present. This section also clarifies what a solar station is. Since holistic optimization for solar stations would include closer to ten or more factors many restriction are made.

The theoretical (2) section includes theoretical concepts behind the thesis. Since model building is a one focus of the thesis Section 2.1 GIS modeling concepts and Section 3.1 PPDAC approach in analytical methodologies deal with model building. Concepts and models behind the spatial data analysis are clarified, in particular, map algebra, digital elevation models and measuring solar radiation are the focuses. Land change, land use pressure and land protection are specific datasets that are used in the thesis. These specific datasets are explained in the Chapter 3 Methodology).

The methodology (3) section includes the presentation of datasets and methods. What is the data that is used, how it is handled, and how to add the value to the raw datasets? The methods that are used to create and add value for the raw dataset are all demonstrated. Basically, the phases of the targeted model of this thesis are thoroughly handled with the support of maps, tables and diagrams. PPDAC, as an analytical methodology tool regarding the work process of the thesis has been used and it is opened in the beginning of the methodology.

Results (4) are clarified by short descriptions, analysis and maps. The output maps created from Harare, Denver and Helsinki dominate the result section. Furthermore, this section answers what kind of results the model produces as end products, the applicability of the model, and answers to the research question that are provided in this section.

The *Conclusions and discussion (5)* section offers a critique of the results and ideas in the thesis. The ideas for the further development of the model which arise during the progress of the thesis are described. This section clarifies if the new model brings something new to the field or not, and how the work settles into the field of location optimization modeling in the energy sector.

2 THEORY

2.1 GIS modeling concepts

GIS models can be divided into categories based on the purpose of the model. A prescriptive model answers questions like “What is the best location in which to site a factory?” (DeMers 2002d). A prescriptive model aims to give an answer and derive a solution for the best location. It does not need to be good or even satisfying, but the best available. The prescriptive model is comparable to a doctor’s prescriptions when treating diseases. Predictive GIS models can be considered under the model-family of prescriptive models. The predictive model aims to predict a phenomenon such as sunniness of the next month or forest fire proceeding (DeMers 2002d). DEM can be formed by different modeling techniques, and it may be an outcome of complex stochastic methodology as well. On the other hand, a DEM is a prescriptive model since it describes the empirical or estimated values of altitude within a certain resolution. Same models may have different GIS model names, depending on the purpose of the model and methodology of the modeling.

GIS models can be divided into categories based on the purpose of the model. A prescriptive model answers questions like “What is the best location in which to site a factory?” (DeMers 2002d). A prescriptive model aims to give an answer and derive a solution for the best location. It does not need to be good or even satisfying, but the best available. The prescriptive model is comparable to a doctor’s prescriptions when treating diseases. Predictive GIS models can be considered under the model-family of prescriptive models. The predictive model aims to predict a phenomenon such as sunniness of the next month or forest fire proceeding (DeMers 2002d). DEM can be formed by different modeling techniques, and it may be an outcome of complex stochastic methodology as well. On the other hand, a DEM is a prescriptive model since it describes the empirical or estimated values of altitude within a certain resolution. Same models may have different GIS model names, depending on the purpose of the model and methodology of the modeling.

Furthermore, prescriptive GIS models are divided into two categories by model comprehensiveness (DeMers 2002d). A model can be either heuristic or atomistic, depending on the area-connectivity of specified criteria. For instance, if a raster grid cell gets a certain value, but will not consider what’s happening on the neighboring cells, the model is atomistic. For instance, in site planning it is important to find large enough areas that have suitable characteristics. This is the reason why the site planning's final end products derived from GIS methods are often heuristic. When modeling floods or other physical dynamic phenomena using GIS, the model has to be heuristic, since in dynamic modeling neighboring cells matter. As a result of this thesis, a spatial information product (SIP) is formed as an outcome of GIS design (DeMers 2002b).

GIS models are also divided into two categories by its logic of methods (DeMers 2002d). The traditional forms of logic in scientific research are deductive and inductive logic. The inductive method tries to build up a general model by collecting or obtaining empirical data and assessing it. In this approach, statistical testing is fundamental (DeMers 2002d). The deductive method approach, in contrast, assesses the empirical data with restricted hypotheses that can succeed or fail; it is a less exploratory and experimental way of handling the data. For instance, data mining is often approached by inductive logic, as that is the exploratory way of finding interesting information. This thesis applies strongly deductive logic, since the criteria of optimal locations are set beforehand into the model. On the other hand, the datasets used in the model are often gathered by inductive logic and strongly tested to fit as a model of reality. Two examples of this are DEM and cloud coverage.

From the perspective of a customer or a decision maker, the end product of the assessment of the optimal location of a site should be a quickly understood graph or map that would clearly display its results at a glance. As a solution to this, a modeler can use division of "good location" and "bad location". This is called a binary map approach or a Boolean approach (DeMers 2002c). A division into good or bad is not always possible and will not serve the purpose. Also, the model created in this thesis could use categorical output maps, where suitability is ranked. However, the binary map approach is applied to ensure readability, and this acquires a total transparency of division of categories into non-suitable and suitable ones.

2.2 Spatial data analysis

This section introduces some fundamental concepts of geographical analysis. In the following sections these simple and complex spatial analyses are introduced: overlay, integrative, terrain, geostatistical and viewshed analysis. In addition, probability map and classical inference are explained due to their use in cloud coverage estimates. Behind of all the analyses is map algebra programming language, which has its own section. The concepts of solar radiation and its foundation, the digital elevation model (DEM), are also introduced.

Quantitative geography can be presented as a theory of three activities (Fotheringham et al. 2000): 1) the analysis of numerical spatial data, 2) the development of spatial theory, and 3) the construction and testing of mathematical models of spatial processes. This thesis concentrates on the first one. By one division, analysis can be either *simple* or *complex*. Simple analysis can be, for instance, an interpretation of a map as visual analysis, or a comparison between two or more maps. Geostatistical analysis and process modeling are referred as complex analysis. Between these two poles there exists a huge range of transformation and analysis types (Harvey 2008). In this thesis, both simple and complex analysis are used.

Overlay analysis is one of the simplest forms of geographical information analysis. With overlay analysis, it is possible to *combine* two or more overlapping raster or vector datasets and their attributes, although fine-tuning is often needed regarding coordinate systems and raster cells matching. The simplicity of the overlay analysis approach makes it attractive to use "as a method to incorporate environmental and social concerns along with engineering perspectives in planning of highways and other large construction projects", as Harvey (2008) says, putting into words the idea of the head promoter of the overlay analysis, Ian McHarg (1969). Overlay analysis also plays a big role in the methodology of this thesis. The idea is that the geographical information that restricts (in this case, land use pressure issues) or impacts the solar station localization may be different in other cases, but most likely it will still be combined with overlay analysis as long as a more advanced method remains to be found. Overlay analysis becomes *integrative analysis* after the results of overlay operations are interpreted (Harvey 2008). Overlay analysis becomes integrated in this thesis, since overlay combinations are placed by analytical explanation.

Besides local altitude factors, solar radiation is dependent on the altitude of the sun and other atmospheric phenomena. Solar radiation analysis is an application of *terrain analysis*, which uses geostatistical geographic information analysis operations. *Geostatistical analysis* is the most complex and important geographical information analysis operation, since its power lies in mathematics. Geostatistics can produce evidence of spatial processes that are not revealed otherwise (Harvey 2008). For instance, water flows may be extremely hard to predict without computing spatial data with complex dynamic 3D analysis. *Viewshed analysis* is a common type of 2D-terrain analysis used for solar radiation calculation. It is not as complex as predicting water flows or subsidence because of its static characteristics. A solar radiation tool will be closely examined in the methodology section.

Probability map is one kind of tool in geostatistics (Figure 7). Assessing the probability of cloud cover percentage from 30 years of empirical monthly data from NASA will be one geostatistical analysis in this thesis. A probability map can be generated by the rules of *Bayesian* or *classical probability* models, also called *Bayesian* or *classical inference* (Fotheringham et al. 2000). Cloud cover estimating by classical inference is used here. This model estimate gives the same weight for cloud cover in March 1980 as for March 2009, whereas a Bayesian model would emphasis the recent value, which is not interesting from the point of view of 20 to 30 years duration of solar power plants.

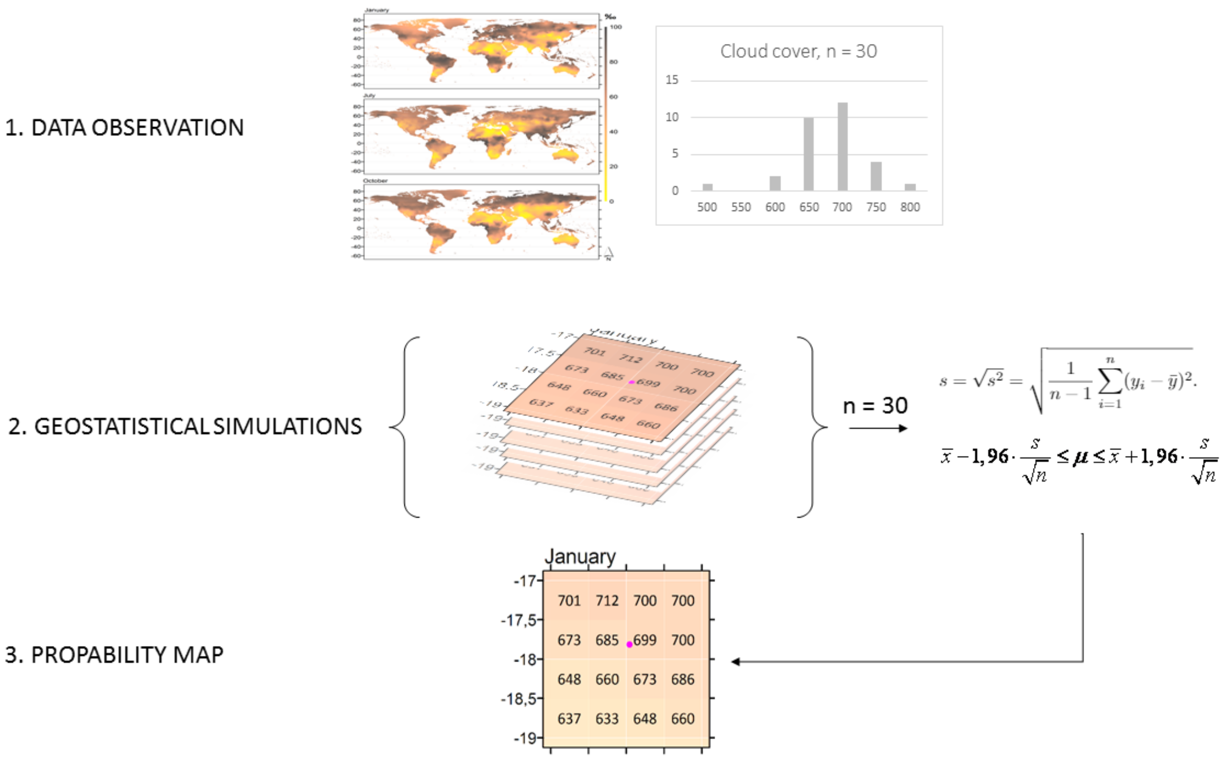


Figure 7. The process of probability map making. Monthly cloud probability is formed using this model.

2.2.1 Map algebra

Map algebra is a programming language that enables, say, a two-dimensional raster layer getting new values; i.e., manipulation of a map (Tomlin 1990). A raster dataset can be considered as any matrix algebra source (DeMers 2002a). Map algebra contains calculations between different raster layers or one layer. These have values for both continuous and discrete coverage. This coverage is a combination of cells, more popularly called pixels. For instance, when using map algebra to sum the values of two different raster layers, the cells in the raster layer must occupy same place. If they do not, raster datasets must be manipulated in a controlled way. If the cells will not match, we can give rules for the calculation, for instance, "use the largest pixel inside another cell". There are also other generalizations, such as a "mean value" that can be calculated from the non-matching cells by using all the cell values that intersect with the output cell. In this case, the values must be measurements, not classes. Map algebra needs an input layer or layers to have something where the calculations are made from. Algebra is made for one pixel at a time. Map algebra can consist of complex calculations like trigonometric functions, but the most general are arithmetic operators like '+', '-', '*' and '/'. An example of this is shown in Figure 8. The map, which has the worst resolution, will be the one that defines the resolution in the output. It means that other resolutions must be dropped off. Before implementing the map algebra it is good to have defined extent, cell size and a mask. Cell size is dependent on the input data cell size (Tomlin 1990).

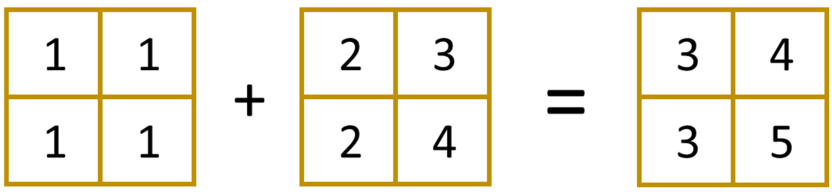


Figure 8. Example of arithmetic operator (+) for spatial data in raster calculations.

With map algebra, we generate a new layer with functions needed. For example, it might be the case that we want to reclassify values on the layer(s). Most raster tools in GIS software (e.g., ArcGIS, QGIS) will accept as an input integer numbers, which means that map values must be reclassified or just round with Map algebra. It can also be that we restrict some of the values to non-values (e.g., null).

Tomlin's (1990) paper presents local functions, but there are other kind of functions as well. Map algebra can be used to calculate different area constructions with different tools. For instance, local, focal, global, zonal, and incremental functions are used. Local is interested in the cell itself and its neighboring cells, focal is interested only in neighboring cells, and global is basically interested all the cells with some weighting. Zonal concentrates on a specific direction, and incremental is for more advanced analysis such as flooding analysis, where the values next to each other are dependent on the source of the water in this case. Solar analysis fits into this type of incremental function as well.

In Tomlin's (1990) book there are many application opportunities with using map algebra. It is possible to form contours, create optimization maps of land development potential, nature protection, mobility/accessibility, intensity of some measured interest (such as population) by observing neighboring points, land use, land cover, and wind exposure. Some of the local functions are called "LocalMinimum", "LocalMaximum" and "LocalAverage*",* etc. Some of these can be analyzed with vector data, but the speed of calculations should be on a higher level. There are also tools that work only for raster.

In the Master's thesis, map algebra is used in many phases; e.g., to calculate the cloud cover estimations and to filter the radiation calculations with cloud coverage estimations. Map algebra is a robust programming language that is included in GIS packages such as GRASS, ERDAS, and ArcGIS. In ArcGIS (greater than 10) Map Algebra is nowadays implemented by Python programming language.

2.2.2 Measuring solar radiation

The sun is practically viewed as a never-ending source of energy, as are wind and geothermal energy. The core of the sun produces radiation energy as electromagnetic radiation. Electromagnetic radiation has horizontal waves that travel by velocity of light. (Lehto et al. 2007) Types of electromagnetic radiations are radio wave, light, infrared radiation, ultraviolet radiation, and x-radiation. Photovoltaic solar cells absorb and conduct not only visible light, but also infrared irradiation, whereas solar thermal stations collect the heat. The higher the frequency of a wave, the more energy is in transit. Figure 9 demonstrates how much theoretically the top of the atmosphere receives energy from the sun on a daily average. The amount of radiation that reaches the ground is enormously smaller since the atmosphere and clouds. Impact of the atmosphere is included in the Area solar radiation tool used in this thesis. Impacts of clouds are estimated by probability mapping technique (see Section 3.3.2 Probability mapping for cloud coverage estimates).

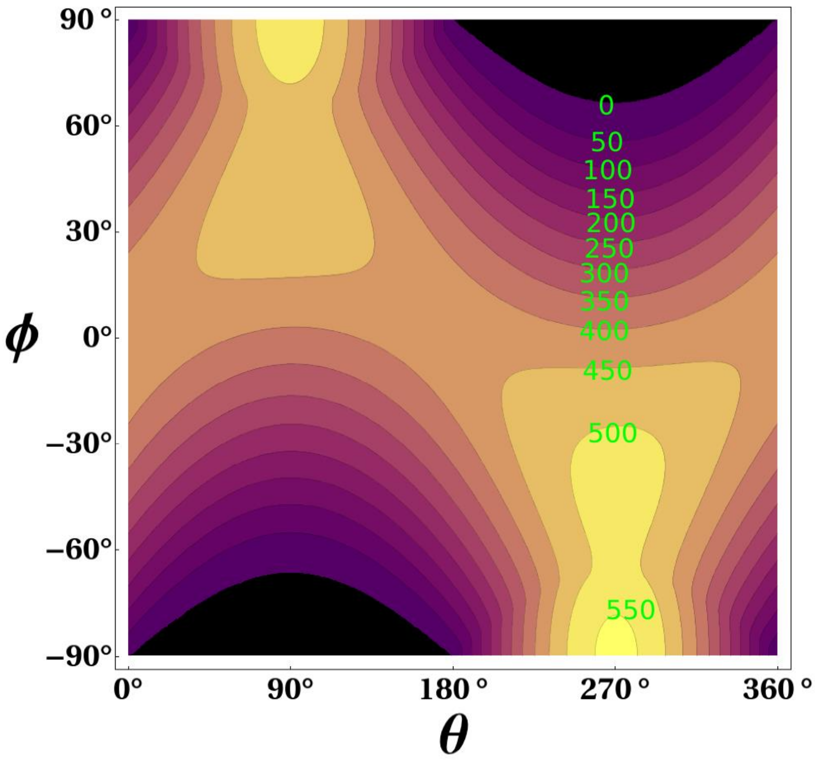


Figure 9. The map shows insolation amount (W/m²) top of the atmosphere. It is made by Incredio (https://commons.wikimedia.org/wiki/File%3AInsolationTopOfAtmosphere.png).

Measuring the insolation on top of the atmosphere is now left behind and focus is directed towards to insolation that reaches the ground surface (Figure 10). There are plenty of ways of estimating the *insolation*. *Solar radiation* can describe the energy amount moving without a time aspect, whereas insolation means the amount of energy that reaches the ground (certain area e.g. square meter) on a certain period of time (e.g. hour). This unit is handful since it is easy to convert it to electricity amount by multiplying the conversion efficiency of the technology (e.g. solar panels) and the amount of insolation. The result indicates the ability to produce electricity with certain technology on a certain area.

The solar radiation here is divided into three types. The first is *direct radiation* that comes directly through the atmosphere as a beam and touches the ground straight. Direct radiation is the most dominant one even though the ratio of radiation types changes by the latitudes and cloud coverage. The second dominant radiation type is *diffuse radiation* that is a radiation which has scattered by molecules and particles in the atmosphere, but still reaching the ground from any direction. The next paragraphs are dedicated in opening the different estimation techniques for insolation and they are considerable leaning on a topical Master's thesis of Holstein's (2015) that comprehensively deals papers with the questions of different measuring techniques of insolation.

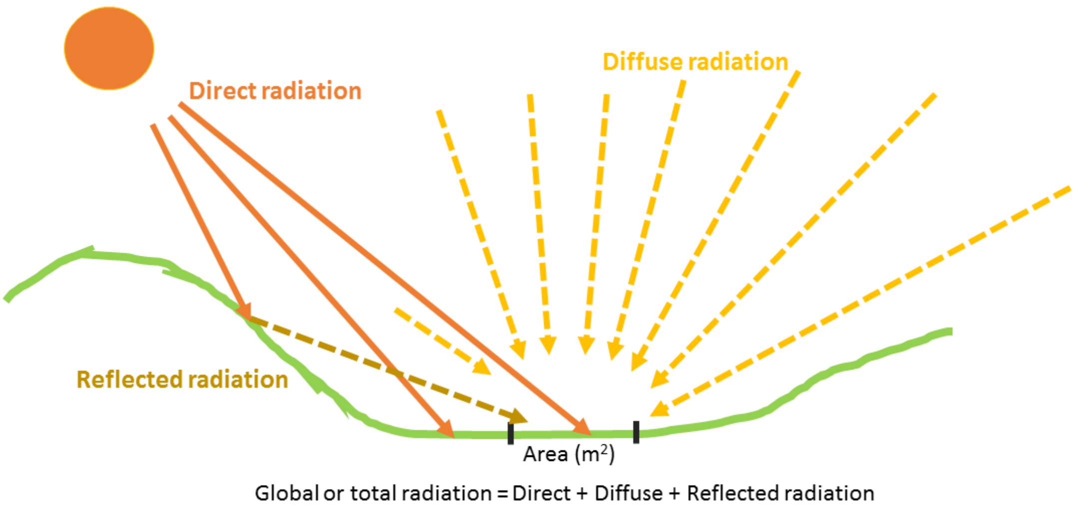


Figure 10. Three different solar radiation types.

Modeling solar radiation is a complex process where many different factors are examined. The factors included are latitude, temporal solar angle, elevation model, surface inclination, surface orientation, ability to reflect and many atmospheric phenomena such as particles condition, gas molecules and water condensation. (Šúri & Hofierka 2004)

Many solar radiation estimation models have been developed in order to measure radiation easier and cheaper than taking multiple sensors to the area to measure the actual radiation. Especially, global radiation estimates have been under development. The models belongs ordinarily into these categories: manual theoretical calculations, satellite data estimates, geostatistics or geographic information system (GIS) models. (Ramachandra & Shruthi 2007; Alsamamra et al. 2009)

Meteorological stations can provide data for manual theoretical calculations or estimates can be indirect as conducted from theoretical or empirical formulas (Ramachandra & Shruthi 2007). Based on the observation data of weather stations statistical model have been developed to identify differences between diffuse and direct radiation (McKenney et al. 2008). Also satellite data enables methods to estimate solar radiation. Even though, geostationary satellites provide effectively continuous data of Earth's radiance, the data is not fine enough of its resolution for photovoltaic applications. Satellite data also may fail in complex topography. (Alsamamra et al. 2009)

Interpolation methods, both deterministic and stochastic, are developed for estimating solar radiation. Splines, polynomial regression, weighted average and different stochastics kriging techniques are used for interpolating measurement points of solar radiation. (Alsamamra et al. 2009; McKenney et al. 2008) Nevertheless, the methods listed above are providing wide range of effective tool for estimating solar radiation, only GIS methods are effectively and reliable enough providing the information of solar radiation in rugged terrain areas (Tovar-Pescador 2006; Alsamamra et al. 2009; Ruiz-arias et al. 2009). Insolation models have been developed for many different GIS platforms (see Section 3.2.1 Data behind radiation estimates).

An area-based model calculates insolation using surface orientation and shadow effects from a digital elevation model. In practice, Area solar radiation tool from ArcGIS 10 software generates an upward-looking hemispherical viewshed, which is generated for each location of DEM by viewshed algorithm. The result of the viewshed analysis is a "fisheye" in Figure 11 in the right. To create viewshed, the viewshed algorithm needs DEM as its input. After the viewshed algorithm excludes the "sky areas" that cannot be seen from the observing ground point because of the terrain barriers, a sunmap and a skymap (Figure 12) can be constructed.

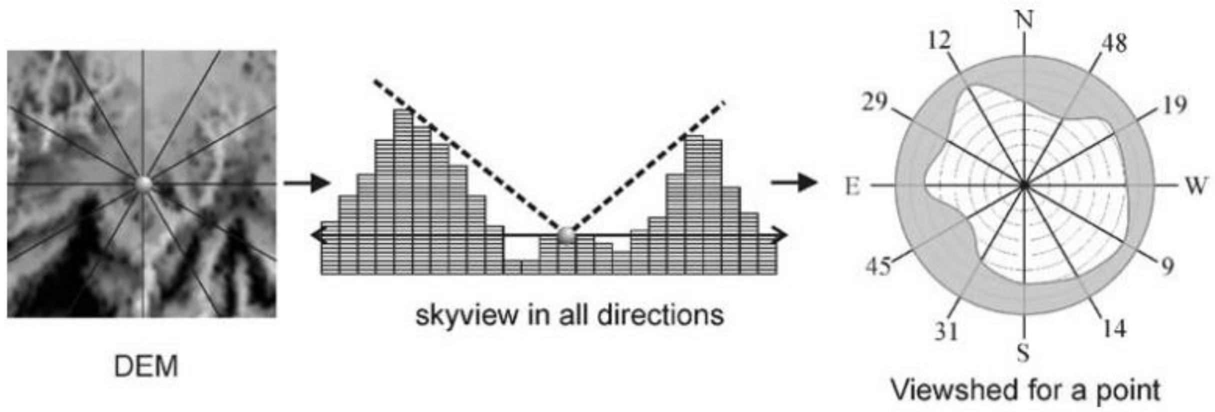


Figure 11. A process of calculating the solar radiation at a certain point or area with a viewshed method. The picture is based on Fu & Rich 1999 (Tovar-Pescador et al. 2006).

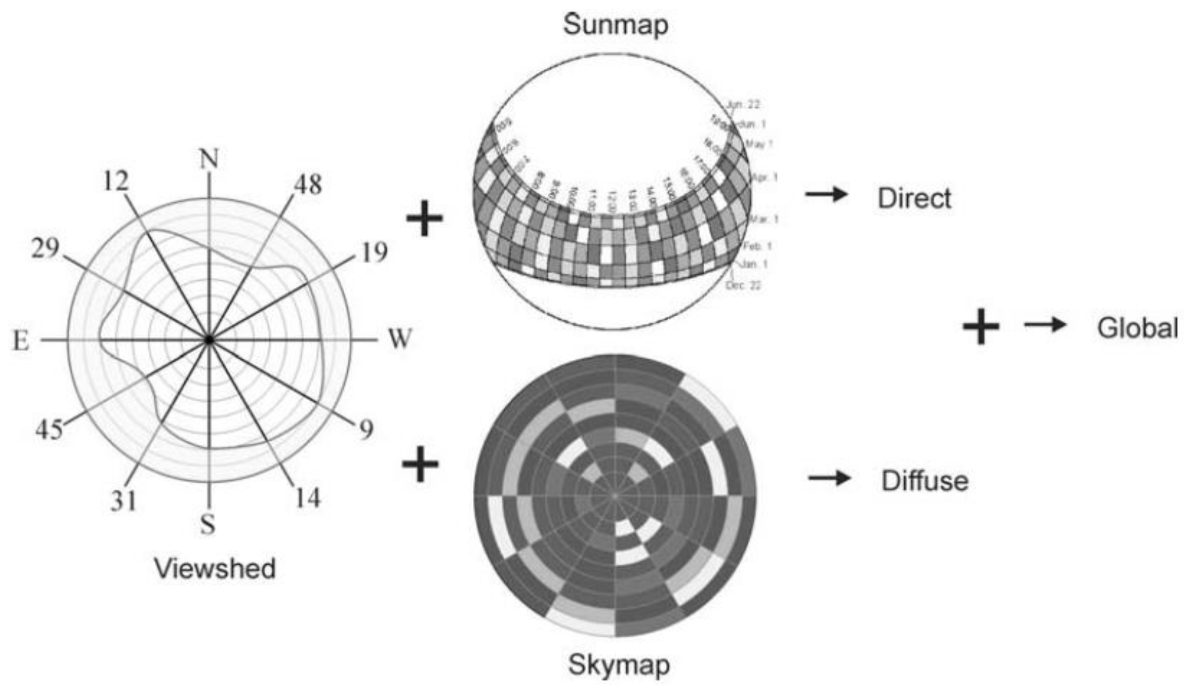


Figure 12. The sunmap (direct radiation) and the skymap (diffuse radiation) calculations can be regionally restricted by the viewshed algorithm result. The picture is based on Fu & Rich 1999 (Tovar-Pescador et al. 2006).

An option for cost-efficient spatial radiation models based on landscape modeling on the earth's surface is a collection station network. Stations are needed densely to produce accurate information of irradiation in rugged terrain areas, so it is very costly. Spatial solar radiation models, which are based on landscape modelled from DEM, are effective in estimation of solar radiation (Dubayah & Rich 1995, cited in Fu & Rich 1999). DEM-based models are available in GIS software, and it is easy to integrate other spatial datasets into the software.

A solar area radiation analyst tool in ArcGIS 10.2.1 software by Esri offers a way to estimate solar radiation or irradiation reaching the earth's surface. The tool needs altitude data, since topography is the most important factor on a local scale when calculating the distribution of solar radiation (Tovar-Pescador et al. 2006). USGS (NASA) offers a free digital elevation model (DEM) covering the whole globe with different resolutions. DEM is an elevation model that consists of continuing raster surface. Every cell has a flat platform, normally including one altitude value as an attribute.

The most important aspects of an algorithm are that it is time- and space-efficient, correct, general, simple and readable. A good algorithm should also provide an answer to a real problem and be implementable, but this is not necessary (Cormen et al. 2009). Solar radiation analysis is computing-intensive, since radiation is calculated by line-of-sight analysis that uses a viewshed algorithm (see Section 3.3.2 Area solar radiation tool). It is a global algorithm, meaning it handles all the cells in the study area in order to get a result to a one cell and this is enabled because of fast viewshed algorithms have been developed (Wang et al. 1996).

Viewshed algorithms are nowadays very effective due to their excluding tactic. Here are the steps of one type of viewshed algorithm: the algorithm first detects invisible DEM cells in the neighborhood using the local surface of the source cell. The second step is to assure that there are no areas on the "black area" that should be involved, so a kind of saturation of *line-of-sight* must be done. The third step uses a line simplification method for finding peaks of the slopes. The fourth step is to find by line-of-sight algorithm the intermediate invisible points. The last step deals with the special cases, meaning those cases that may or may not be visible (Wang et al. 1996).

2.3 Land use pressure

2.3.1 The Land system model

A Land use model for the year 2040 is a predictive GIS model. The Land system (later LS) is a sophisticated model compared to old land cover models in the context of land use pressure. Van Asselen and Verburg (2012) have designed a new way to assess land change on a global scale. Traditionally, land cover has been focused on when assessing land-change phenomena, as in the studies of DeFries and Rosenzweig (2010) and Hansen et al. (2010). Land cover includes biophysical spatial data and is classified as, e.g., agricultural, urban, or deforested land areas. Nevertheless, van Asselen and Verburg (2012) argue in their work that when classifying the actual use of land, and even further assessing land use pressure, it is crucial to take into account the intensiveness of the human use of land. The researchers' result is a model known as Land Systems (LS), which is observed next and with more detail in Section 3.2.2 Land use datasets. The aim of the paper is to classify, map, and characterize land systems on a global scale and analyze determinants for these systems (van Asselen & Verburg 2012).

Land systems include information regarding land cover and, compared to older models, use land use management intensity and livestock integration as novel inputs (Figure 13). The model is built using a hierarchical classification method. Spatial determinants are strong socio-economical and biophysical indicators of human-environmental interaction (van Asselen & Verburg 2012). Data is more closely observed in the data section.

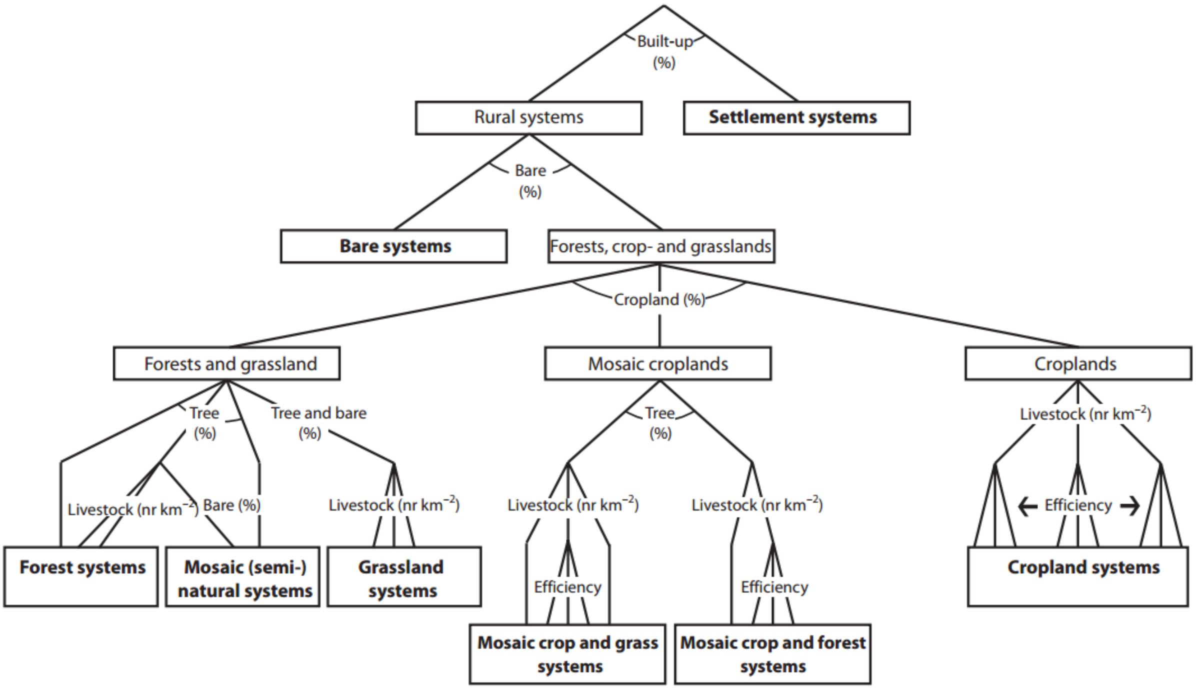


Figure 13. Van Asselen and Verburg (2012) have designed the land systems model with a hierarchical classification method. Land cover, livestock and agricultural intensity are individual determinants here.

The above-described land systems model is not the first representation of human-environment interactions on a global scale. For instance, Boserup (1965) took population density as a main factor in his theory regarding land use intensity. However, van Asselen and Verburg (2012) claim that the population classification criterion does not necessarily characterize land systems as directly as the three other aspects: agricultural intensity, land cover, and livestock density. For instance, intensively managed cropland systems do not necessarily correlate with high population; e.g., intensive cropland systems with livestock do not necessarily have high population density (van Asselen & Verburg 2012). Ellis and Ramankutty (2008) provided a novel classification of the earth's biomes as increasingly detailed anthropogenic biomes, or anthromes. These are globally significant ecological patterns by which interactions between humans and ecosystems are examined (Ellis & Ramankutty 2008).

2.3.2 Nature and biodiversity protection in land management

One of the restricting aspects in finding optimal locations for renewable energy production is biodiversity protection (Santangeli et al. 2015; Montesino Pouzols et al. 2014). The Convention on Biological Diversity (CBD) has adopted Aichi Biodiversity Target 11, which is to expand the network of protected areas to 17 % of the terrestrial world by 2020 (Montesino Pouzols et al. 2014). CBD is an agreement struck by the United Nations in 1992, which was ratified by 168 countries by the end of 2015 (CBD webpage). This verifies that the political will to protect nature exists, and due to that this theme is included in this thesis.

A recent and thorough study on effective prioritization of protection areas has been done by Montesino Pouzols et al. (2014). One of their main results is that nature protection should be coordinated at a global level instead of at smaller regional or national levels. Nowadays, nature protection is overshadowed by *parochialism,* which leads to protections that do not consider the global entirety of threatened species, but instead focus on national needs. These species include 24,757 terrestrial vertebrates assessed by International Union for the Conservation of Nature and are called 'red list of threatened species'. Researchers show that 1) globally coordinated protection of 17 % of the terrestrial surface of the planet would triple the protection of species; 2) if global land use in 2040 remains as predicted, 1000 threatened species would lose 50 % of their present effective ranges, and 3) there is a huge gap between national and global conservation. They created a classified global spatial dataset as a guide to prioritizing areas that should be conserved (Montesino Pouzols et al. 2014). In the data section this data is addressed in detail, and in the method section the dataset is integrated into the model. This dataset is one-of-a-kind, and it is hard to find comparable data.

3 METHODOLOGY

3.1 PPDAC approach in analytical methodologies

In the field of spatial analytics and geostatistics, the quantitative approach is a legitimate paradigm. The researcher is objectively limited to data collection and interpretation, gaining empirical measured data as the basis of the analysis, models and theories. This study follows a quantitative approach and uses tools from positivistic analytical methodology, as spatial analysis is often approached in the wider context of analytical methodologies. An analytical process is often complex and needs iterations to yield an outcome that is actually helpful (De Smith et al. 2009: 87), For instance, this thesis aims to produce a usable model as a part of a pragmatic planning infrastructure, whereas a purely theory-based deductive waterfall-model-shaped-model will not necessarily counter the outcome expectations in the real world. This is why an iterative process method PPDAC (Figure 14) is supporting this Master's thesis process as a methodological tool. The application of the model in this thesis in Section 5.3 Building a GIS model for solar power plant location optimization.

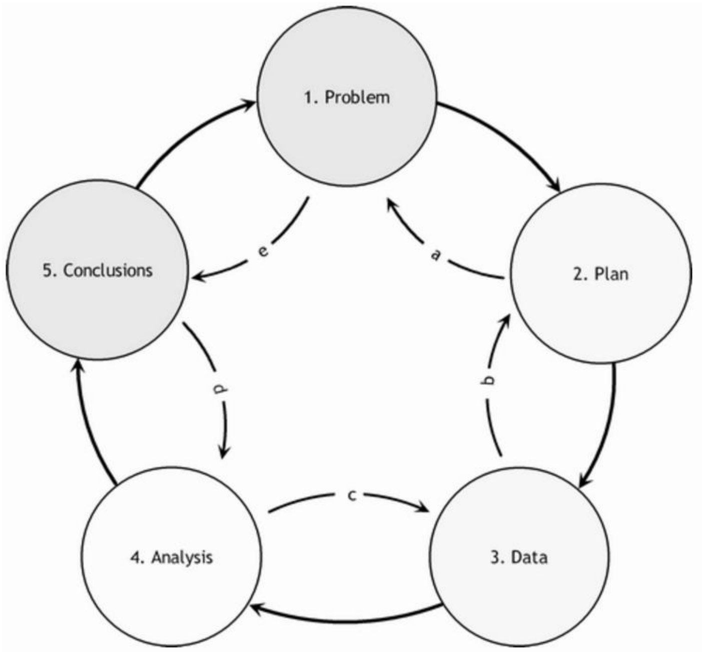


Figure 14. Iterative analytical process by MacKay and Oldford (2000) illustrated by de Smith, Goodchild and Longley (2009), page 86.

The formulation of the *problem (1)* is the first task in this study. The problem definition in the field of solar energy eventually needed at least two iterations to fit in an actual niche; i.e., to get interesting *conclusions (5)*. The real-world fact of increasing land use pressure was not included in the model at first, which made the formulation of questions unoriginal. Solar radiation magnitude has been already broadly investigated (see section: Measuring solar radiation), but land use pressure is often excluded in the research for finding optimal locations. The availability of land use data and guidance of thesis' supervisor lead to a reformulation of the problem. The task of fitting the research questions to the interesting conclusions continued until the last days of work on this thesis.

The PPDAC method suggests to *plan (2)* the thesis, going through all the next phases: data, analysis and conclusion. The sketch of an abstract works well as a planning tool, since all phases are outlined rapidly, but it will not bring out the truth of the iterative fixing through the work. The plan became more accurate during the process. It was strongly affected by my core competency, demand for this kind of research, availability of the datasets, time limitations, and my own interest in the solar energy business. The plan made me realize that it is impossible in a Master's thesis to create a detailed result map for the entire globe. This was the reason why the model evolved to examine one greater city area at a time.

Part of the *datasets (3)* used here are pre-existing spatial global datasets. Pre-existing dataset means that they are produced by third parties (De Smith et al. 2009). Here, half of the datasets are obtained by research groups accompanied by my supervisor. Since getting the needed datasets was not an issue going forward, there was time to explore the way the datasets were conducted. The way these were conducted guided me to a certain pool of methods.

Using spatial *analysis (4)* is an effective way to narrow down optimal areas, as power plants depend on many geographical datasets. Methods for the quantitative datasets were one of the easiest tasks to decide and motivate a Master's student in geoinformatics. However, one of the hardest and most important tasks was to visualize the results of the optimal location, since 12 different output maps were supposed to be seen at a glance due to the radiation distinctions between different months. Eventually, many iterations of method-picking were made so that the result output maps were easy to view.

The *conclusions* of the thesis are maps that show the prescriptive measures of solar radiation and exclusion of possible conflicted areas. The outcome aims to support the decision-making and minimize the prospects of failure for the projects-planning session.

3.2 Global data

The datasets used in this thesis are mostly collected or modelled in a deterministic manner. As an example, the digital elevation model, on the basis of radiation estimations, is conducted from two sensors (interferometric synthetic aperture radar) that were placed on the shuttle. By this two-sided elevation information, the elevation model was constructed. The cloud cover is estimated to a certain ratio (percentages, %) using NASA's satellite images. Also, the terrestrial nature conservation area was prioritized following the deterministic datasets of existing threatened vertebrate species. Nevertheless, the Land system dataset, which is an integration of three separate datasets (Table 2), is done by a hierarchical classification method in which two of the classifications of the original datasets are based on enlightened expert knowledge. This dataset is non-numeric, as distinct from all other datasets used here. See datasets in Table 1.

All the datasets to be used are in a raster format, and most of these compose a continuous coverage. The Land system dataset is discrete, with its values having categorical classes. The coverage itself is without gaps. Discrete values are not a problem, since the dataset is simply used as a mask to exclude land use forms that might cause a land use conflict with extensive solar energy production. Data cell sizes varies by datasets, which is taken into account to make sure that the data is compatible. The methods later presented are designed for raster datasets.

Table 1. The datasets used are in raster format (geographical coordinate system EPSG 4326). One degree (d) is 360 arc-seconds (a.s.) and example "E30m" means 30-meter resolution on the Equator.

DATA	RESOLUTION	TYPE	NUMERICAL TYPE	YEAR
DEM (STRM)	1 a.s., E30m	numerical	ratio	2014
Cloud cover	0,5 d., E50km	numerical	ratio	1980-2009
Radiation	1 a.s., E30m	numerical	ratio	2016 (own)
Radiation (after clouds)	1 a.s., E30m	numerical	ratio	2016 (own)
Nature conservation	60 a.s., E2km	numerical	ordinal	2014
Land system: livestock	180 a.s., E6km	numerical	ordinal	2007
Land system: agriculture	(continuous)	numerical	ordinal	2010
Land system: land cover (3)	500m & 300a.s.	categorical	-	2003-2009

3.2.1 Data behind radiation estimates

3.2.1.1 Digital elevation model

The division into two digital elevation models is next: digital surface model (DSM) and digital terrain model (DTM). DSM includes the highest point of elevation information at a certain point; for instance, the digital surface model includes the elevation information for canopies and houses or ground. Again, DTM includes only the ground elevation information. Both are contained in an upper class of the digital elevation model (DEM). In this thesis, urban areas and important forest areas are excluded from suitable areas. Also, the solar stations that are under study here are designed to the ground. This is the reason why DTM is selected here instead of DSM.

There are many accurate digital elevation models (DEM) to use to create estimations of solar radiation for utility-scale power stations. If the purpose of estimation was solar panels on a roof or a constructed urban area, a 30-meter resolution for DEM would definitely be too inaccurate. However, utility power stations need broad areas, and the purpose of this thesis is to restrict urban or other significantly competing areas as “unsuitable” areas. In the following, a justification for adequate resolution and selection of the final DEM data over rival DEMs is offered using relevant papers.

The resulting model for this thesis has to be usable also in a complex terrain area. Therefore, papers that are used here to justify resolution have rugged topography study areas. Picking a certain DEM with accurate enough resolution from Table 3 is based on a study by Rexer & Hirt (2014) and GIS domain websites.

Table 2. Specs of found different available digital elevation models (DEMs). Arc-second (a.s.), “E1km” is 1 kilometer resolution on the Equator. The geographical coordinate system is EPSG 4326 in all.

NAME	DEM TYPE	RESOLUTION	EXTENSION	PROVIDER
GTOPO30	DTM	30a.s., E1km	whole globe	NASA
STRM CGIAR	DTM	3a.s., E90m	60°N, 56°S	NASA & NGA
ASTER GDEM 2	DSM	1a.s., E30m	83°N, 83°S	NASA & METI
SRTM1 (chosen)	**DTM**	**1a.s., E30m**	**60°N, 56°S**	**NASA & NGA**
ALOS 3D	DSM	1a.s., E30m	partly globe	NTT DATA, RESTEC
ALOS 3D (not free)	DSM	5m	whole globe	NTT DATA, RESTEC
GMTED1	DTM	7,5a.s., E250m	whole globe	NASA

Tovar-Pescador et al. (2006) have concluded that a 20-meter resolution in DEM generates a correlation of 0.75 (i.e., multiple correlation squared), when monthly solar radiation estimates are compared to the radiation sensor data on the ground. In their study area, sized 10 kilometers times 5 kilometers in Sierra Nevada Natural Park, the altitude varies by 1 kilometer. Empirical measures from radiometric stations are always taken on a cloudless day from 14 different radiometric stations that detect global radiation. Differences, i.e., residuals, between detected and estimated values are shown in the left in Figure 15. Radiation estimates were calculated by Solar Analyst software (Esri), using global radiation measure as an assessed target. On the right side graph (Graph 1), a similar test has been conducted. A similar amount of testing spots is used, but the average value for each month is calculated, which due generalization, i.e., manipulation of data, makes the correlation statistic higher, 0.91. Resolution of the DEM is 30 meters.

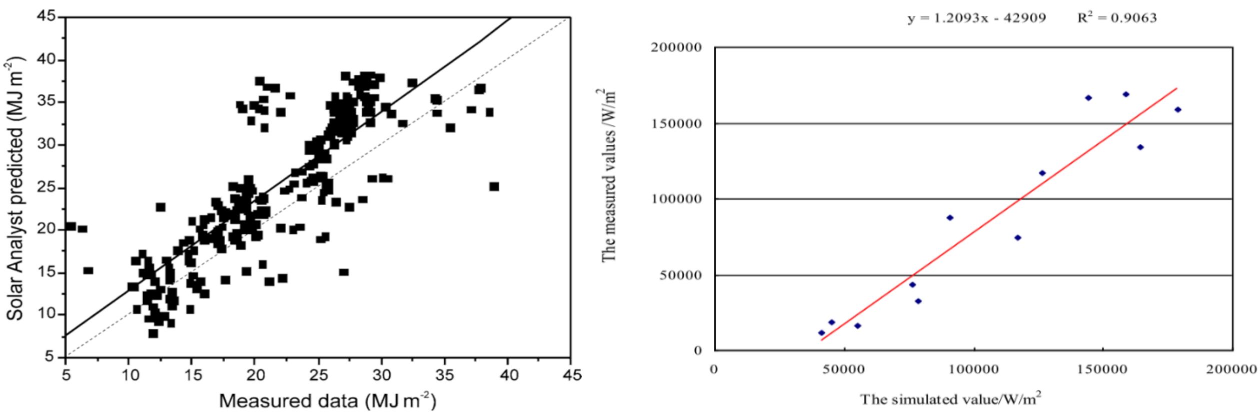


Figure 15. The left graph indicates correlation between measurement and estimate in Sierra Nevada Natural Park (Tovar-Pescador et al. 2006) *using 20 meter DEM. At the right side graph* (Mei et al. 2015)*, resolution is 30 meters and the place is mountainous area in China. ESRI's similar tool is used, but a 10-year different version.*

In Table 3, the accuracy of estimates is shown. Monthly radiation estimates are calculated using global radiation analysis tool (Solar Analyst) and 30-meter resolution in DEM (Mei et al. 2015). The values are correlations that show the compatibility of estimations and empirical measurements in different conditions and from different points on the compass. Both Mei et al. (2015) and Tovar-Pescador et al. (2006) lack the name or the origin of the data in the papers, which would help the latter's comparison of 30-meter-resolution DEMs.

Table 3. DEM with 30-meter resolution and Solar Analyst tool creates highly correlated estimations of global radiation (Mei et al. 2015)*. Comparison is made to radiation sensors.*

Aspect (°)	Plane				Plateau				Hill				Mountain			
	month															
	1	4	7	10	1	4	7	10	1	4	7	10	1	4	7	10
North (0° ~ 22.5°)	0.91	0.97	0.99	0.94	0.80	0.93	0.96	0.86	0.80	0.94	0.97	0.86	0.80	0.94	0.97	0.86
Northeast (22.5° ~ 67.5°)	0.94	0.98	0.99	0.96	0.80	0.91	0.94	0.85	0.79	0.91	0.94	0.84	0.81	0.93	0.96	0.86
East (67.5° ~ 112.5°)	1.02	1.01	1.01	1.01	0.98	0.97	0.97	0.97	1.00	0.98	0.98	0.99	0.99	0.99	0.99	0.98
Southeast (112.5° ~ 157.5°)	1.09	1.03	1.02	1.06	1.16	1.03	1.00	1.10	1.25	1.06	1.02	1.16	1.31	1.08	1.04	1.20
South (157.5° ~ 202.5°)	1.12	1.04	1.02	1.08	1.21	1.06	1.03	1.15	1.35	1.10	1.04	1.23	1.38	1.11	1.05	1.25
Southwest (202.5° ~ 247.5°)	1.09	1.03	1.02	1.06	1.20	1.05	1.02	1.13	1.24	1.06	1.02	1.16	1.32	1.09	1.04	1.21
West (247.5° ~ 292.5°)	1.01	0.99	0.99	1.00	0.96	0.97	0.97	0.96	0.99	0.98	0.98	0.98	0.98	0.98	0.99	0.98
Northwest (292.5° ~ 337.5°)	0.89	0.96	0.97	0.92	0.79	0.91	0.95	0.84	0.80	0.91	0.94	0.84	0.81	0.93	0.96	0.85
North (337.5° ~ 360°)	0.91	0.97	0.99	0.94	0.81	0.93	0.97	0.86	0.71	0.86	0.91	0.75	0.72	0.89	0.94	0.78

More justification to find a suitable resolution for DEM must be found, since solar radiation calculations are time-consuming and a lower resolution would save time in calculations. The perfect paper by Ruiz-arias et al. (2009) was found: "A comparative analysis of DEM-based models to estimate the solar radiation in mountainous terrain". They compare resolutions of 20 and 100 meters in DEMs using four different software tools: r.sun in GRASS, SRAD that can be run in ArcGIS, Solei-32 for Windows, and Solar Analyst by Esri. The study area is the same National Park of Sierra Nevada as in Tovar-Pescador's et al. (2006) research. The whole experiment consists of 523 records from 14 radiometric stations. These records were taken on both clear and cloudy days and are shown separately in the results. DEM, day of the year, latitude and time steps were similar and atmospheric conditions differ, whereas diffuse proportions may differ significantly (Ruiz-arias et al. 2009).

The result was that Solar Analyst estimates of global radiation were slightly better with 20-meter resolution DEM, whereas r.sun and SRAD and Solei-32 made it better with 100-meter resolution DEM. Solei-32 and r.sun produced the most accurate estimations, and Solar Analyst produced good estimates when parameters were for clear conditions and underestimates for cloudy conditions (Figure 16). The accuracy of the tools was especially tested by using root-mean-square error (RMSE)

(Ruiz-arias et al. 2009). In this thesis, the conditions are calculated for a clear sky with Solar Analyst and manipulated later by a cloud cover dataset for which Solar Analyst extension tools remains a valid choice. Since the 20-meter resolution produced better results for Solar Analyst, the 30-meter resolution is used instead of the 100-meter one. Also, the developers of Solar Analyst used a 30-meter resolution DEM while testing the tool (Fu & Rich 1999). Sensitivity testing to the resolution of the DEM showed that the datasets produced by 20- and 100-meter resolution DEMs in Solar Analyst concluded with a root-mean-square deviation of 5%, whereas others were near 9%. This indicates that the Solar Analyst datasets were most similar when the resolution was changed. It is most likely due to the heavy weight of transmissivity of the cloudy days in the model (Ruiz-arias et al. 2009).

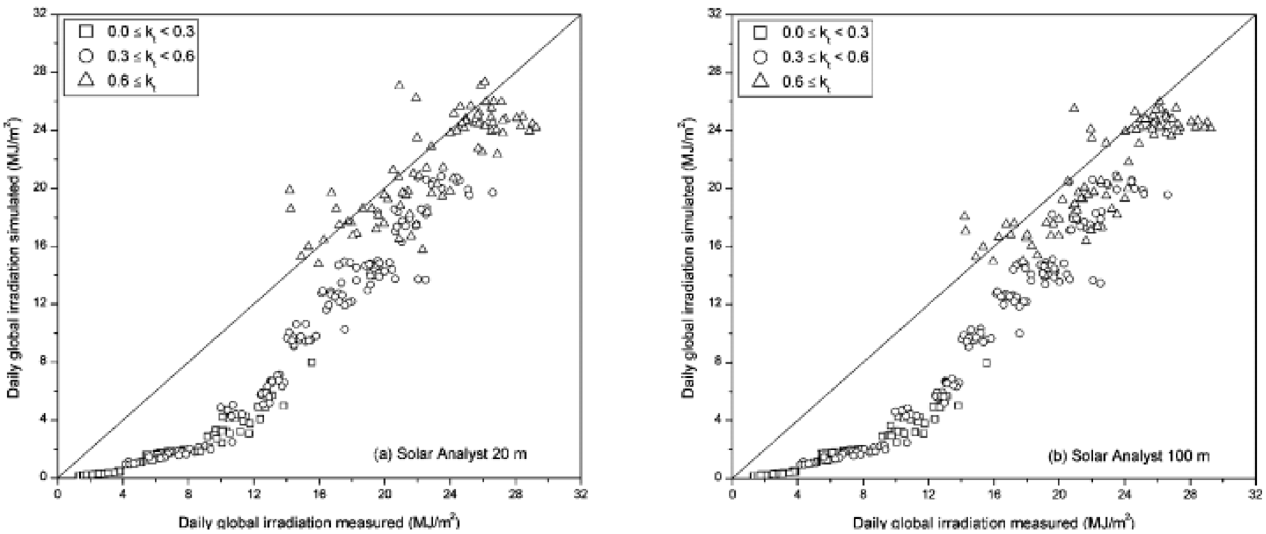


Figure 16. Solar Analyst's estimates are tested with 20-meter (left) and 100-meter (right) resolution DEM's. Triangles indicates clear sky, squares cloudy sky and balls from the between.

Table 2 listed three products of the 30-meter resolution DEM: ASTER GDEM2, STRM1 and ALOS 3D. Since ALOS 3D is still covering the world as a patchwork, it is excluded immediately. JAXA (ALOS 3D's producer) is also creating commercial DEM with 5-meter resolution, and it is not certain when the global 30-meter resolution DEM is going to be available. Up to that, ALOS is DSM, not DTM. In one study (Rexer & Hirt 2014), ASTER GDEM 2 is still not comparable to the SRTM DEMs CGIAR, and furthermore ASTER is DSM, not DTM. The conclusion of the comparison is that SRTM1 is the best digital elevation model in the model. The only negative side to SRTM1 is that it covers 60th parallel North to 56th parallel South, which excludes the northern countries of Canada and Russia from 60th parallel North up to the northern pole. Antarctica is excluded in the South. This lack of DEM in these areas is covered by GMTED 7,5 arc-second resolution DEM, which is a better version of GTOPO30. The resolution is around 100 meters in wide (latitude) and 250 in height (longitude); as the resolution gets better, the further it is from the equator. If national DEMs are found and the resolution is between 10 to 30 meters, they are accepted in the model. In this thesis, Helsinki in Finland is one validation city, and the Finnish Land Survey's 10-meter resolution DEM is used. Since the resolution is high, the study area is narrowed to 40km * 100 km area to reduce the calculation time due to the coastal location of Helsinki (see Figure 16).

SRTM 1 is downloaded via USGS' Earth Explorer, a web map application (webpage: http://earthexplorer.usgs.gov/). The raster grid size for DEM is 1 degree. For instance, Harare's DEM is downloaded in 4 raster tiles. Denver, on the other hand, has 9 tiles due to its proximity to the Equator. The tile areas are getting smaller the further from the Equator, since latitudinal degrees stay the same but distances shorten. Tiles or raster layers are moved to a Mosaic dataset in a File Geodatabase. The Mosaic dataset can combine different raster datasets into a single file, and this is enabled by functions of File Geodatabase. This single raster file as an output can be inserted into the Area Solar Radiation tool.

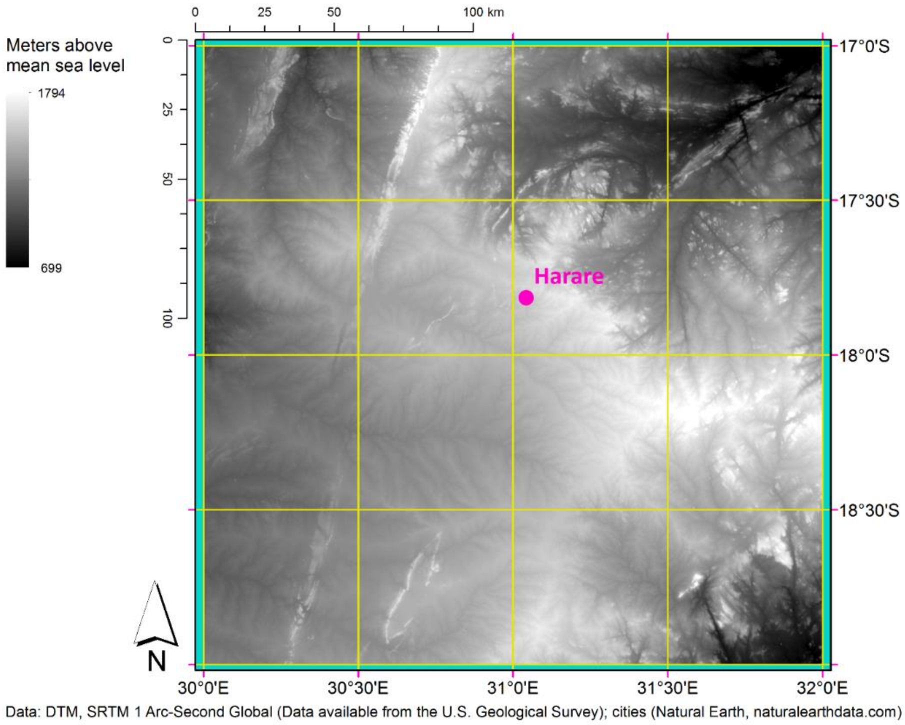


*Figure 17. Digital elevation model (SRTM1, 1 arc-second) of Harare. The size of the study area is 211 km * 222 km. ADD SCALE VERTICAL AND HORIZONTAL*

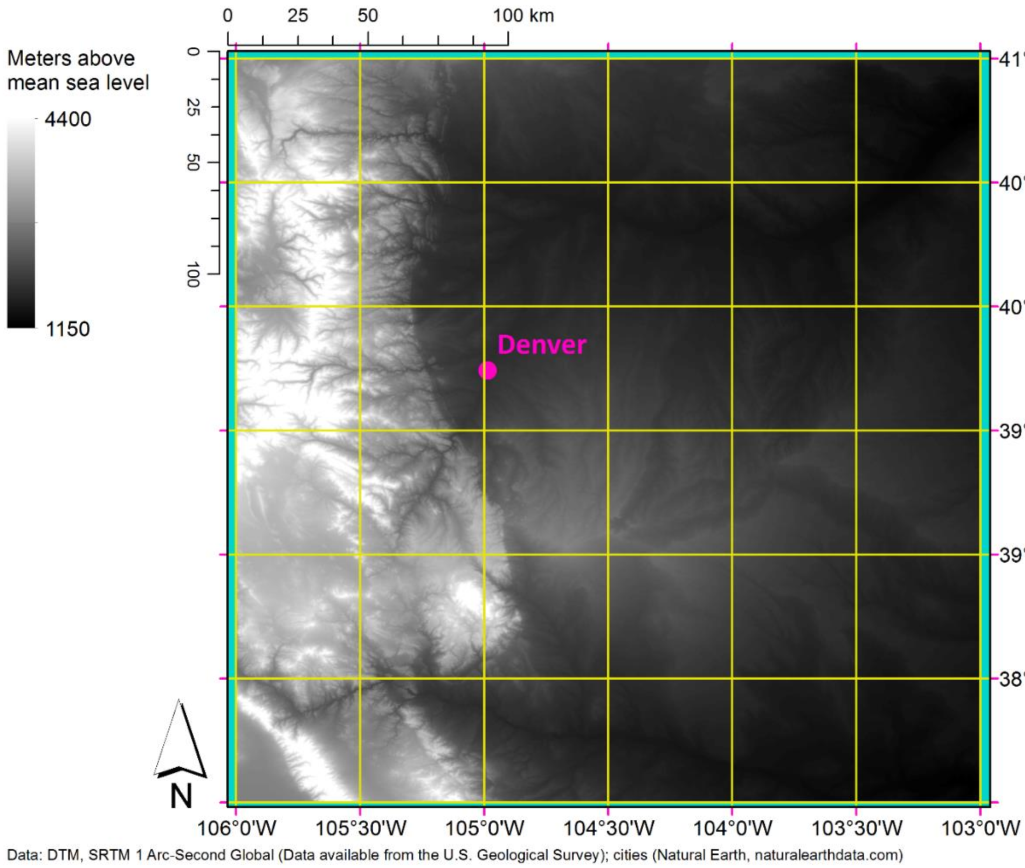


*Figure 18. Digital elevation model (STRM1, 1 arc-second) of Denver. The size of the study area is 260 km * 334 km.*

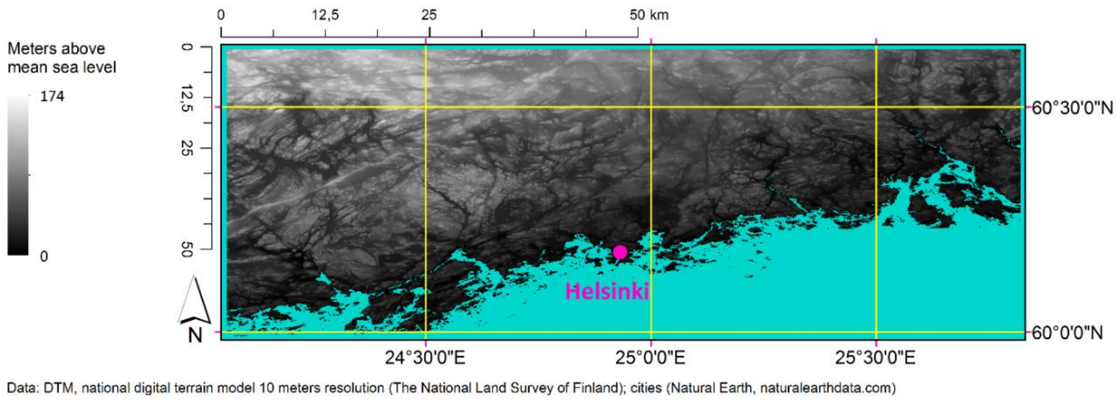


*Figure 19. Digital elevation model (national, 10 meter) of Helsinki. The size of the study area is 100 km * 40 km. The national coordinate system is converted to WGS84 (ESPG:4326) that Area solar radiation reads. The maps after this are converted back to the national coordinate system.*

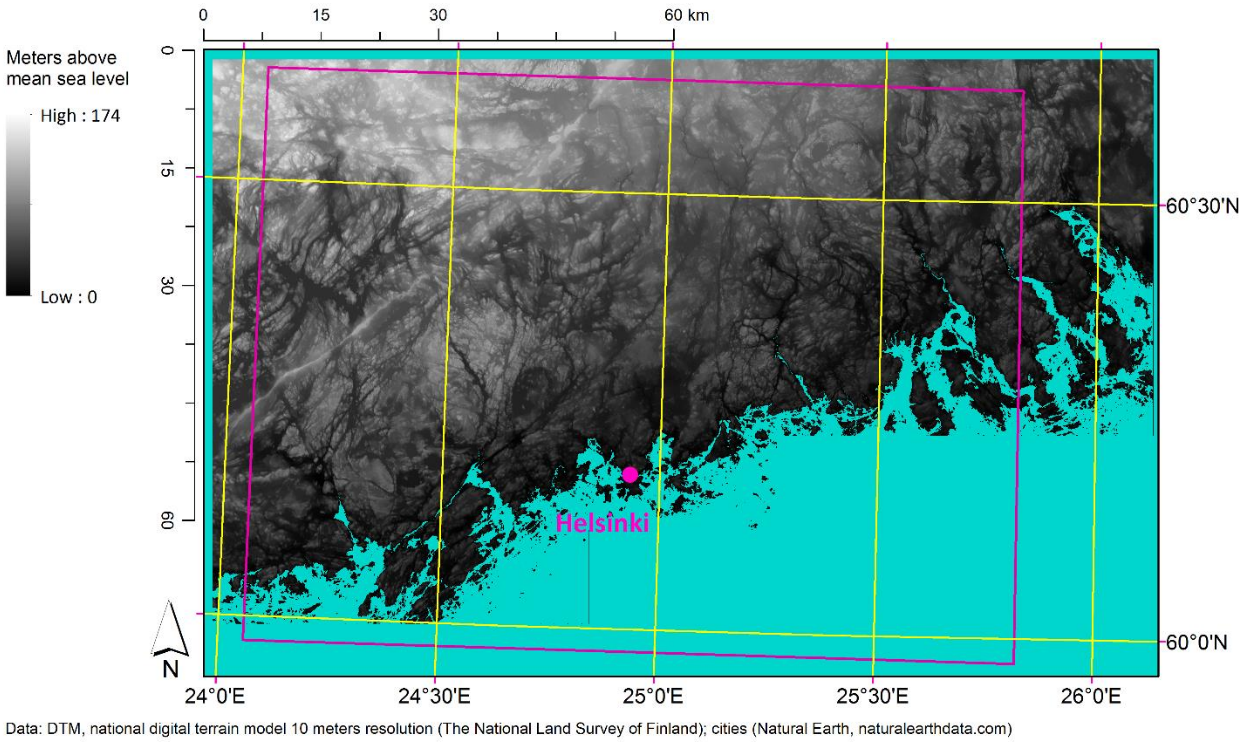


*Figure 20. Digital elevation model (national data, 10 meter resolution) of Helsinki. The size of the study area is 100 km * 40 km. The coordinate system is national Finnish geographical coordinate system (EPSG:3067).*

3.2.1.2 Cloud cover as filter

Solar radiation estimates are not going to fit reality without a thorough assessment of the impact of cloud cover. Many studies uses cloud cover data from CRU-TS Climate database when the historical monthly average is accurate enough (e.g. Pogson et al. 2013; Hasting et al. 2009). In this study, the monthly average is accurate enough, since radiation estimates are done per month as well. Figure 21 shows how the cloud coverage varies between different months. The particular cloud cover used in this study is CRU-TS version 3.10.01, distributed by CGIAR CSI. Cloud cover is valued as percentages, and its resolution is half a degree (NCAR & UCAR 2016).

30 years of monthly cloud cover data depict a representative sample of monthly cloud cover, although the cloud cover sample period and its yearly weights due to climate change should be considered. Shahzad, Renguang and Iftikhar (2014) use 30-year period monthly data in radiation calculations over Pakistan. Their data is provided by the Pakistan Meteorological Department's cloud cover data from 21 sensor points, and was verified by MODIS satellite Level 2 cloud data.

In Area solar radiation tool, it is possible to give atmospheric conditions as parameters, but the creators of extension tools say that information about transmittivity and diffuse proportions are not typically available (Fu & Rich 1999). Nevertheless, the situation must have changed over those 15 years, so this study will not include cloud cover in radiation calculations by Area solar radiation tool, but cloud coverage is addressed separately. The cloud cover effect on radiation is calculated by the equation that NASA's school uses (NASA 2015).

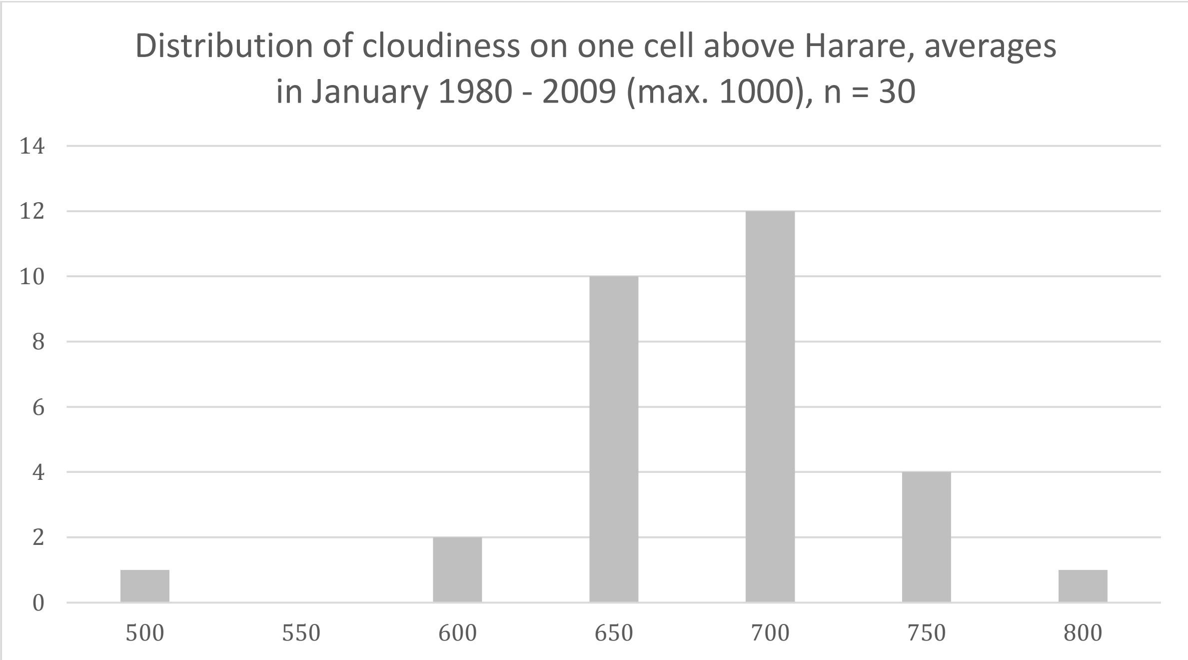


Figure 21. The distribution of monthly average on each year starts to remind a normal distribution while sample size is 30.

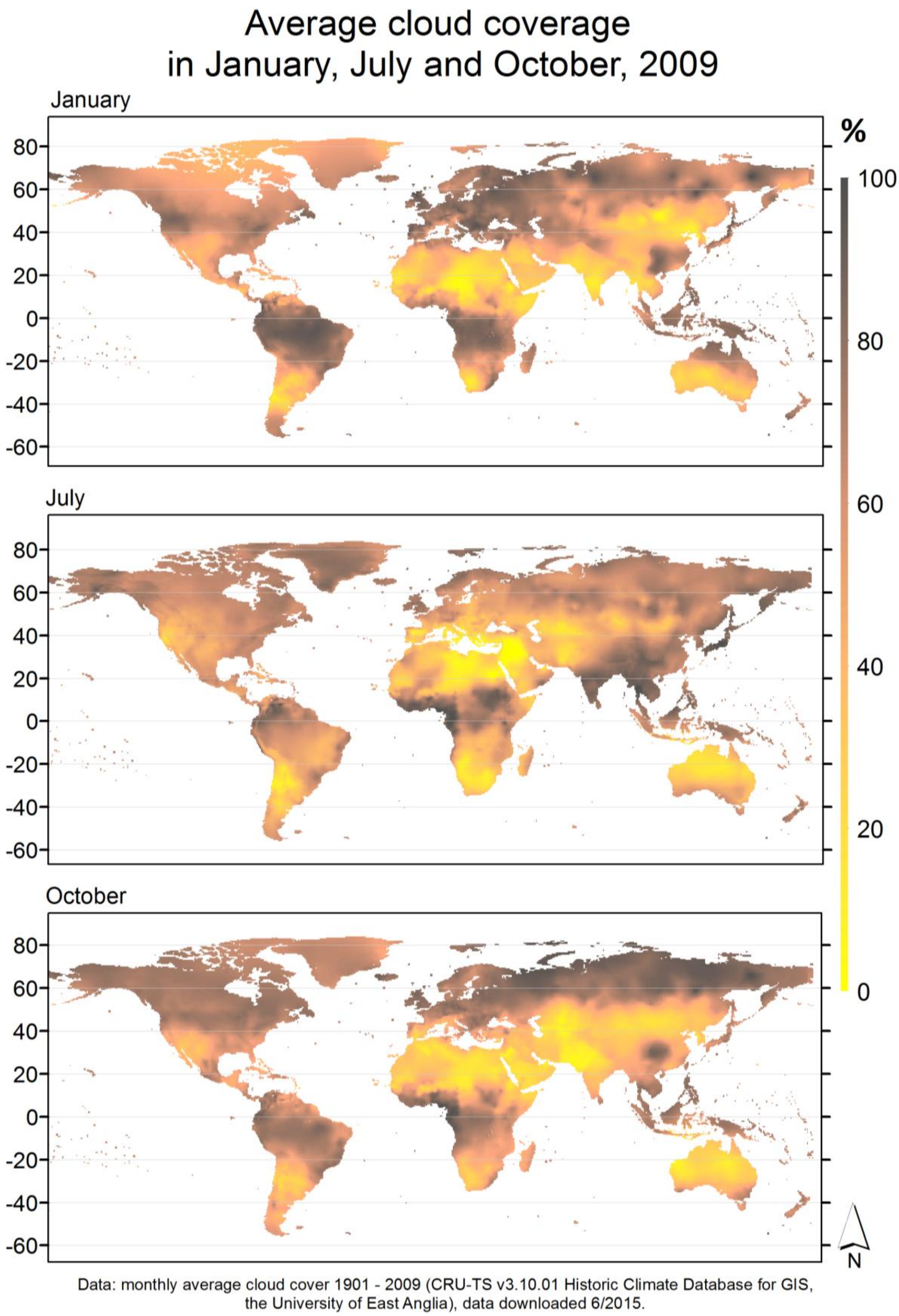


Figure 22. The cloud coverage varies geographically between different seasons. These cuts are taken from the year 2009 data.

3.2.2 Land use datasets

3.2.2.1 Land systems

Land system (LS) data is an integrated assessment model (IAM) that is often "used to assess environmental consequences of interactions between different systems, such as economic, social, and biophysical systems" (van Asselen & Verburg 2012). It is implemented in the thesis model as a land-use pressure factor. In their model, van Asselen and Verburg combine the determinants of land cover, livestock system and agricultural intensity (Figure 23). The land cover variables are percentages of tree cover, bare soil cover, cropland cover, and built-up areas. MODIS Vegetation Continuous Field dataset was used for the first two. MODIS 500m satellite data was used for discovering the built-up areas, and the cropland cover is work of Ramankutty et al. (2008).

Main group	Classification factor	Resolution	Unit	Reference
Land cover	Tree & bare cover	500 m	%	Hansen *et al.* (2003)
	Cropland cover	5 arcminute	%	Ramankutty *et al.* (2008)
	Built-up area	500 m	%	Schneider *et al.* (2009)
Livestock	Livestock density	3 arcminute	nr km^{-2}	FAO (2007)
Agricultural intensity	Efficiency of agricultural production	Interpolated from point data	ratio (0–1)	Neumann *et al.* (2010)

Figure 23. Factors behind the Land system dataset. (van Asselen & Verburg 2012, page 3127)

In Figure 24 below, the proportions of different classifications of land use intensiveness are shown. Black dots are population proportions of the whole Earth's population. The histograms are proportion of land system types of the area of the whole Earth. To learn more about model testing with logistic regression behind the hierarchical classification, see van Asselen & Verburg (2012). Global presentation of the Land system dataset is seen in Figure 25 in a form of a map.

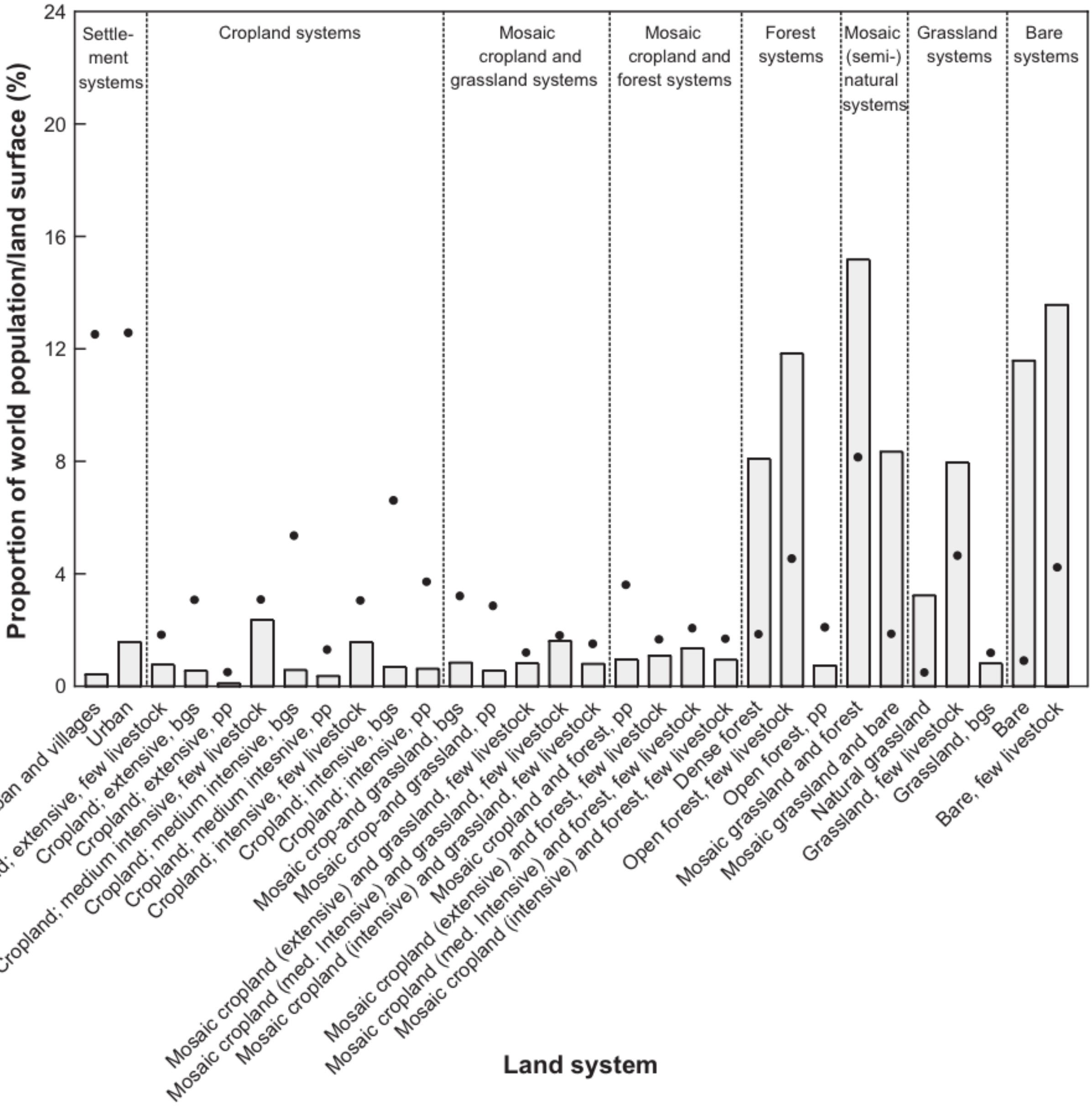


Figure 24. The classification of the land system dataset by van Asselen & Verburg (2013), *p. 3136*

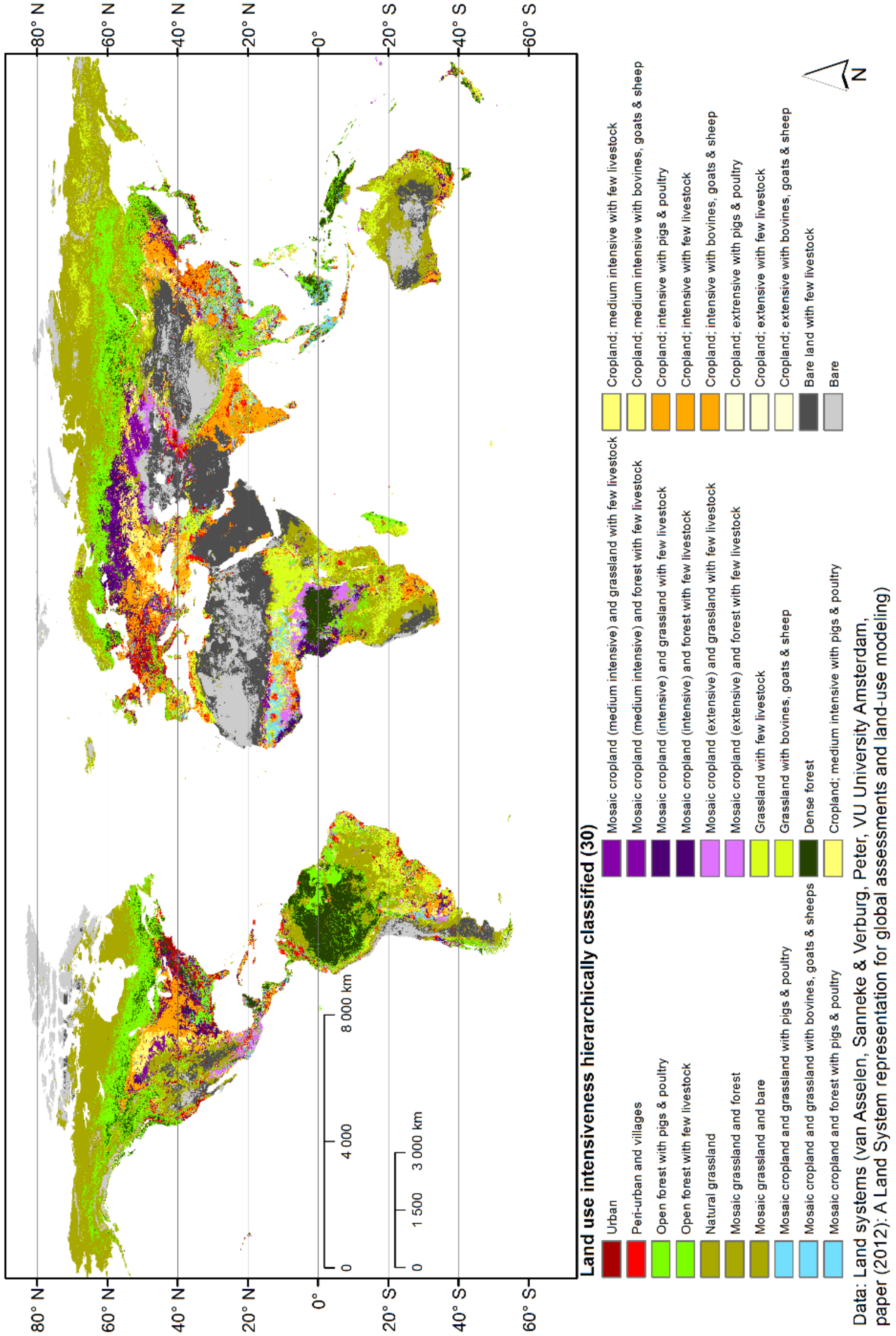


Figure 25. Land systems dataset by van Asselen & Verburg (2013) *visualized.*

3.2.2.2 Nature protection prioritization

The nature protection dataset used in this thesis was created by Montesino Pouzols et al. (2014) in order to save as many threatened vertebrates as possible by optimizing protection areas with GIS methods on a global scale. 168 countries in the world have ratified an UN agreement that by the year 2020, 17 % of the terrestrial land area of the world will be protected to save threatened vertebrates. Their approach in prioritizing the areas is that instead of national consideration of the best conservation areas, protection should be coordinated at a global level in order to save species effectively (Montesino Pouzols et al. 2014). The research was published in Nature, and the dataset they produced is a remarkable tool that is widely implemented in decision-making.

The dataset just referred to contains many input datasets. Union for the Conservation of Nature (IUCN) have created a distribution data for 24,757 terrestrial vertebrates and 827 terrestrial eco-regions that are the base data for the dataset. Montesino Pouzols and others used a newly developed prioritization method that creates a spatial conservation prioritization dataset while assessing the information of projected land use change. They show that the land use change that is estimated to happen by the year 2040 prevent carrying out possible protection aims. (Montesino Pouzols et al. 2014) More information about the projected land use change can be read in the paper of Laurence et al. (2012). The starting point for Montesino Pouzols et al.'s (2014) research was expanding existing nature protection networks, not to create new networks from scratch. The suggested globally optimized 17 % of the terrestrial nature protection areas are easy to implement in this thesis' model. The raster dataset they produced is harmonized to many resolutions and the resolution implemented into this model is 1 arc per minute. (Montesino Pouzols et al. 2014) Figure 26 presents the dataset.

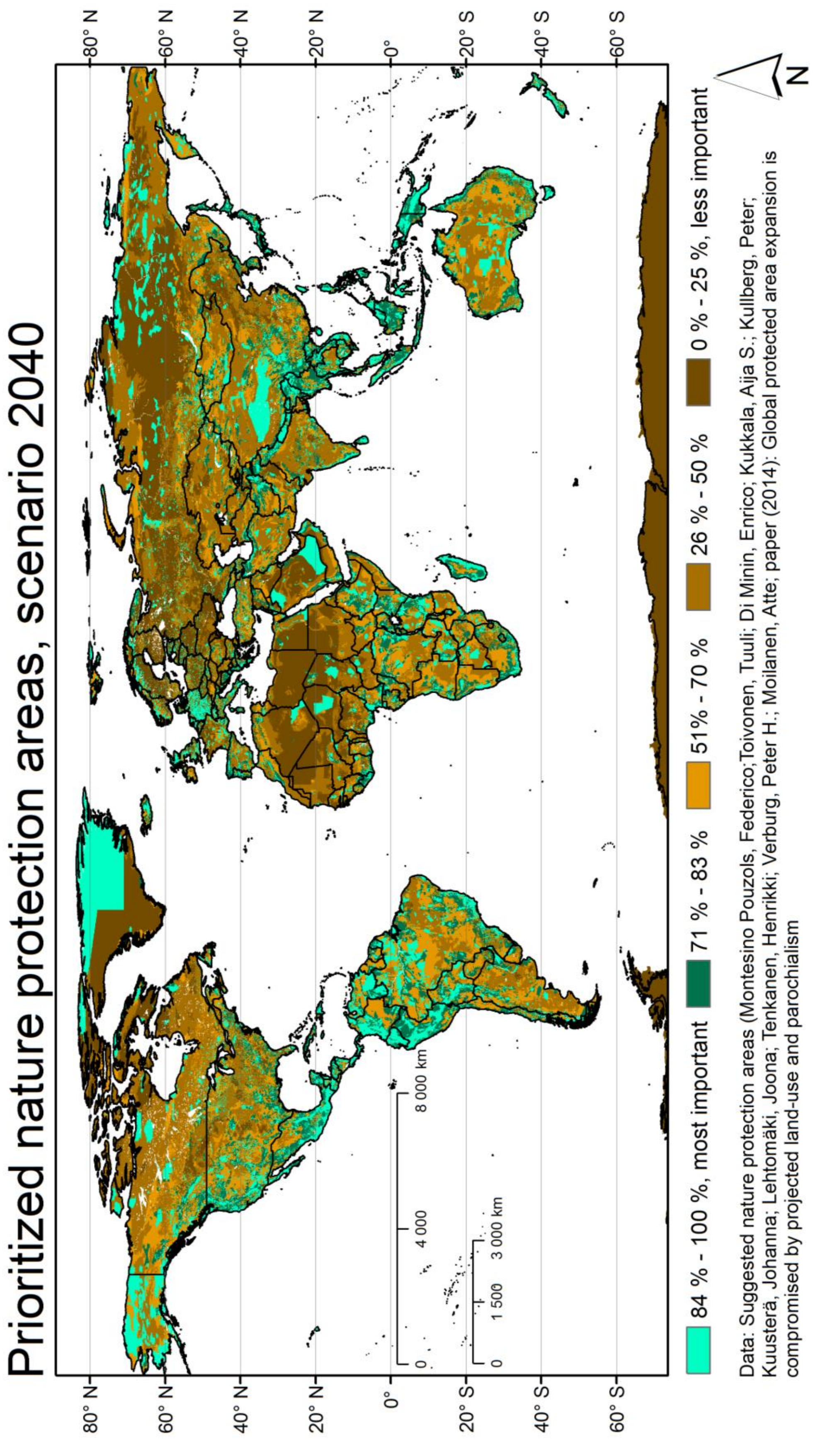


Figure 26. The global dataset used in the model to protect threated vertebrates.

3.3 Methods

The new model built in this thesis is presented in Figure 27. The dataset in brown-colored boxes are presented in the Section 3.2 Global data and the method are presented next. Table 4 also shortly explains why the method is used and the “Local/Zonal/Global” column tells if the result value in a cell is depending on neighboring cells or not (Section 2.2.1 Map algebra). The grey area of the model is including further development ideas (see Section 5.5 Future development).

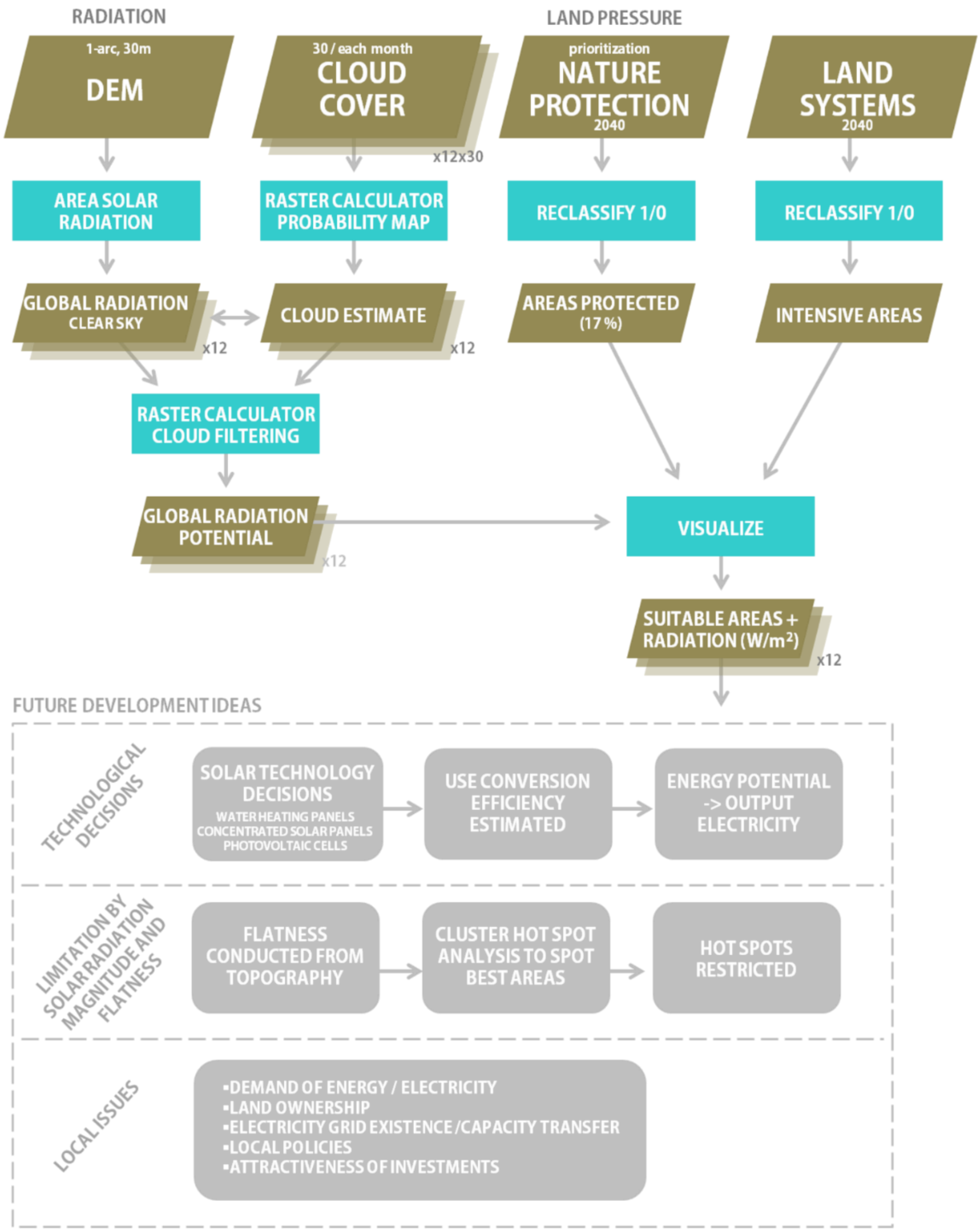


Figure 27. Datasets and methods for assessing optimal locations for utility-scale solar production.

All the datasets are put in Esri's database format file geodatabase (.gdb), as a raster or mosaic dataset to make the calculations more quickly. A file geodatabase is applied in ArcGIS software, and it makes it possible for datasets to be arranged by indexing logic. This is very important, when large datasets are used and need to be divided for calculations. Some of the tools of ArcGIS will not be enabled if the data isn't in the format of file geodatabase.

Table 4. Short explanations of the purpose of the method and its relation to neighboring cells.

Method	**Why used?**	**Global / Local / Zonal**
Area solar radiation	Calculate radiation estimates	Global
Probability map	Calculate cloud cover estimates	Local
Overlay	Exclude unsuitable areas	Local
Boolean / Reclassify	Decide which are suitable and unsuitable	Local
Future:		
Cluster analysis	Find "large enough" hotspots for stations	Zonal
Hotspot analysis	Find "large enough" hotspots for stations	Zonal

3.3.1 Area solar radiation tool

Fu and Rich (1999) developed Solar Analyst toolsets used within ArcGIS. These tools can calculate global (also called total) radiation and radiation types separately: direct and diffuse radiations. Estimates can be calculated for either points or areas. Behind the tool is a hemispherical viewshed algorithm that captures the topographical barriers to solar radiation, like mountains. Beside the DEM that is one of the inputs, diffuse radiation proportion and atmospheric transmittivity should also be defined as parameters (Fu & Rich 1999). More about the theory of radiation and the tool is in the theory section.

There are various competing tool providers for calculating solar radiation on desktop programs. Open source tools are provided, e.g., by GRASS with r.sun tool, and by SAGA-GIS with daily insolation, whereas commercial tools are provided, e.g., by Esri with Area solar radiation tool and by GeoModel Solar with their photovoltaic simulation software SolarGIS. Esri's Area solar radiation (later ASR) tool is chosen here, since it is easy to use and well documented by Pinde Fu and Paul Rich in their conference proceeding design and implementation of solar analyst (Fu & Rich 1999). More importantly the quality of estimation of the ASR tool has been validated by numerous researches (Table 5) and in the Global data section with DEMs.

Table 5. These studies have used Solar Analyst extension's toolsets for calculating estimates of global radiation. Most of mentioned are also in Holstein's list (2015).

PURPOSE	LOCATION	AUTHORS	YEAR
Soil temperature	Gunnison Basin, Colorado, U.S.	Fu and Rich	2000
Wildland fire fuels mapping	Moscow Mountain, Idaho, U.S.	Falkowski, Gessler, Morgan, Smith & Hudak	2004
Variation in canopy composition	Mont St. Hilaire, Quebec, Canada	Arii, Hamel & Lechowichz	2005
Solar energy in complex topogr.	Sierra Nevada Natural Park, Spain	Tovar-Pescador, Pozo-Vázquez, Ruiz-Arias, Batlles, López & Bosch	2006
Rooftop photovoltaic potential	Boston Massachusetts, U.S.	Esri (Boston Redevelopment Authority)	2008
Rooftop photovoltaic potential	Honolulu, Hawaii, U.S.	Meder & Pennetier	2008
Solar energy in complex topogr.	Sierra Nevada Natural Park, Spain	Ruiz-arias, Tovar-Pescador, Pozo-Vázquez, Alsamamra	2009
Solar radiation in built environments	Xinjiang, China	Li	2009
Concetrating solar power	Wilayat Duqum, Oman	Charabi & Gastli	2010
Rural solar technologies	Poultney, Vermont, U.S.	van Hoesen & Letendre	2010
Residential photovoltaic potential	Durham, North Carolina, U.S.	Mulherin (Thesis)	2011
Terrain and cloud impact on solar radiation	Pakistan	Shahzad, Renguang & Iftikhar	2014
Terrain impact on solar radiation	Hiamusi, Heilongjian, China	Mei, Fan & Mao	2015

In the theory section (Section 2.2.2 Measuring solar radiation) the fundamentals of the solar radiation calculation have been introduced, if the background of the model is of interest. In the following, the parameters of the ASR tool are handled.

DEM is SRTM 1 with 30-meter, or more precisely, 1 arc-second resolution. DEM is a mosaic dataset and the tool calculates the average latitude of the tile set, which is 20th parallel south in Harare. The tool can take only one latitude's parameter, which makes it necessity to handle one latitude at a time. The validation area of Greater Harare is from 17th parallel south to 19th parallel south. The inaccuracy of the northern and the southern areas is accepted. It is tested by the tool makers that 200 sky-size resolution is very accurate, and that the 500 resolution brings only 0.0001% more accurate results and multiples the calculation time radically. The whole year is calculated by monthly interval, and the year has been chosen to be 2040 by the forecast timing of the other datasets. The calculation day is in the middle of the month, and the frequency during that day is half an hour. Outputs are created for each month. (Figure 28)

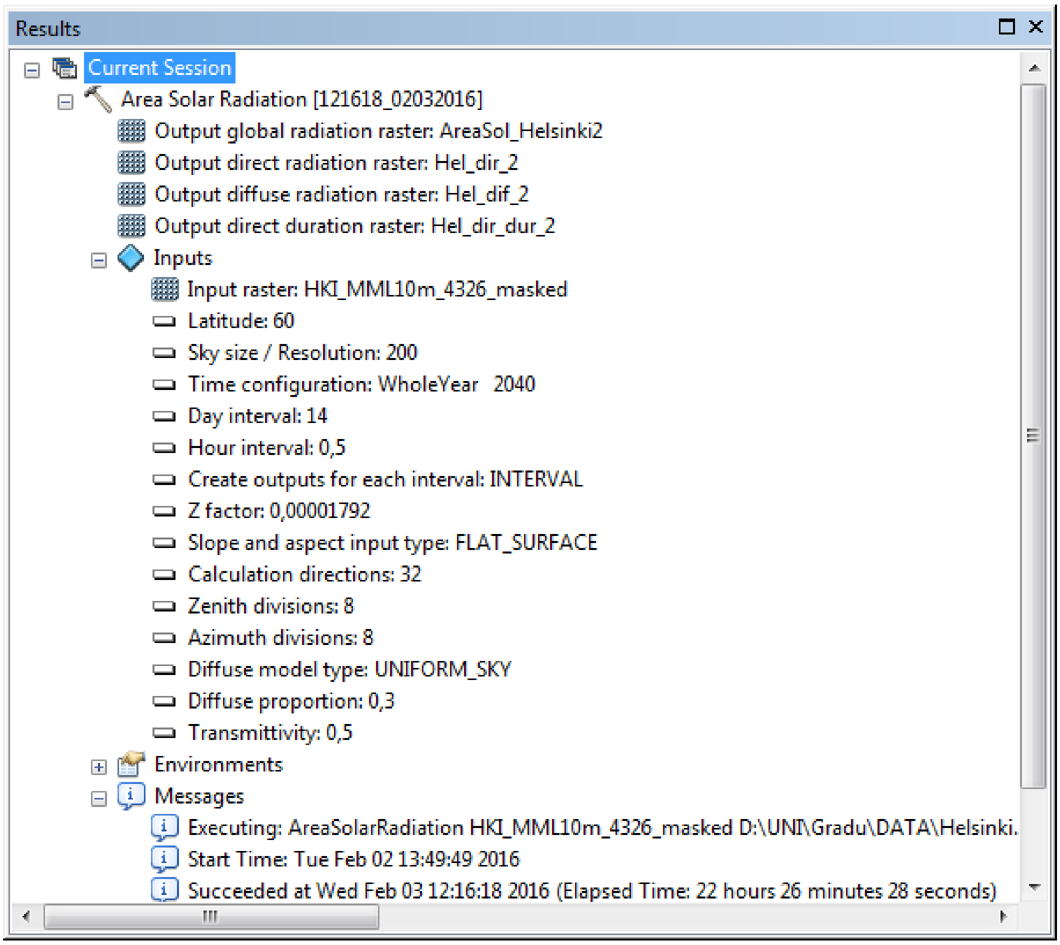


Figure 28. Area solar radiation tool's parameters, input DEM and outputs requested.

Next comes the topographic parameters. The Z-factor has to be entered because the coordinate system is in geographical degrees, and the phases of ASR need parameters in meters. One meter is 0,00000956 degrees on the 20th parallel south, whereas 1 degree on the Equator is 111,3 kilometers.

In this phase of screening possible areas for solar stations, global radiation is calculated for the flat surface. After screening out the impossible, high land-use pressure areas, more accurate tilted solar panel angles can be calculated in the interesting areas. This work wll not cover that phase. Calculation directiosn are chosen to 32 as recommended in the ASR tool for complex terrain areas. The radiation parameters are otherwise kept as defaults, since these are accurate and beneficial to the time consumed with making the calculations and. Diffuse proportion and transmittivity are defaults for clear sky conditions. Only the diffuse model type has been changed from uniform sky to standard overcast sky. This means that the incoming diffuse radiation flux varies while the zenith angle changes. The uniform sky would keep the diffuse radiation flux constant.

Direct, diffuse and global radiation, as well as the duration of direct radiation, are the output layers. On each month for each type an output raster is produced. All the outputs are requested for observational purposes. If direct radiation creates a considerably different pattern, the main radiation assessed should be changed to direct radiation since direct radiation is the main source of power in solar technology, as explained in the theory section. The global radiation is chosen to be the presented radiation. When deciding the classes and colors to the output maps the distinction between different values should be clear but to make the cities comparable the solar radiation values need to follow the same classification schema. Classification method chosen is equal interval to ensure readability for decision makers. As a tool for classification is used a draft (Figure 29) of historgrams of the resulted values of solar radiation in Harare, Denver and Helsinki and the results are presented in forms of maps (Figure 30-32).

Figure 29. A sketch to plan classification and its coloring.

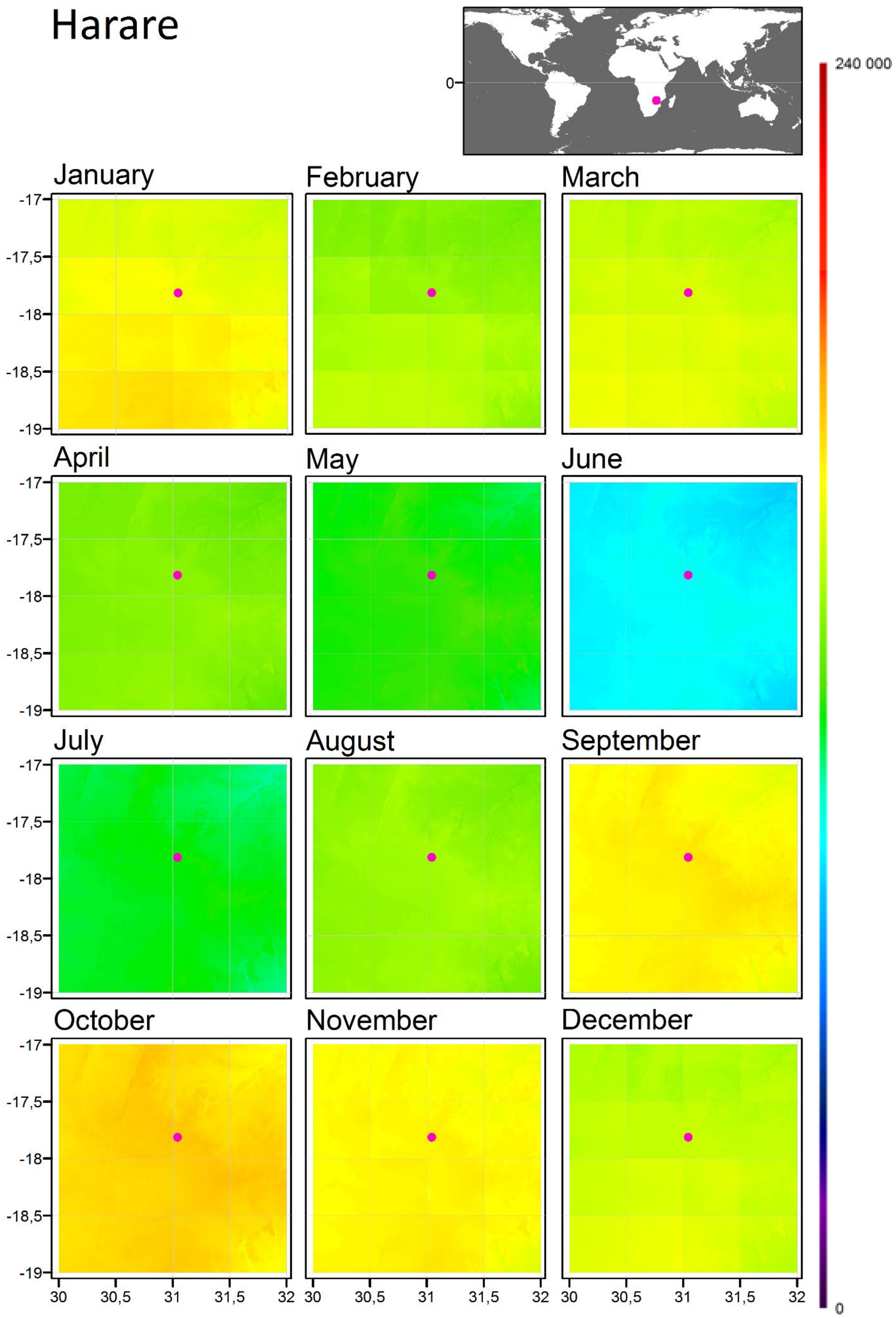


Figure 30. Final solar radiation estimates in Harare (Wh/m^2).

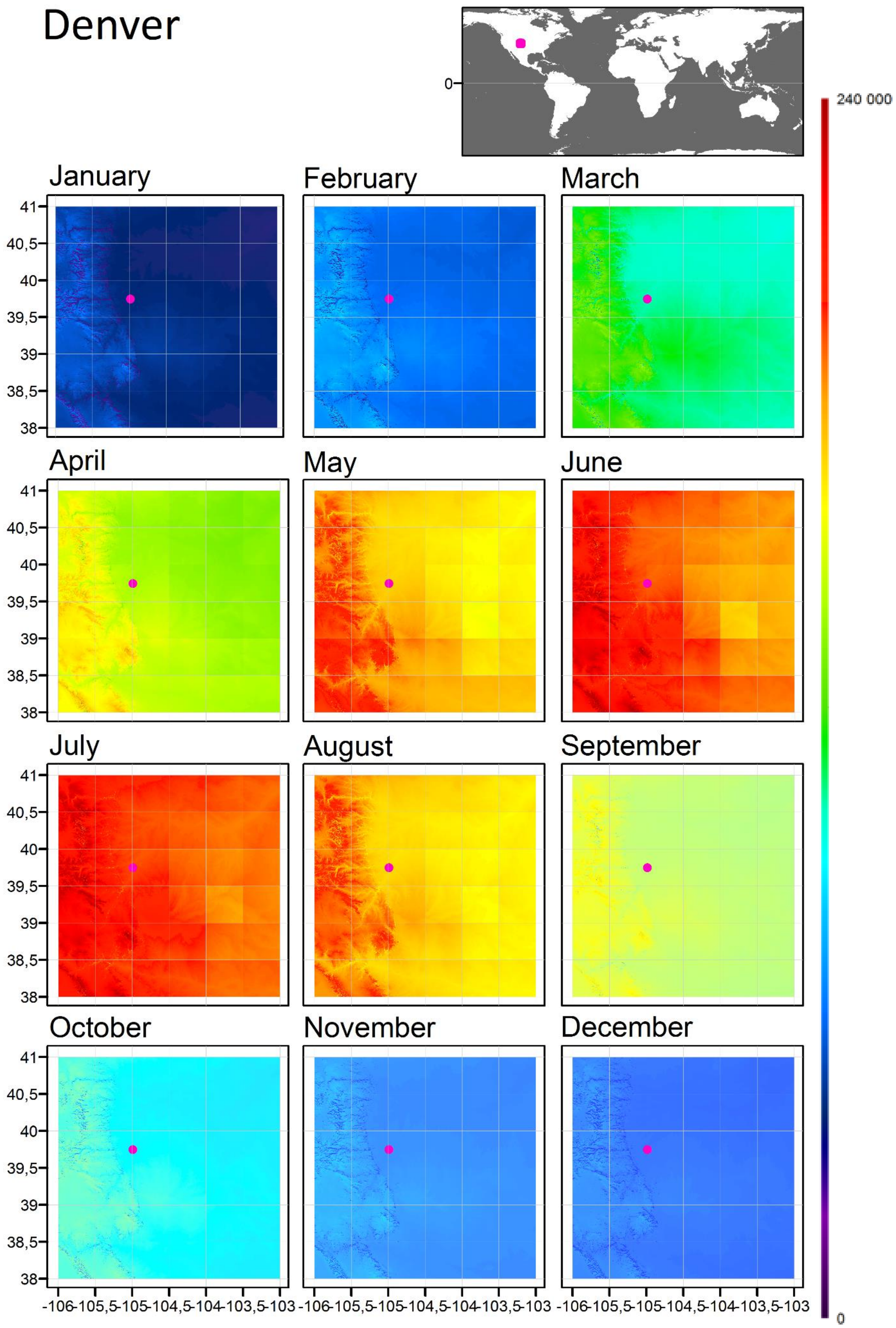


Figure 31. Final solar radiation estimates in Denver (Wh/m²).

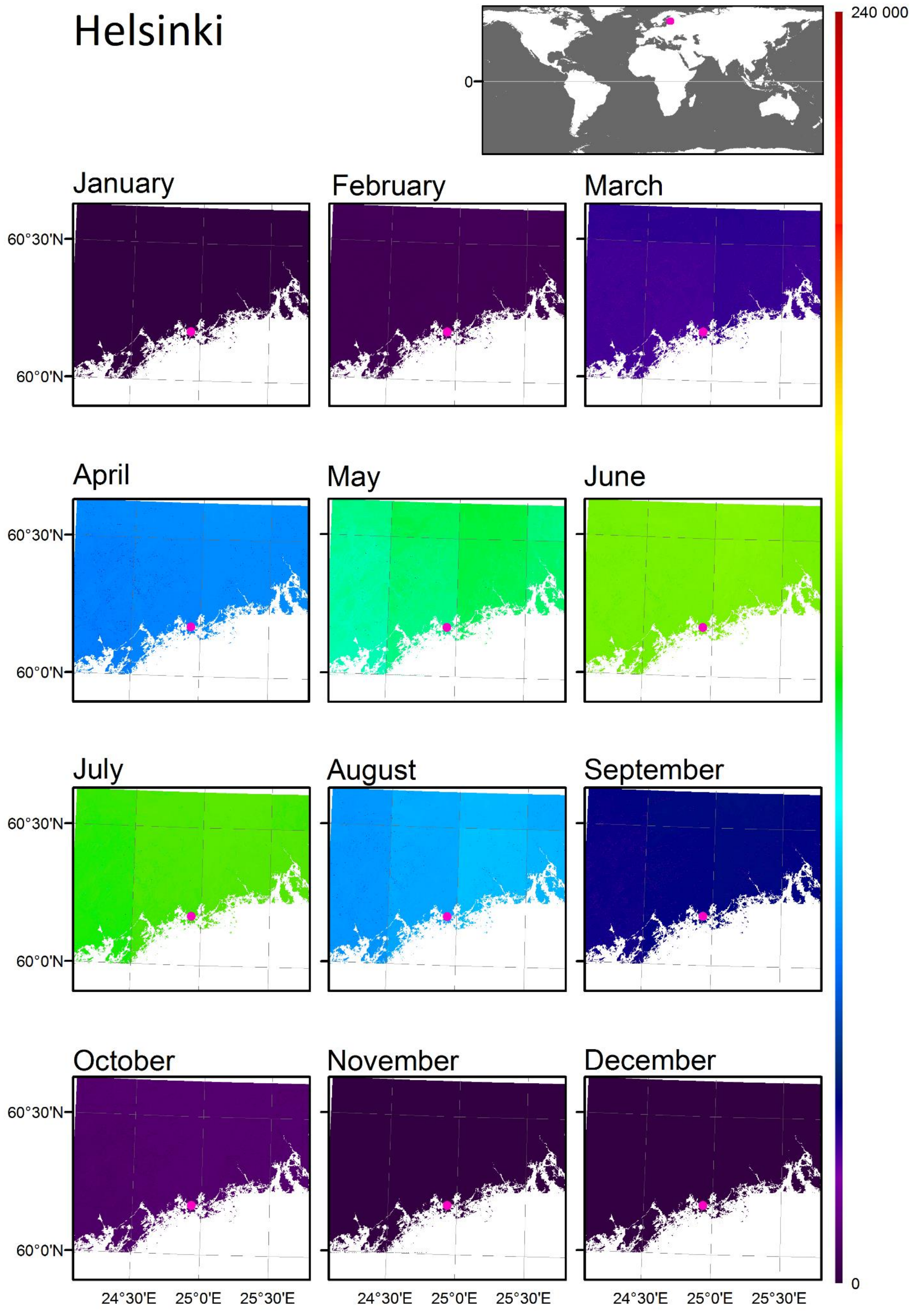


Figure 32. Final solar radiation estimates in Helsinki. (Wh/m²)

3.3.2 Probability mapping for cloud coverage estimates

Cloud coverage estimates are calculated by the raster calculator tool in ArcGIS software. When assessing radiation, one certain uncertainty is cloud coverage and its thickness. This is the reason why radiation estimation should include estimates of cloud coverage. There are experimental data of cloudiness available globally with half-degree (0,5) resolution. Data has a monthly average through years 1900 and 2009, but this study used a 30-year data period from 1980 until 2009. It is an official period of time for defining weather. Even though 30 measurements aren't that large a sample in statistics, it can be assumed that weather data, especially monthly averages, are nearly normally distributed. This is an opportunity to use formulas that are available for experimental data that imitates the normal distribution curve.

Probability mapping means that stochastic functions are used in calculating map values. In raster layers stochastic functions are easy to implement by raster calculator tool, which uses map algebra, especially when the data is compatible. For further background, see Section 2.2.1 Map algebra in theory section.

Since radiation estimates are better too low than too high from the perspective of production capacity, the goal is to find a probable moderate upper limit statistic. We can assume that cloud cover value (per mil, ‰) in the same raster cell on the same month would start to follow a normal distribution, which is a theoretical distribution, while values grow infinitely. This is essential for using classical statistical inferences in estimation.

A one-sided confidential interval is used for estimating a statistic for each cell separately. By using a one-sided confidential interval it is possible to estimate by probability (here 97.5 %), the average statistic of future measurements, which is less than the upper limit statistic that is the result of the test. The upper limit statistic in this thesis' model is a conservative cloud cover estimate. In practice this value underestimates solar radiation, and the figure of probability wanted could be loosened up to 90%.

One cell in the study area of Harare can be an example for how the one-sided confidential interval is calculated. There is a monthly cloud cover average in per mil (‰) from each January during the

30 years (1980-2009), altogether 30 measurements on one cell. The phases (Figure 33, 34) for calculating the statistic are:

- First empirical mean is calculated by summing up the next values and dividing them by 30 (495, 600, 603, 633, 636, 639, 641, 642, 654, 658, 667, 672, 673, 684, 684, 690, 693, 702, 705, 709, 712, 713, 713, 715, 722, 728, 731, 750, 760, 780)
- Then residuals are calculated; those are the differences between observed values and empirical mean (residual = 680.1333 – observed value)
- Residuals are summed up after they are squared, and then the sum is divided by degrees of freedom, which is the sample minus one (30 – 1 = 29)
- The last figure obtained is the sample variance (s^2) and when its square rooted a standard deviation.
- A one-sided upper confidence limit is obtained when a probability degree is chosen (here 97,5 % → 1,96) and standard deviation and sample size is placed to the right side of the formula, right from a one-sided upper confidential limit sign, μ
- The answer is next: it is 97,5 % probable that the average of cloud cover measurement for Januaries is the coming years is less than 77,23 %.

$$s = \sqrt{s^2} = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(y_i - \bar{y})^2}.$$

Figure 33. Sample standard deviation, which is inserted to the formula.

$$\bar{x} - 1,96 \cdot \frac{s}{\sqrt{n}} \leq \mu \leq \bar{x} + 1,96 \cdot \frac{s}{\sqrt{n}}$$

Figure 34. 1.96 is a constant for standardized normal distribution curve, when finding the upper limit of 97,5 % one-sided confidential interval. This is tool for finding confidential intervals.

In the model, this calculation is done for all months and all cells, and it is automated due to raster calculator abilities. These datasets that are estimates of the cloud coverage of each month are ready to be applied to global radiation estimates, which is the next phase.

In Figure 34 it is seen how cloud coverage (n) impacts global radiation. If cloud coverage is from 0% up to 50%, the insolation of the global radiation is still over 90% of the original global radiation. When cloud coverage rises over 60% thickness, the global radiation the reaches the ground starts to rapidly decrease. The formula to handle cloud coverage estimates is used by NASA (Shodor 2016).

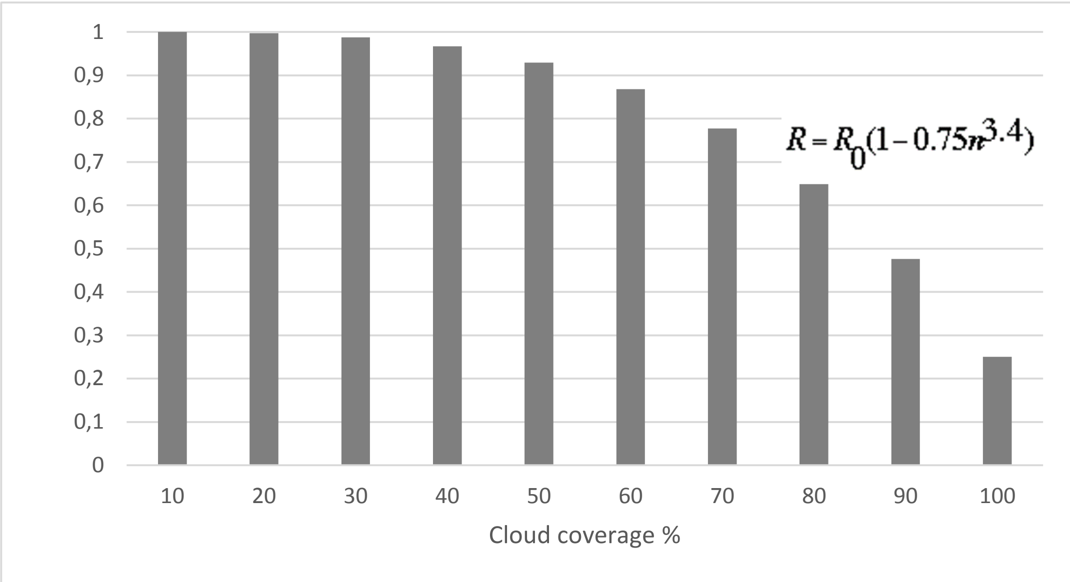
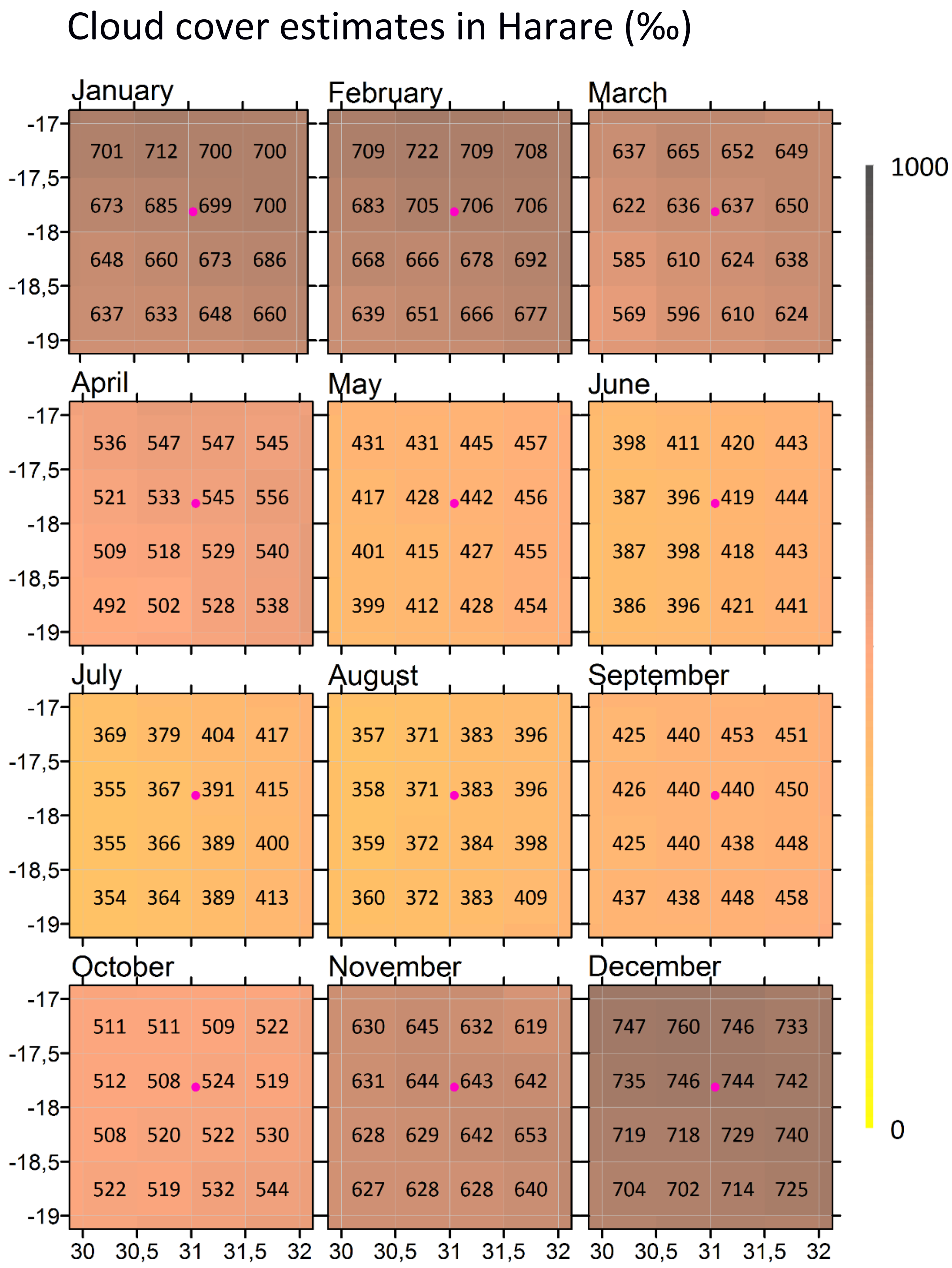
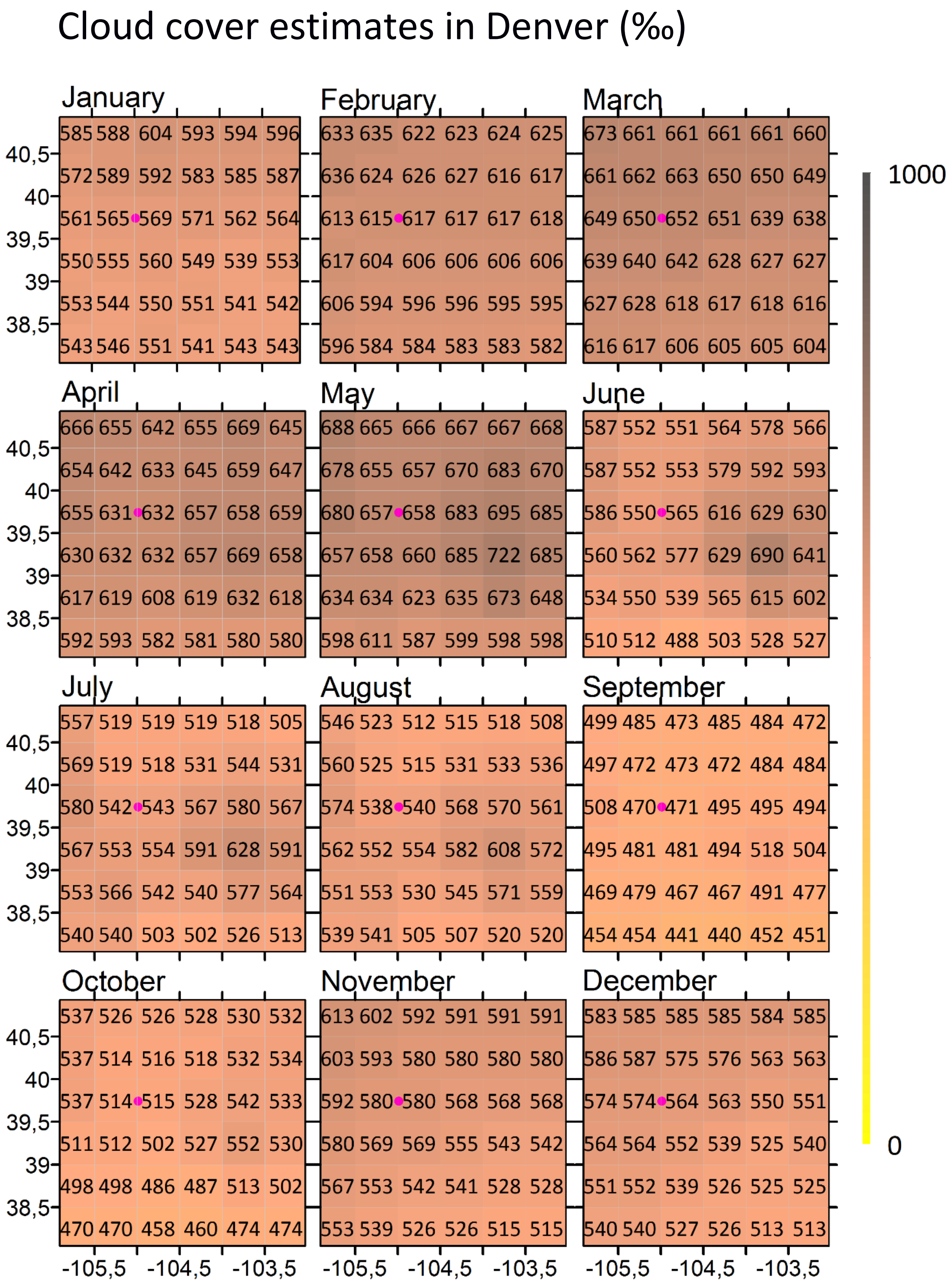


Figure 35. Cloud cover lowers the amount of global radiation reaching the ground. R indicates global radiation after clear sky global radiation (R_o) is filtered by cloud coverage (n=[0.00-1.00]).



Data: Original monthly average cloud cover 1901 - 2009 (CRU-TS v3.10.01 Historic Climate Database for GIS, the University of East Anglia), data downloaded 6/2015.

Figure 36. Harare's cloud cover estimates for each month (0-1000 ‰).



Data: Original monthly average cloud cover 1901 - 2009 (CRU-TS v3.10.01 Historic Climate Database for GIS, the University of East Anglia), data downloaded 6/2015.

Figure 37. Denver's cloud cover estimates for each month (0-1000 ‰).

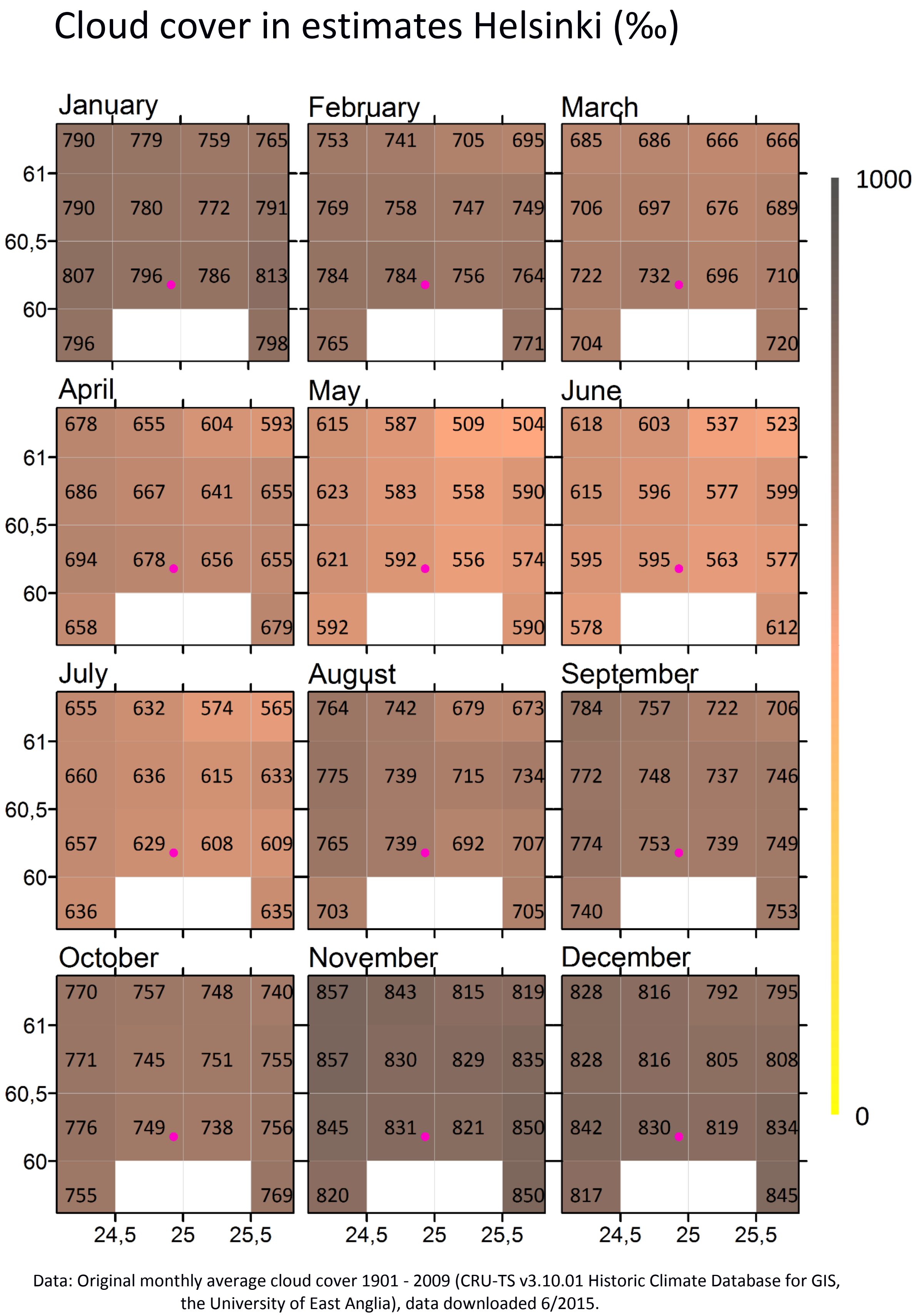


Figure 38. Helsinki's cloud cover estimates for each month (0-1000 ‰).

3.3.3 Combination: overlay

Overlay is one of the most used geographic information analysis types. A simple definition of overlay is that two vector or raster datasets with their attributes are joined (Harvey 2008). This model in particular uses these simplest forms of overlay analysis. Three dataset were joined in a manner that two of the dataset restricted unsuitable areas. These two datasets were nature protection data (Figure 40) and land use intensiveness (Figure 39). The third dataset, estimates of the global radiation is not restricting the areas in this thesis, but is planned to be important restricting factor later. In the overlay the global radiation is the only layer that preserve its original values.

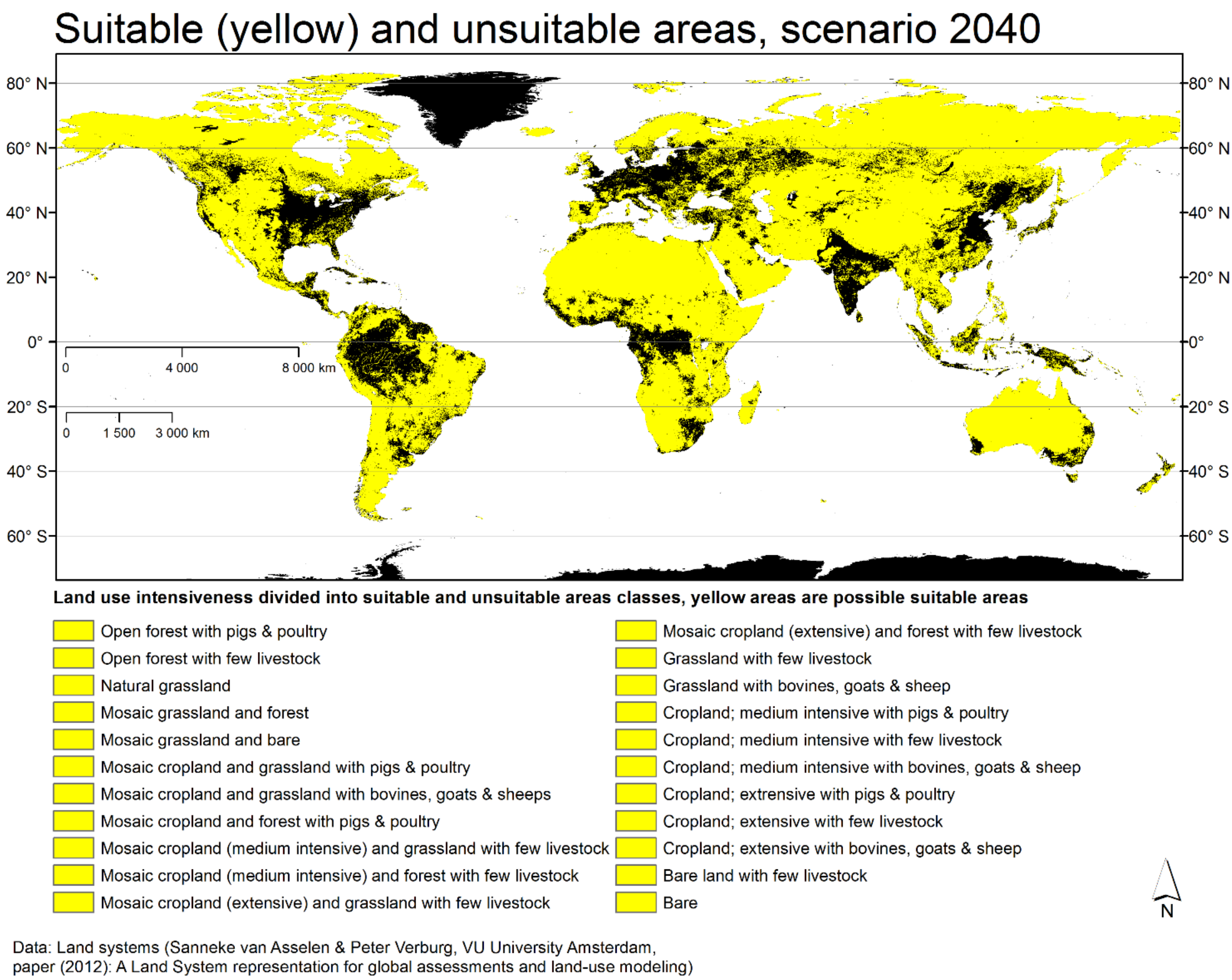


Figure 39. The areas of the world are divided into suitable and suitable in this thesis using the hierarchical classification of land use intensiveness.

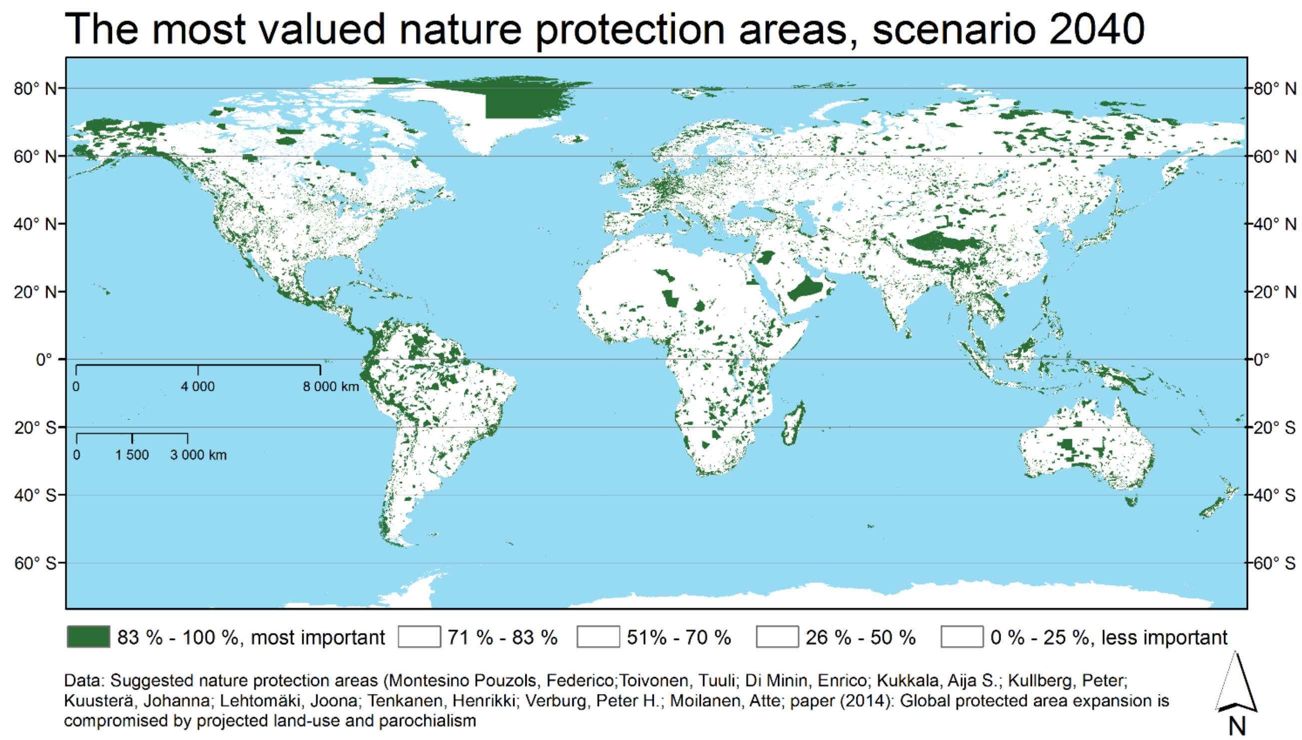


Figure 40. The unsuitable areas concerning the model are in green. Those are the most prioritized areas from the perspective of threatened vertebrates.

4 RESULTS

4.1 Model building process

The GIS model building process is compressed to form an iterative PPDAC process frame (Figure 41). The problem phase (1) includes clarification of the need for the model, preliminary discussions with proper and as a result of discussions finding a niche. The plan phase (2) is more a boundary crossing unit. The data (3), analysis (4) and conclusion (5) phases all have a strong impact on the plan phase. The plan covers the assessment of GIS as a tool for the conclusions and results wanted. In the plan phase, earlier studies are looked through on a level needed to make the decision to continue model building. The criteria of the model has to be decided on some level in the plan phase in order to determine a framework for finding correct data and methods of analyzing. The process the model building follows is reasonable to choose already in the plan phase. This model was planned to be built using PPDAC model in this phase.

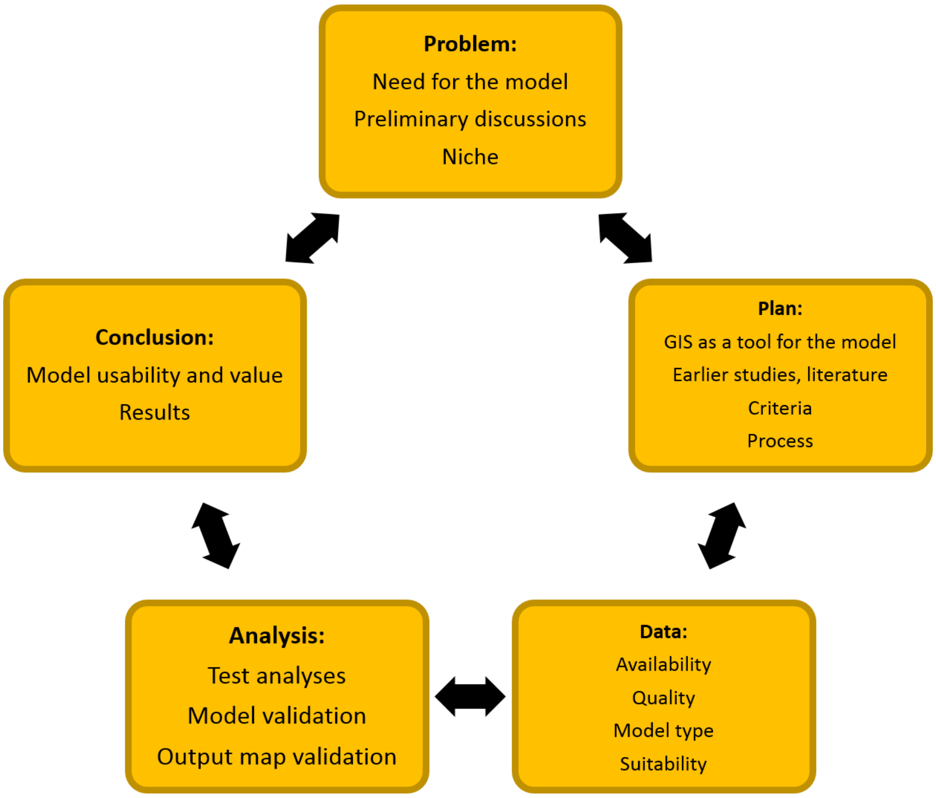


Figure 41. The iterative PPDAC process applied to the model building.

The data phase (3) is an operative phase as is the analysis phase (4). Availability, quality and suitability of the data is investigated in the data phase. The GIS model type is decided after examining the possibilities the datasets offer. In the operative analysis phase the test analyses are executed to find out if the model offers satisfying information considering the resources used (e.g. time, money). The model should be validated by the outputs it returns. Next sections of this thesis are dedicated to that. The conclusion part (5) of the PPDAC frame concentrates on the model's ability to encounter the niche from the problem phase. If the usability and the value of the model does not match the need of the model the process continues. This is an iterative process and the boundaries of the phases can be crossed when necessary.

4.2 Output maps from the location optimization model

The location optimization model cuts the suitable areas for solar power plant siting remarkably. The size of validation areas i.e. greater city areas varies greatly due to tile size of DEM and placement of the city center in relation to mentioned DEM tiles. Said that, the potential suitable areas in Harare are cut 60 % (study area size 46 840 km^2), in Denver 45 % (98 000 km^2) and in Helsinki 70 % (7 200 km^2). The land use pressure factors (See Figure 41) determine areas that are cut out. The pipeline of the model is presented in Figure 42 with the case of Harare in January. The legend for the "Case

Harare" is usable in Helsinki and Denver as well (Figure 43). The actual output maps including information for 12 months for each of the three cities are in Figures 44 to 46 and detailed data inspection is done in the next section. The matter of the exact suitable areas in each of the cities is not addressed here per se. This is because the necessary threshold of the global radiation magnitude for solar power plants is not determined within this thesis. The following five pages present the pipeline and the output maps.

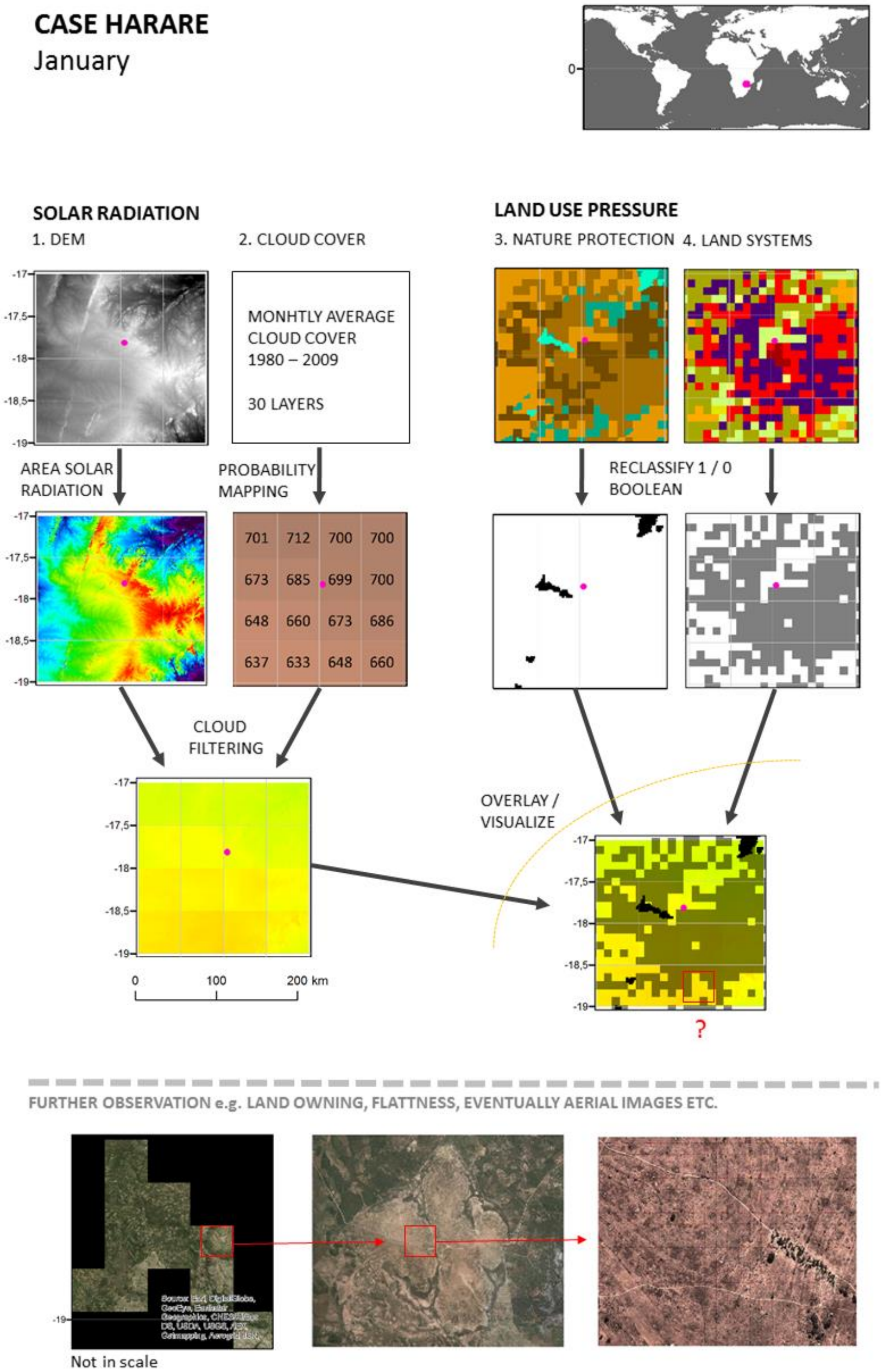


Figure 42. An example of the model applied to the Harare city area in January (legend: Figure 42).

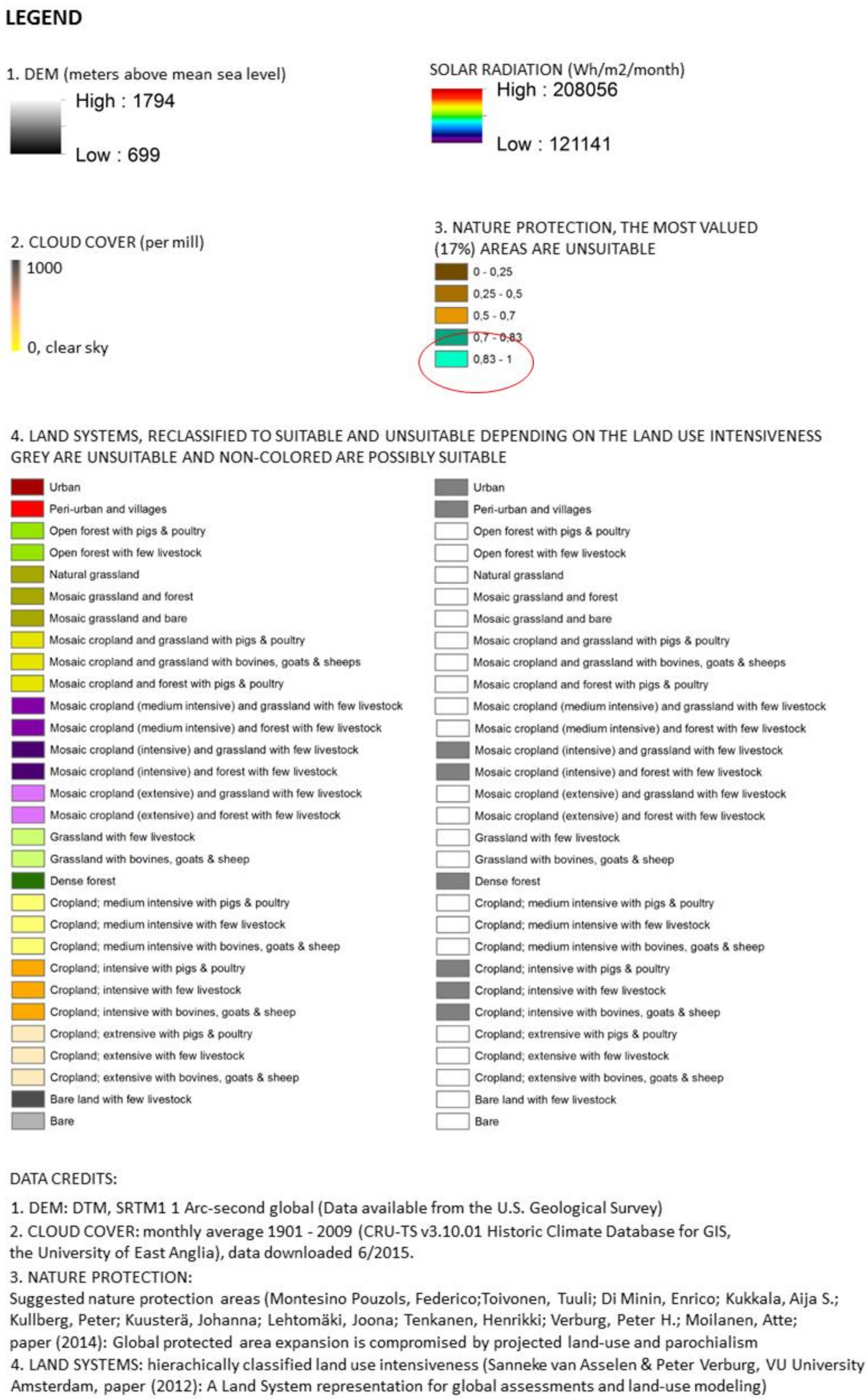


Figure 43. Legend for the data in Figure 42 which can be used to Figures 44-46 as well.

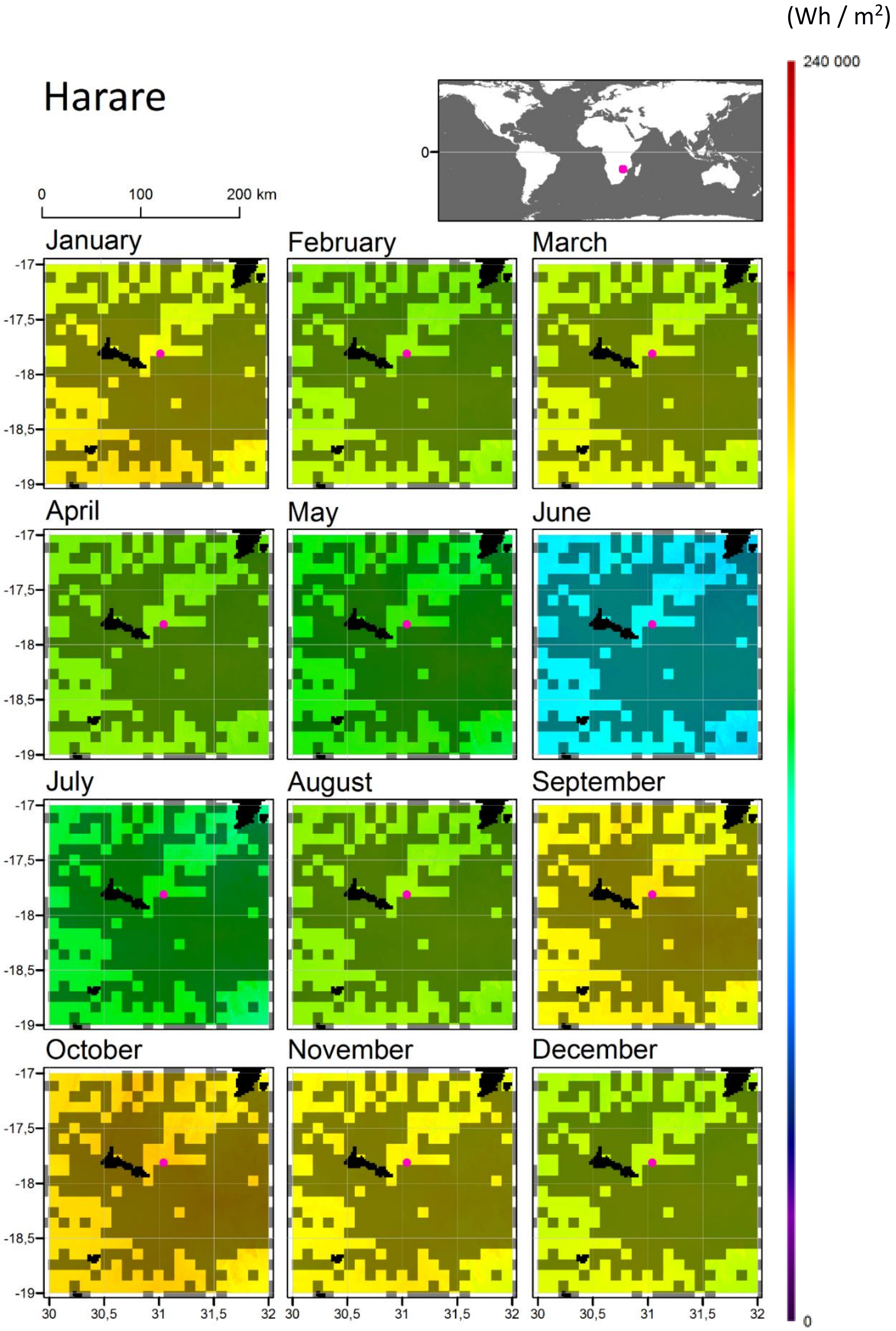


Figure 44. The model output maps in Harare. Global solar radiation is the colorful layer (0 - 240 000 Wh/m²). Black and grey areas are unsuitable for further location investigation.

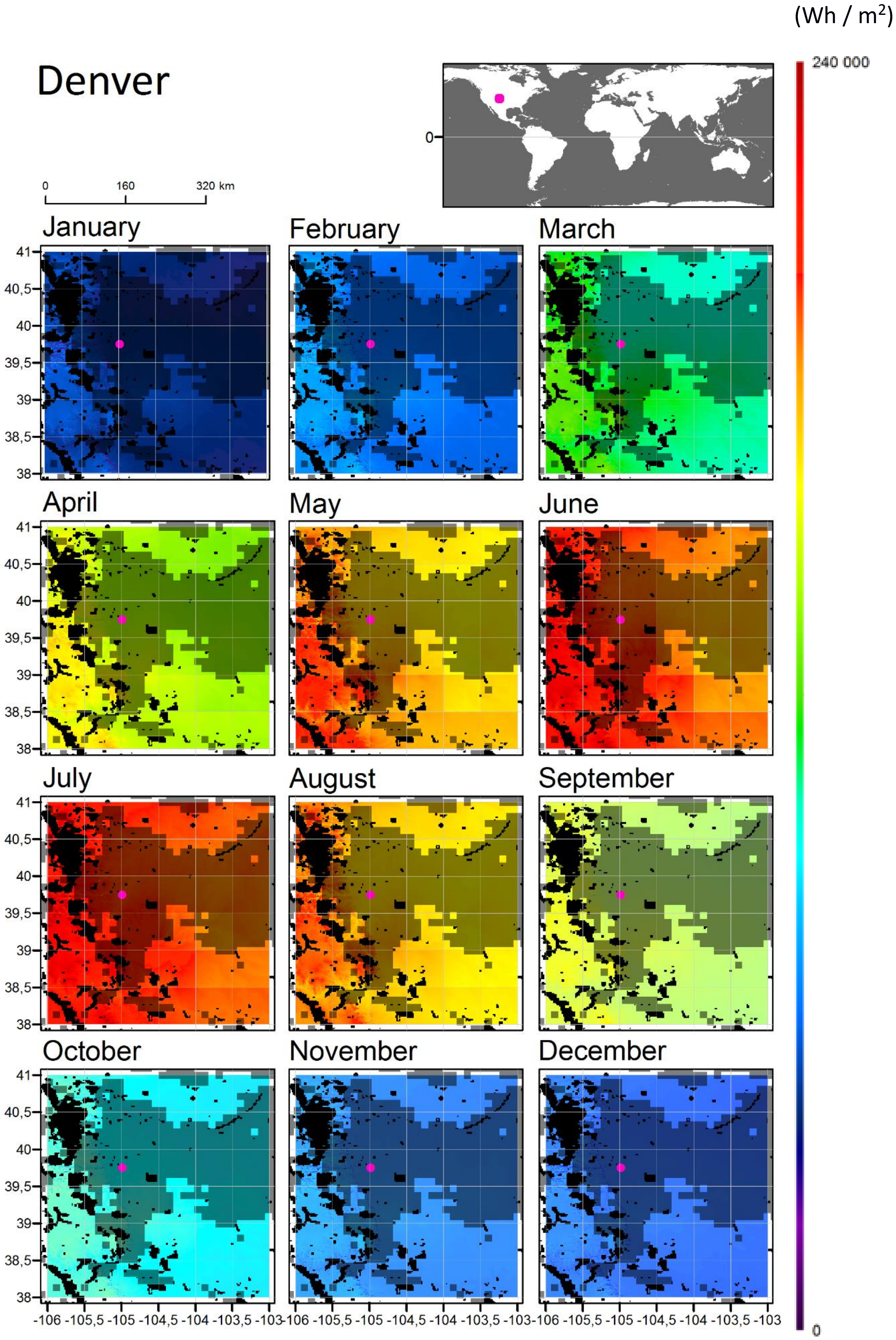


Figure 45. The model output maps in Denver. Global solar radiation is the colorful layer (0 - 240 000 Wh/m²). Black and grey areas are unsuitable for further location investigation.

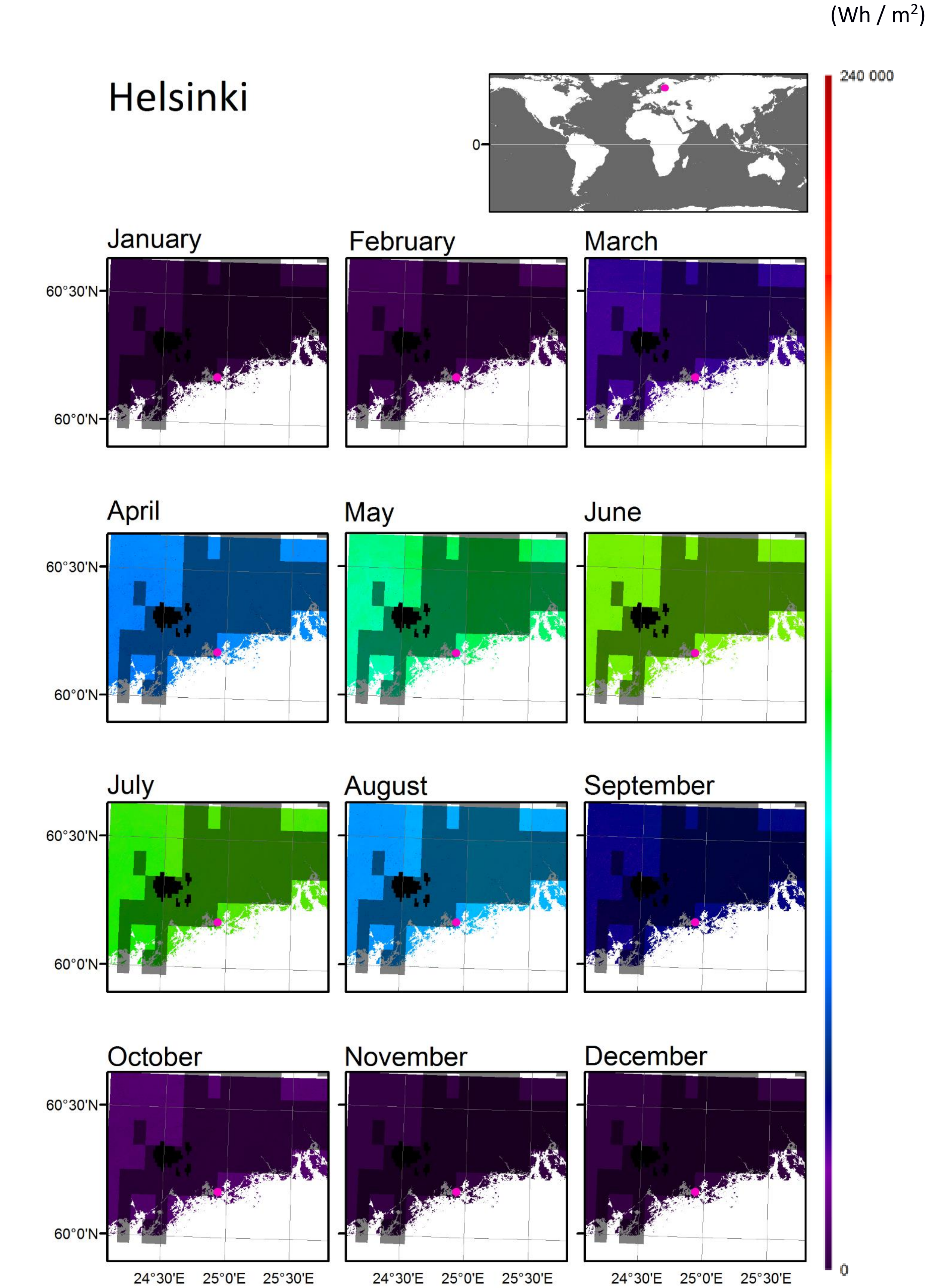


Figure 46. The model output maps in Helsinki. Global solar radiation is the colorful layer (0 - 240 000 Wh/m²). Black and grey areas are unsuitable for further location investigation.

4.3 Trends in output data

This section presents data outputs from the model proceedings in three validation areas: Harare, Denver and Helsinki. Temporal and spatial distinctions within a city and between cities can be measured when comparing the magnitude of global radiation. Clear spatial patterns can be noticed from the restricted areas which are a result of binary classification for land use intensiveness (Land systems) to suitable and unsuitable areas. Urban areas, semi-urban areas and intensive agriculture and livestock areas are located in the immediate surrounding of the cities forming continuous surfaces starting from the city center and classified as unsuitable areas for solar power plants.

Harare, Denver and Helsinki got clearly distinct values in global radiation magnitude (see Table 6 and Figure 47). The maximum values are the most interesting since solar power plants are seeking for the maximums. Maximum global radiation values in validation cities in January varied the most compared to the maximums of other months. Helsinki's maximum value in January was 426 Wh/m^2/month, which is the amount of electricity needed to watch a plasma TV less than half an hour. Corresponding figure for Denver is 70 000 Wh/m^2/month and Harare 171 000 Wh/m^2/month, although, Harare's summer season is in January in consequence of its location on the southern hemisphere. The distinction in June was decreased between Harare (134 178 Wh/m^2/month) and Helsinki (94 294 Wh/m^2/month), but Denver gained most global radiation (258 724 Wh/m^2/month) in comparison between different seasons and cities.

Table 6. Statistics of the final solar radiation estimates (watt-hour/m^2/month).

	January			June		
	max.	**min.**	**mean**	**max.**	**min.**	**mean**
Harare	171 866	93 231	158 591	94 294	10 397	88 792
Denver	70 684	4 994	43 803	258 724	18 722	195 259
Helsinki	426	63	340	134 178	24 486	129 418

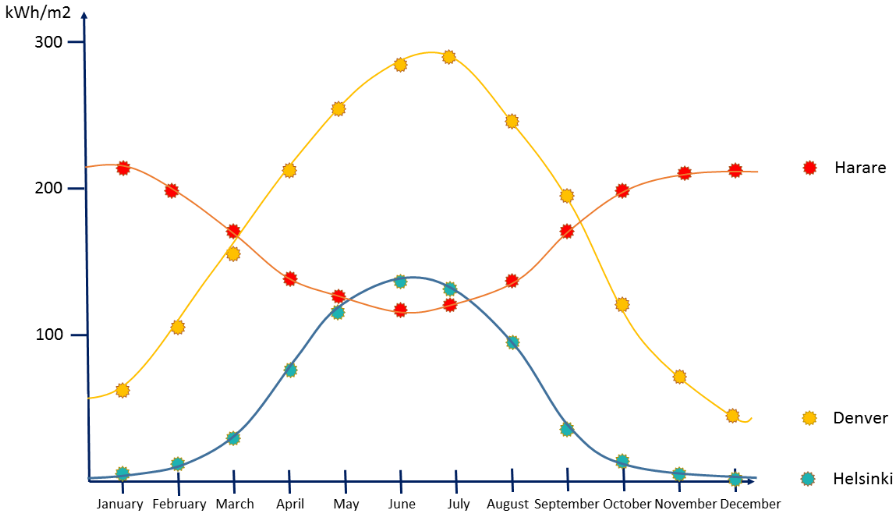


*Figure 47. Maximum global radiation for each month in each validation city. These values are the maximum values for a single 30m*30m-sized cell.*

The results handled above present temporal differences in the maximum global radiation in a single city and compare maximum values between different cities. The following results concern the global radiation output values within a single city from a spatial point of view (Figure 48). Spatially the biggest variance in global radiation is found in Denver in June, 60 000 Wh/m^2/month, whereas in June in Harare and Helsinki the variation is only around 10 000 Wh/m^2/month. Outliers were excluded here. In January the spatial variation are 20 000 Wh/m^2/month, but in Helsinki global radiation is between 200 to 400 Wh/m^2/month, which means in practice that solar production is on hold. There is significant spatial and temporal difference in global radiation values which makes it important to calculate the accurate global radiance using accurate digital elevation model.

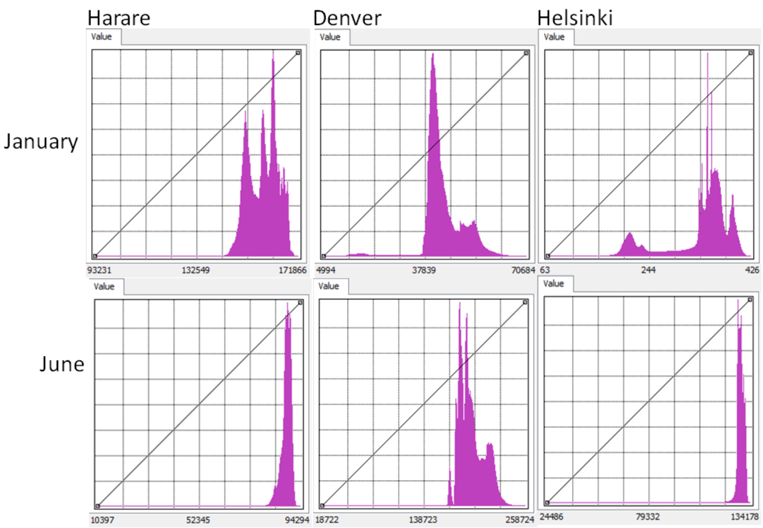


Figure 48. The final solar radiation estimates in Harare, Denver and Helsinki. The scales varies. To see statistics better see Table 6.

5 DISCUSSION

5.1 New GIS model fastens the process of location optimization

The new GIS prescriptive model constructed in this thesis provides an excellent decision making tool for solar power plant site investigators to restrict the unsuitable areas from the city under consideration. The results of this thesis show in Harare, Denver and Helsinki that the yet suitable areas cover up to 50 % of the examined areas. The model is comprehensively pre-validated by reliable datasets used in the model. Although the model restricts unsuitable areas efficiently, it does not address local issues such as landowning or public policies.

The model is targeted to be a useful application for the energy industry. Its efficiency will be proved when applied to industry-scale decision-making processes. This section underlines the major findings of the thesis, suggestions for future research, and the relevance of the findings to existing academic debate on the subject. These themes are addressed through the research questions presented in the introduction of the thesis. Here those are repeated:

Question 1: **What are the main criteria for a good location for utility-scale solar power plant?**
Question 2: **How to build a GIS model for solar power plant optimization?**

5.2 Main criteria for a good location for utility-scale solar power plants

The assessment of the main criteria to be used in the model is mainly based on the literature. The availability of the suitable datasets influenced the choice of criteria. As a result, these criteria were chosen: solar radiation magnitude, land use pressure (i.e., land use intensiveness) and nature protection. Other criteria that could not be applied to the model were investment attractiveness, planning and zoning regulations, landowning, and existing transmission networks. The main reason for excluding these important criteria are the lack of easily usable global datasets. The comprehensive list of the criteria can be found in Figure 6.

Solar radiation is a mandatory factor in the model, since it is the source of solar energy. Nevertheless, it became obvious during the process that investigations of the radiation magnitude have become non-unique (e.g. Schmalensee & Bulovic 2015; Shahzad et al. 2014, Tovar-Pescador 2006; Ruiz-arias et al. 2009; Fu & Rich 1999; Watson & Hudson 2015; Brewer et al. 2015; Mei et al. 2015; Breyer et al. 2010; Pogson et al. 2013). Different applications measuring the radiation are widely used in the industry and among decision-makers. Global radiation as a single factor, no matter how accurate, cannot alone add value to the research field or to the decision-makers in the industry. More geographical criteria are needed in order to get industrially valuable results.

A major finding in the study is that the dataset of land use intensiveness gives an efficient tool to screen the potential areas for solar parks in the greater city where the model is allocated. In the beginning of the research work, it was assumed that the land use pressure element would be a remarkable factor. This is based on the basic law of urbanization, in which the demand for the land is high near the cities. As the results of the model indicate, the possible suitable areas for solar energy production decrease already by even 50 percent in the validation areas (Harare, Denver and Helsinki). Since the dataset was unique of its kind, the results are not compared to other land use datasets, such as land cover, that will not tell as much about the intensiveness of the land use. According to the model, the odds of target investigations in the future of a land area that has great value decrease.

In the validation areas (Harare, Denver and Helsinki) there was no remarkable nature protection pressure according to the dataset used that would decrease the land area suitability for solar power plants. This is due to suggested nature protection areas being highly concentrated in certain areas where a great amount of vertebrates live; e.g., the tropical areas. No matter if this criteria was not "efficient" in order to restrict the suitable land; its importance will not collapse, as 165 countries in the world have pledged to save threatened vertebrates by area conservation. The resulting dataset of prioritized nature protection areas by Montesino Pouzols et al. (2014) is highly reliable, and the comprehensive implementation of the proposed expansion of the protected area network could save important areas from unwanted development and destruction of biodiversity-rich areas.

The closest model found in the literature compared to the model built in the thesis is Watson's and Hudson's (2015) model. The distinct value of the model to the thesis is that it is local due to local datasets, so it will not work globally, only in southern Great Britain. Watson and Hudson use multi-criteria decision-making and GIS for categorizing the best areas for solar and wind energy production on a larger scale. They use a binary model that takes into account agriculture, historically important areas, landscape, residential areas, wildlife, and solar radiation. The criteria common to both models are agriculture and residential areas at some level, wildlife and solar radiation. Nevertheless, their model will not work globally, but it has the important elements for closer investigations. The limitation of the model of this thesis is the limited accessibility of globally available and, simultaneously, locally interesting datasets.

The major finding itself is that the new model is built using important criteria of land use intensiveness, nature protection, and solar radiation estimates for 12 months separately. The model has proven to work, and industry shows interest in it with respect to the criteria picked.

5.3 Building GIS model for solar power plant location optimization

How to build a geographical information tool for more leverage in a decision-making infrastructure is the core learning process of the whole thesis. The process of building a model is iterative and experimental, and this is the reason why the PPDAC process model was chosen. It turned out to be an excellent tool for handling the model-building changes along the way. The model changed along with demand, open datasets, time, model testing, new findings from the literature, and understanding the problem better. This is the process that is described in the results (Figure 41).

At the first phase, the need for the study was examined and evaluated. GIS expertise is an excellent starting point for discussions with possible end-users of the location optimization model. The idea of a location optimization model that helps decision-makers in the energy industry started to arise in these discussions. It seemed that geographical information is a requirement for a good decision-making instrument in terms of optimizing solar energy production. Specifically, utility-scale production over residential solar energy production was chosen, since the demand for siting solar power plants is a current issue. The keen demand for information tools for utility-scale solar power plant location optimization was crucial to the decision to study this further. Later, the research problem developed as the thesis instructor pointed out the importance of land use pressure.
In the planning phase, the suitability of GIS tools was examined. This was done by researching the software capabilities and limitations, as well as investigating models created in earlier studies. In

order to choose the relevant criteria, theories on the fields of land use planning, solar energy, and site planning were examined. Choosing the criteria turned out to be a useful research question for the thesis. There are many GIS models that, for instance, targets siting in agriculture, factories or other environments, but not the combination of solar radiation, land use intensiveness and prioritized suggested nature protection areas (see Section 1.5 Earlier studies). A process flowchart of the model was drawn.

The final form of the model was highly influenced by the data phase. The study perspective turned from local to global during the research work. Investigating a certain small area would have afforded possibilities of pointing out the actual sites for solar power plants. Now instead, the model is a universal tool but always require further local studies, and knowledge of local policies, etc. The fact that the incredible valuable global datasets regarding land use intensiveness and nature protection were available with relatively high (1 km x 1 km) resolution led to building this global model. There were two factors behind the model type selection: the model is predictive, since it estimates solar radiation, and prescriptive since the results point to possible areas for further investigations.

The test analyses during the analysis phase caused iterative re-planning of the model. For instance, the study area had to be cut since processing of the radiation estimates appeared to be very time-consuming. The study area of Harare was in the end 200 kilometers per edge of the square. The validation of the model was not required because the datasets applied in the model had already been validated. The model and its output maps were tested in three different greater city regions to enable comparison of the results and thus better evaluation of the model.

In the last conclusion phase, it turned out that the results of the output maps matched the original need set in the first phase. The new model gives an efficient tool for decision-makers to screen suitable areas for utility-scale solar energy production. That is because land use pressure is a criterion that efficiently indicates the unsuitable areas. The model is new, validated and it aims at contributing to solar energy industry by providing information regarding land acquisitions.

5.4 Limitation

Even though the model works for its purpose there are several limitations in it. For instance, the information, who owns the land is crucial, but is not implemented in the model (see Section 5.5 Further development). One of the challenging parts of this work was that the model is designed for global use. Consequently, the data resolution and the themes of the data are not telling enough of the local situation.

The accuracy of the datasets is one limitation of the model. The resolution of the land use intensiveness dataset is only around 10 kilometers in the Equator area. This can exclude some partly suitable areas from the results. On the other hand, the solar radiation estimates are as optimal as GIS software can produce (see Section 3.2.1 Digital elevation model).

To ensure that the results meet the original need it is useful to iterate with “the product owner” during the research work. In this case, instead of the industry representatives, the instructor had the role of iterative discussions partner. The discussions with industry continues after the thesis inspection.

5.5 Future development

A locally used model needs various geographical datasets or other assessment tools that are not yet implemented in this model. Suggestions for future work are presented in Figure 49 below. As the model develops, experts in the fields of locating energy production, solar energy, and business strategy, as well as scientists from the fields of meteorology, land-use planning, and physics should be interviewed in order to create a more effective model to fulfill expectations regarding this kind of model. For instance, to consulting a meteorologist could lead to more a sophisticated model regarding cloud cover estimates than the NASA version currently used in the model. Another example of material for further discussion is intensiveness of land use. What is the threshold of intensive land use while planning land acquisitions?

The further development ideas are here divided into three different themes, although the "local issues" box includes subjects from various fields. Some of the development points include a need for geographical datasets such as land ownership or the flatness of the area, but some of the points are nationwide and no need for geographical differentiation exists.

The themes are not in chronological order. For example, investment attractiveness from the "local issues" is strongly related to hotspot areas, since the electricity generation has to be at a profitable level. Hotspots again are affected by estimates of electricity production capacity, when again the production capacity is related to solar technology and its ability to utilize certain types of solar radiation.

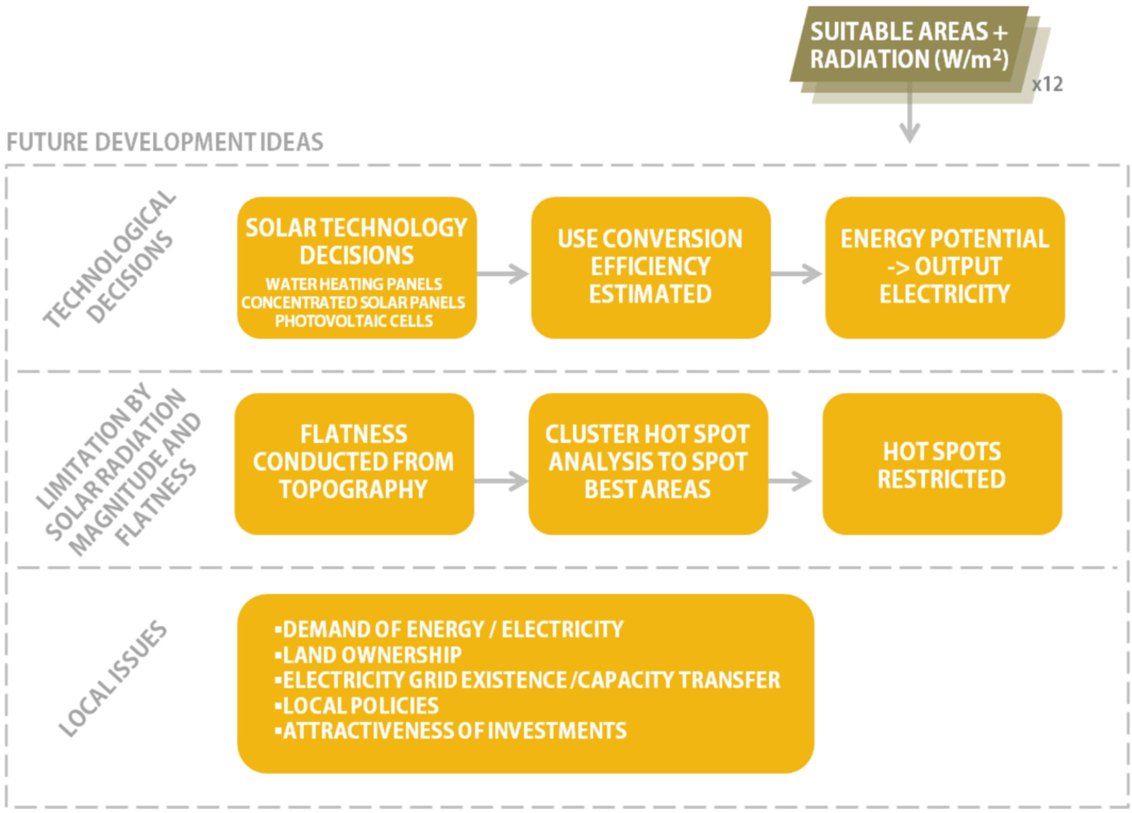


Figure 49. Ideas for further development of the model (part of the model, see the entire model in Section 3.3 Methods).

The local issues are strongly related to theory of DeMers (2002b) regarding siting. The book elaborates the different aspects of siting in land-use planning (see 1.5 Earlier studies). The aspects that are considered important from DeMers' list for further suggestions are zoning regulations, legal restrictions, and landowning. Also, aesthetic values are important when reaching a detailed level.

Different solar technologies are presented in the Introduction (see Section 1.1 Insights for solar production). The choice of certain technology affects the amount and type of radiation technology that can be utilized. This is then a crucial phase before estimating any output electricity amounts. For instance, some technologies may utilize reflected, diffuse, and direct radiation, but there are types of technology that use only direct, the strongest type of radiation. The material, shape, and tilt angle of the technology impacts this. The model built for this thesis estimated global radiation on a flattened space, since the technology was not chosen and no restrictions to different technologies were made. This approach will not limit technologies, but requires more detailed solar radiation calculations when the technology is chosen. On the other hand, the solar radiation output maps could be generated for each technology separately. This is possible with the ArcGIS tool Area solar radiation when the parameters (e.g., tilt angles and measured radiation type) are changed. To defend the choice of global (direct + diffuse + reflected) radiation use in this thesis, direct radiation often covers around 80 % of the global radiation (see Section 3.3.1 Area solar radiation tool), which keeps the ratio of the radiation magnitude consistent and seems to lead to the best spots in the final.

Further geographical information that could be added to the model are flatness of the areas by clustering the large-enough flat areas into a suitable area. It is easier to manage a solar power plant that has steadily changing topography. Before or after finding steady topography, it is reasonable to find hotspots in order to have a high magnitude of solar radiation. Determining the threshold of profitable radiation magnitude depends on the excepted proceedings and expenditures during the whole life cycle of the power plant. This question also leads to Herbert Simon's concept of "bounded rationality", which examines the difference between satisfying and optimal that was discussed in the theory section. Hotspots may be found by cluster or hotspot analysis, also available in ArcGIS software but in many other GIS software types as well.

The last idea for future development of the model is local issues. The local demand for electricity needs to be assessed, as well as the capacity and distribution of current transmission networks. As land ownership, local public policies and investment attractiveness are crucial for location optimization for utility-scale solar power plant, these are suggested inclusions for the model. If the model is considered useful as a part of a decision-making infrastructure, it might be worthwhile to build an automated GIS model to execute the pipeline of the model more easily.

5.6 Critics towards references

In this thesis, references were well-distributed in terms of assessing the topic and the emphasis of the thesis. The most emphasized references were GIS modeling and datasets, solar energy, site or land-use planning, as well as biodiversity and nature (Figure 50). Other energy subjects and agriculture were the next most popular, having five related papers on each subject. There were a couple of news items regarding the solar energy industry and a few more references concerning urbanization and economic growth, mainly in Africa.

There is a huge pro-solar movement going on and the distinct between objective and biased research is hard to solve. Overall, over 95 % of the references were papers found in the databases bought by the University of Helsinki. Those should be strictly peer-reviewed papers. The papers that handled the data used in the model are all highly valued and especially the nature protection data has strong background and it is published in Nature in 2015. Five books that handles the fundamentals of GIS are used. Two of those are advanced level literature in Master's studies. In Figure 51 can be seen the chronological distribution of the papers used.

GIS modeling and data:	IIIIIIIIIIIIIIIIIIIII
Solar energy:	IIIIIIIIIIIIIII
Biodiversity & nature:	IIIIIIIII
Site planning:	IIIIIIII
Land use planning:	IIIIII
Agriculture:	IIIII
Energy:	IIIII
Urbanization & economic growth:	IIII
Solar energy news:	II

Figure 50. Themes of the references and the amount of references regarding each theme.

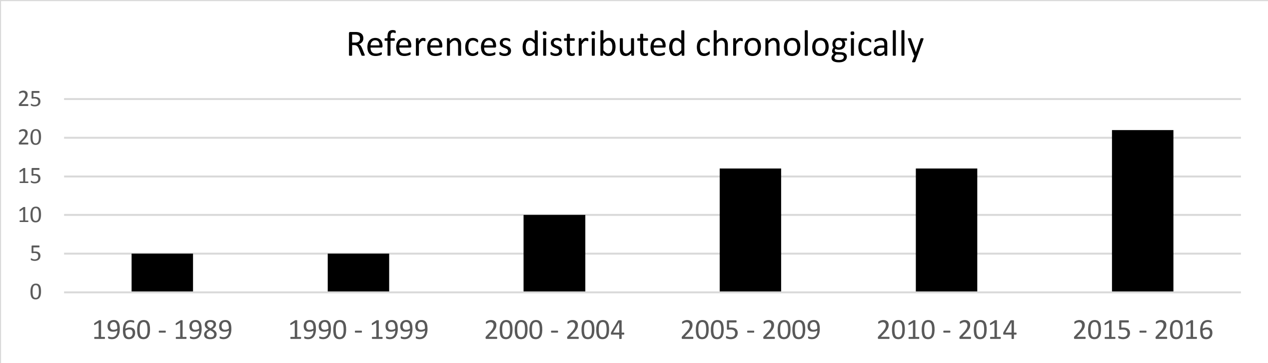


Figure 51. Most of the references were from the 21th century and around 40 out of 80 were published later than 2010.

To understand better electricity distribution, Chapter 3, "Modelling of energy service networks" from the book *Modelling: distributed energy resources in energy service network* is an excellent comprehensive overview of the subject. Another interesting book that was not utilized in this thesis due to time limitations was a book about advanced modeling regarding the economics of solar power. The name of the book is *Renewable and Efficient power systems,* and of particular interest is its excellent Chapter 5, "Economics of distributed resources". For example, the logic of balancing demand and supply electricity between different states in the U.S. of are thoroughly discussed in terms of the knowledge of distribution networks, which is one of the development targets of this thesis model.

6 CONCLUSIONS

The new model built in this thesis effectively restricts the unsuitable areas for utility-scale solar power plants in the greater city region. The dataset of land use intensiveness is the core of the efficient restriction. The output maps of the model show clearly which locations are prioritized for nature protection, but the validation areas of the model did not contain much of these areas. Spatial and temporal changes in solar radiation in a local environment indicate that the radiation estimates should be measured each month separately by using a digital elevation model as its basis. The model needs further development in order to work when applied to industry. The ideas of future development are included into the model as a sketch. The model is in a phase where it could be tested as an application of energy industry, so future development may continue.

Decision makers in e.g. energy companies have the possibility with the model focus their future investigations to the most interesting areas where to locate new solar power plants. The model cuts expenses when screening the optimal areas. It also gives tools for decision makers to carry out social responsibilities when important nature protection and land use pressure issues are added to the model.

REFERENCES

Alsamamra, H. et al., 2009. A comparative study of ordinary and residual kriging techniques for mapping global solar radiation over southern Spain. *Agricultural and Forest meteorology*, 149(8), pp.1343–1357.

van Asselen, S. & Verburg, P.H., 2012. A Land System representation for global assessments and land-use modeling. *Global Change Biology*, 18, pp.3125–3148.

van Asselen, S. & Verburg, P.H., 2013. Land cover change or land-use intensification: simulating land system change with a global-scale land change model. *Global Change Biology*, 19(12), pp.3648–3667. Available at: http://doi.wiley.com/10.1111/gcb.12331.

Bloomberg, 2015. What Apple Just Did in Solar Is a Really Big Deal. Available at: http://www.bloomberg.com/news/articles/2015-02-11/what-apple-just-did-in-solar-is-a-really-big-deal [Accessed December 4, 2015].

Boserup, E., 1965. The conditions of agricultural growth. *Population Studies*, 20(1), pp.1–108. Available at: http://www.jstor.org/stable/2172620?origin=crossref.

Brewer, J. et al., 2015. Using GIS analytics and social preference data to evaluate utility-scale solar power site suitability. *Renewable Energy*, 81, pp.825–836. Available at: http://www.sciencedirect.com/science/article/pii/S0960148115002979.

Breyer, C. et al., 2010. Fuel-parity: New very large and sustainable market segments for PV systems. In *2010 IEEE International Energy Conference and Exhibition, EnergyCon 2010*. pp. 406–411.

Breyer, C. & Gerlach, A., 2013. Global Overview on Grid-Parity Event Dynamics. *Progress in Photovoltaics: research and applications*, 21(1), pp.121–136.

Brook, B.W. & Bradshaw, C.J.A., 2015. Key role for nuclear energy in global biodiversity conservation. *Conservation Biology*, 29(3), pp.1–11.

Carson, L., 2014. Solar Energy. *Australian Energy Resource Assessment: Chapter 10 Solar Energy*, pp.261–284.

Chiabrando, R., Fabrizio, E. & Garnero, G., 2009. The territorial and landscape impacts of photovoltaic systems: Definition of impacts and assessment of the glare risk. *Renewable and Sustainable Energy Reviews*, 13(9), pp.2441–2451.

Convention on Biological Diversity, 2015. Convention on Biological Diversity. , p.CDB Homepage. Available at: https://www.cbd.int/information/parties.shtml [Accessed November 20, 2015].

Cormen, T.H. et al., 2009. *Introduction to algorithms* 3rd ed., Cambridge, MA: MIT Press. Available at: https://teemu-tomcat.linneanet.fi/vwebv/holdingsInfo?bibId=458367.

Cova, T. & Church, R., 2000. Contiguity constraints for single-region site search problems. *Geographical Analysis*, 32(4), pp.306–329.

DeFries, R. & Rosenzweig, C., 2010. Toward a whole-landscape approach for sustainable land use in the tropics. *Proceedings of the National Academy of Sciences of the United States of America*, 107(46), pp.19627–19632.

DeMers, M.N., 2002a. Characterizing the functional operations. In *GIS modeling in raster*. New Mexico: John Wiley & Sons, Inc., pp. 58–93.

DeMers, M.N., 2002b. Conceptualizing the model. In *GIS modeling in raster*. New Mexico: John Wiley & Sons, Inc., pp. 121–139.

DeMers, M.N., 2002c. *Model formulation, flowcharting, and implementation*, New Mexico: John Wiley & Sons, Inc.

DeMers, M.N., 2002d. Modeling Essentials. In *GIS modeling in raster*. Nex Mexico: John Wiley & Sons, Inc., pp. 94–120.

Deshpande, M.V., 2009. *Elements of Electrical Power Station Design*, PHI Learning.

Dubayah, R. & Rich, P., 1995. Topographic solar radiation models for GIS. *International Journal of Geographical Information Science*, 9(4), pp.405–419. Available at: http://www.informaworld.com/openurl?genre=article&doi=10.1080/02693799508902046&magic=crossref||D404A21C5BB053405B1A640AFFD44AE3.

Ellis, E.C. & Ramankutty, N., 2008. Putting people in the map: Anthropogenic biomes of the world. *Frontiers in Ecology and the Environment*, 6(8), pp.439–447.

Farfan, J. & Breyer, C., 2016. Structural Changes of the global power generation capacity towards sustainability and the risk of stranded investments, in preparation.

Fotheringham, A.S., Brundsdon, C. & Charlton, M., 2000. *Quantitative geography : perspectives on spatial data analysis* 1st ed., London: SAGE Publications Inc.

Fu, P. & Rich, P., 1999. Design and implementation of the Solar Analyst: an Arcview extension for modeling solar radiation at landscape scales. In *Proceedings of the Nineteenth Annual ESRI User Conference*. Kansas, pp. 1–33. Available at: http://professorpaul.com/publications/fu_rich_1999_esri.pdf.

Green, M.A. et al., 2015. Solar cell efficiency tables (Version 45). *Progress in Photovoltaics: Research and Applications*, 23(1), pp.1–9. Available at: http://dx.doi.org/10.1002/pip.1160.

Hansen, M.C., Stehman, S. V & Potapov, P. V, 2010. Reply to Wernick et al.: Global scale quantification of forest change. *Proceedings of the National Academy of Sciences of the United States of America*, 107(38), p.E148. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2944703&tool=pmcentrez&rendertype=abstract.

Harvey, D., 1985. *The Urbanization of Capital: Studies in the History and Theory of Capitalist Urbanization* 1st ed., The Johns Hopkins University Press.

Harvey, F., 2008. *A primer of GIS : fundamental geographic and cartographic concepts* 1. ed., New York: 2008 The Guilford Press. Available at: http://yjs.jxust.cn:801/gis/(S(ctn1zaqsp4zfsth5p0ghf30h))/uploads/20142303032322.pdf.

Harvey, F., 1997. From Geographic Holism to Geographic Information System. *The Professional Geographer*, 49(1), pp.77–85.

Hastings, A. et al., 2009. The development of MISCANFOR, a new Miscanthus crop growth model: towards more robust yield predictions under different climatic and soil conditions. *GCB Bioenergy*, 1(2), pp.154–170. Available at: http://doi.wiley.com/10.1111/j.1757-1707.2009.01007.x.

Hayers, A., 2015. The Economics Of Tesla Batteries. *Investopedia*. Available at: http://www.investopedia.com/articles/investing/050715/economics-tesla-batteries.asp.

Holstein, B.E., 2015. *Analyzing photovoltaic potential using a geographic information system: a case study of Prince William County public schools*. George Mason University. Available at: http://digilib.gmu.edu/xmlui/bitstream/handle/1920/9787/Holstein_thesis_2015.pdf?sequence=1&isAllowed=y.

Kemenes, A., Forsberg, B.R. & Melack, J.M., 2007. Methane release below a tropical hydroelectric dam. *Geophysical Research Letters*, 34(12), pp.1–5.

Lambie-Mumford, H. & Dowler, E., 2014. Rising use of “food aid” in the United Kingdom. *British Food Journal*, 116(9), pp.1418–1425. Available at: http://dx.doi.org/10.1108/BFJ-06-2014-0207.

Laurance, W.F. et al., 2012. Averting biodiversity collapse in tropical forest protected areas. *Nature*, 489(7415), pp.290–294.

Lehto, H. et al., 2007. *Fysiikka 8 : Aine ja säteily* 1st ed., Helsinki: Kustannusosakeyhtiö Tammi.

MacKay, R.. J.. & Oldford, R.W., 2000. Scientific Method , Statistical Method and the Speed of Light. *Statistical Science*, 15(3), pp.254–278. Available at: http://www.jstor.org/stable/2676665.

McKenney, D.W. et al., 2008. Spatial insolation models for photovoltaic energy in Canada. *Solar Energy*, 82(11), pp.1049–1061.

Mei, X., Fan, W. & Mao, X., 2015. Analysis of Impact of Terrain Factors on Landscape-scale Solar Radiation. *International Journal of Smart Home*, 9(10), pp.107–116. Available at: http://dx.doi.org/10.14257/ijsh.2015.9.10.12.

Montesino Pouzols, F. et al., 2014. Global protected area expansion is compromised by projected land-use and parochialism. *Nature*, 516(7531), pp.383–386. Available at: http://www.nature.com/doifinder/10.1038/nature14032.

NASA, 2015. Cloud Cover and Solar Radiation 82. , 990, pp.3–5. Available at: http://scool.larc.nasa.gov/lesson_plans/CloudCoverSolarRadiation.pdf [Accessed November 9, 2015].

NCAR & UCAR, 2016. CRU TS3.21 Gridded precipitation and other meteorological variables since 1901 - See more at: https://climatedataguide.ucar.edu/climate-data/cru-ts321-gridded-precipitation-and-other-meteorological-variables-1901#sthash.aE7nmntd.37iR48DB.dpuf. Available at: https://climatedataguide.ucar.edu/climate-data/cru-ts321-gridded-precipitation-and-other-meteorological-variables-1901 [Accessed January 14, 2016].

Pacione, M., 2001. *Urban geography : a global perspective*, London: Routledge. Available at: http://helka.linneanet.fi/cgi-bin/Pwebrecon.cgi?BBID=1560756.

Pogson, M., Hastings, A. & Smith, P., 2013. How does bioenergy compare with other land-based renewable energy sources globally? *GCB Bioenergy*, 5(5), pp.513–524. Available at: http://doi.wiley.com/10.1111/gcbb.12013.

Potts, D., 2015. Debates about African urbanisation, migration and economic growth: what can we learn from Zimbabwe and Zambia? *The Geographical Journal*, pp.1–14. Available at: http://doi.wiley.com/10.1111/geoj.12139.

Ramachandra, T. V & Shruthi, B. V, 2007. Spatial mapping of renewable energy potential. *Renewable and Sustainable Energy Reviews*, 11(7), pp.1460–1480.

Ramankutty, N. et al., 2008. Farming the planet 1. Geographic distribution of global agricultural lands in the year 2000. *Global Biogeochemical Cycles*, 22(1).

Rexer, M. & Hirt, C., 2014. Comparison of free high resolution digital elevation data sets (ASTER GDEM2, SRTM v2.1/v4.1) and validation against accurate heights from the Australian National Gravity Database. *Australian Journal of Earth Sciences*, 61(2), pp.213–226. Available at: http://www.tandfonline.com/doi/abs/10.1080/08120099.2014.884983.

Ruiz-arias, J.A. et al., 2009. A comparative analysis of DEM-based models to estimate the solar radiation in mountainous terrain. *International Journal of Geographical Information Science*, 23(8), pp.1049–1076.

Santangeli, A. et al., 2015. Global change synergies and trade-offs between renewable energy and biodiversity. *GCB Bioenergy*, pp.1–14. Available at: http://doi.wiley.com/10.1111/gcbb.12299.

Schmalensee, R. & Bulovic, V., 2015. *The future of solar energy : an interdisciplinary MIT study*, Available at: https://mitei.mit.edu/system/files/MIT Future of Solar Energy Study_compressed.pdf.

SEIA, 2015. Solar Industry Data, Solar Industry Breaks 20 GW Barrier - Grows 34% Over 2013. Available at: http://www.seia.org/research-resources/solar-industry-data [Accessed November 23, 2015].

Shahzad, S., Renguang, W. & Iftikhar, A., 2014. Impact of Terrain and Cloud Cover on the Distribution of Incoming Direct Solar Radiation over Pakistan. *Journal of Geographic Information System*, 2014(6), pp.70–77. Available at: http://dx.doi.org/10.4236/jgis.2014.61008.

Shodor, 2016. Solar Radiation Cloud Cover Adjustment Calculator. Available at: http://www.shodor.org/os411/courses/_master/tools/calculators/solarrad/ [Accessed January 13, 2016].

de Smith, M.J., Goodchild, M.F. & Longley, P.A., 2009. *Geospatial analysis : a comprehensive guide to principles, techniques and software tools* 5th ed., Leicester: Matador cop. Available at: http://www.spatialanalysisonline.com/HTML/.

Solar Energy Industries Association, 2015. Major Solar Projects Map. Available at: http://www.seia.org/map/majorprojectsmap.php [Accessed December 9, 2015].

Šúri, M. & Hofierka, J., 2004. A new GIS-based solar radiation model and its application to photovoltaic assessments. *Transactions in GIS*, 8(2), pp.175–190.

The Economist, 2015. African energy: The leapfrog continent. *The Economist*. Available at: http://www.economist.com/news/middle-east-and-africa/21653618-falling-cost-renewable-energy-may-allow-africa-bypass.

Tomlin, C.D., 1990. Chapter 4: Characterizing Individual Locations. In *Geographic Information Systems and Cartographic Modeling*. Redlands, CA, pp. 45–65.

Tovar-Pescador, J. et al., 2006. On the use of the digital elevation model to estimate the solar radiation in areas of complex topography. *Meteorological applications*, 13(03), p.279. Available at: http://doi.wiley.com/10.1017/S1350482706002258\nhttp://onlinelibrary.wiley.com/doi/10.1017/S1350482706002258/abstract.

Tsoutsos, T., Frantzeskaki, N. & Gekas, V., 2005. Environmental impacts from the solar energy technologies. *Energy Policy*, 33, pp.289–296. Available at: http://www.sciencedirect.com/science?_ob=MImg&_imagekey=B6V2W-49JHJV2-2-1&_cdi=5713&_user=1922361&_pii=S0301421503002416&_origin=search&_coverDate=02/28/2005&_sk=999669996&view=c&wchp=dGLzVlb-zSkzS&md5=6eedf1ba4afd738926a997579796c78d&ie=/sdarticle.pdf.

Wang, J., Robinson, G.J. & White, K., 1996. A fast solution to local viewshed computation using grid-based digital elevation models. *Photogrammetric Engineering & Remote Sensing*, 62(10), pp.1157–1164.

Watson, J.J.W. & Hudson, M.D., 2015. Regional Scale wind farm and solar farm suitability assessment using GIS-assisted multi-criteria evaluation. *Landscape and Urban Planning*, 138, pp.20–31. Available at: http://www.sciencedirect.com/science/article/pii/S0169204615000249.

Williams, T.H.L., 1985. Implementing LESA on a geographic information system: A case study. *Photogrammetric Engineering and Remote Sensing*, 51(12), pp.1923–1932.

World Bank, 2015. World bank: Access to electricity (% of population). Available at: http://data.worldbank.org/indicator/EG.ELC.ACCS.ZS?order=wbapi_data_value_2012+wbapi_data_value+wbapi_data_value-first&sort=asc [Accessed November 22, 2015].