SPATIAL MODELLING OF RENEWABLE ENERGY
INTEGRATING REMOTE SENSING DATA

Thesis submitted in partial fulfillment of the requirements of
the degree Doctor rer. nat. of the
Faculty of Forest and Environmental Sciences,
Albert-Ludwigs-Universität
Freiburg im Breisgau, Germany

By

Shifeng Wang

Freiburg im Breisgau, Germany
2010
Acknowledgements

I sincerely express my gratitude to Prof. Dr. Barbara Koch for accepting me as a doctoral student at the Department of Remote Sensing and Landscape Information Systems (FELIS). She introduced me to a new field of renewable energy, helped me to completely “enter" the field and encouraged me to think critically and analytically. Without her invaluable supervision, support and contribution, I would not have obtained such interesting results in my research. I am indebted to her.

I delightedly acknowledge another supervisor Prof. Dr. Dr. h.c. Dieter R. Pelz from the Department of Forest Biometry. He gave me a lot of wonderful suggestions, comments and criticisms for my research. Thank him for his help.

Thanks to my colleagues at FELIS, especially Dr. Matthias Dees and Mr. Christian Schill, for their friendship, patience and support and the enlightening discussions on the topics of remote sensing, GIS, mathematical, renewable energy and others. They have left an unforgettable impression on me. Thanks to Christian Schill for the translation of the German summary.

I am deeply grateful to the forest program at the International Institute for Applied Systems Analysis (IIASA), especially Dr. Michael Obersteiner, Dr. Fritz Steffen, Dr. Sylvain Leduc and Mr. Kraxner Florian, etc. There are so many people who helped me when I stayed in Vienna, Austria and still helped me after I left Vienna that I can not list them all. I am indebted to them.

I would also like to express my honest appreciation to my parents and my brother. Thank you for supporting me with my studies in Germany, always encouraging me and pushing me forward. Thank my brother for taking good care of my parents, supporting me and making me concentrate on my research in Germany. A special thanks to Dr. Sicong Wang for supporting me and taking good care of me.

Thank Marylou Khalfa (Swansea University, U.K.) for help in editing English grammar and style.

The study is funded by GEOBENE (Global Earth Observation - Benefit Estimation: Now, Next and Emerging) project.

Shifeng Wang (汪诗锋)
28, June 2010, Freiburg, Germany
# CONTENTS

LIST OF TABLES ........................................................................................................ iv
LIST OF FIGURES ..................................................................................................... v
ABBREVIATIONS ..................................................................................................... vi
ABSTRACT ............................................................................................................... vii
ZUSAMMENFASSUNG ........................................................................................... ix

1 INTRODUCTION ................................................................................................. 1

1.1 RESEARCH BACKGROUND ........................................................................... 1
1.2 RESEARCH SIGNIFICANCE .......................................................................... 3
  1.2.1 Energy and population ............................................................................. 3
  1.2.2 Energy and climate .................................................................................. 4
  1.2.3 Chance of renewable energy ................................................................. 4
  1.2.4 Characteristics of renewable energy ....................................................... 6
  1.2.5 Development of renewable energy ........................................................... 6
1.3 RESEARCH MOTIVATION ............................................................................... 8
1.4 RESEARCH QUESTIONS AND OBJECTIVES .............................................. 9
1.5 ORGANIZATION OF THE DISSERTATION .................................................. 11

2 LITERATURE REVIEW ...................................................................................... 12

2.1 OVERVIEW OF MODELING RENEWABLE ENERGY .................................... 12
2.2 MATHEMATICAL APPROACHES OF RENEWABLE ENERGY OPTIMIZATION MODELS... 14
2.3 ENERGY SCENARIO ...................................................................................... 15
2.4 EQUILIBRIUM THEORY .............................................................................. 16
2.5 REFERENCE ENERGY SYSTEM .................................................................... 18
2.6 SPATIAL WEIGHT ......................................................................................... 20
2.7 SPATIAL AUTOCORRELATION ..................................................................... 21
2.8 SPATIAL MODEL SPECIFICATION TEST ..................................................... 24
2.9 SPATIAL TREND MODEL ............................................................................. 26
2.10 STATISTICAL MODEL SELECTION CRITERIONS ....................................... 28
2.11 OPTIMIZATION MODELS FOR SUPPLY .................................................. 30
2.12 STATISTICAL MODELS FOR CONSUMPTION ......................................... 34

3 DESIGN OF OFLR OPTIMIZATION MODEL FOR SUPPLY .......................... 37

3.1 MODEL OBJECTIVES AND SCOPE ................................................................ 37
3.2 SOLAR AND WIND ENERGY POTENTIALS WITH REMOTE SENSING DATA ...... 38
3.3 SIMPLIFICATION OF SUPPLY CHAIN ......................................................... 41
3.4 REGULAR GRID OF SUPPLY AND DEMAND REGIONS ............................... 44
3.5 SPATIALLY EXPLICIT .................................................................................... 45
3.6 Flexibility ......................................................................................................................... 46

4 DESCRIPTION OF OFLR MODEL ........................................................................... 48

4.1 Economic theory .......................................................................................................... 48
4.2 Symbol system ............................................................................................................. 49
4.3 Subscripts and sets ..................................................................................................... 51
4.4 Variables description ................................................................................................. 51
4.5 Parameters description .............................................................................................. 52
4.6 Constraint equations .................................................................................................. 53
  4.6.1 Economics ............................................................................................................... 53
  4.6.2 Renewable energy potential ................................................................................... 53
  4.6.3 Renewable energy plants ....................................................................................... 54
  4.6.4 Transportation plants ............................................................................................ 56
  4.6.5 Demand regions ..................................................................................................... 58
  4.6.6 Remote sensing-based constraints ......................................................................... 58
4.7 Objective function ....................................................................................................... 60
4.8 Diagram ....................................................................................................................... 62

5 DESCRIPTION OF IST MODEL FOR CONSUMPTION .............. 64

5.1 IST model .................................................................................................................... 64
5.2 Spatial statistical model selection procedure ............................................................. 65

6 CASE STUDIES .............................................................................................................. 67

6.1 Case study of OFLR model ........................................................................................ 67
  6.1.1 Materials ............................................................................................................... 67
  6.1.2 Experimental design .............................................................................................. 73
  6.1.3 Results .................................................................................................................... 76
6.2 Case study of IST model ............................................................................................. 88
  6.2.1 Materials ............................................................................................................... 88
  6.2.2 Experimental workflow ......................................................................................... 92
  6.2.3 Results .................................................................................................................... 94

7 DISCUSSION .................................................................................................................... 99

7.1 Discussion of OFLR model ....................................................................................... 99
7.2 Discussion of IST model .............................................................................................. 102
7.3 Discussion of case study of OFLR model ................................................................. 104
7.4 Discussion of case study of IST model ...................................................................... 107

8 CONCLUSIONS AND FUTURE WORK .................................................................... 112

8.1 Conclusions ................................................................................................................ 112
8.2 Future work ........................................................................................................... 114

REFERENCES............................................................................................................. 115

APPENDIX I: CODES OF 49 U.S. STATES AND RELATED NAME ............ 125
List of Tables
Table 1-1. Energy consumption in 2004........................................................................................................... 4
Table 1-2. Selected renewable energy indicators ............................................................................................ 7
Table 2-1. Comparison of Moran’s I and Geary’s C......................................................................................... 23
Table 3-1. Cost and performance parameters of HVAC and HVDC with transmission of 5GW... 43
Table 3-2. Countries adopting HVDC ............................................................................................................ 43
Table 4-1. Suitable factors of land categories .................................................................................................. 54
Table 4-2. Reference cost of solar thermal plant with 200MW ................................................................. 56
Table 4-3. HVDC with voltage ±800kV, unit capacity 5GW .......................................................................... 57
Table 6-1. Installed capacity of photovoltaics in 2003..................................................................................... 69
Table 6-2. Data list of case study of OFLR model .......................................................................................... 70
Table 6-3. Suitable factors of land cover ......................................................................................................... 72
Table 6-4. Photovoltaics prices of five European countries, 2006 ................................................................. 73
Table 6-5. Electricity demands of five European countries, 2006 ................................................................. 73
Table 6-6. Locations of demand regions and corresponding electricity demands ........................................ 79
Table 6-7. Potential photovoltaics locations and solar potential ................................................................. 80
Table 6-8. Optimal locations and capacity sizes of photovoltaics, Scenarios I and II ................................. 82
Table 6-9. Optimal capacity sizes of HVDC, Scenarios I and II ................................................................. 82
Table 6-10. Electricity prices of demand regions, Scenarios I and II ........................................................... 84
Table 6-11. Optimal locations and capacity sizes of photovoltaics, Scenario III ........................................ 84
Table 6-12. Optimal capacity sizes of HVDC, Scenario III .......................................................................... 84
Table 6-13. Electricity prices of demand regions, Scenario III .................................................................... 85
Table 6-14. Optimal locations and capacity sizes of photovoltaics, Scenario IV ....................................... 86
Table 6-15. Optimal capacity size of HVDC, Scenario IV ......................................................................... 86
Table 6-16. Electricity prices of demand regions, Scenario IV .................................................................... 87
Table 6-17. Sources of datasets of case study of IST model ................................................................. 89
Table 6-18. Estimated values of parameters of spatial trend models ......................................................... 94
Table 6-19. Estimated values of parameters of IST models ......................................................................... 96
Table 6-20. Moran’s I test for residuals of spatial trend models and IST models ........................................ 97
Table 7-1. Theoretical solar potential, geographical solar potential and improvement ................................. 105
Table 7-2. Estimated values of parameters of single spatial trend models ................................................. 108
Table 7-3. Estimated value of parameters of single IST models ............................................................... 109
Table 7-4. Moran’s I test for residuals of all single spatial trend models and single IST models ............ 109
List of Figures

Fig 1-1: Selected renewable energies and their renewable products ........................................ 1
Fig 1-2: Renewable energy share in final global energy consumption in 2006 .......................... 5
Fig 2-1: Equilibrium theory ....................................................................................................... 17
Fig 2-2: Simple regional RES ..................................................................................................... 19
Fig 3-1: World 22-year (July 1983 - June 2005) annual average solar direct normal irradiation... 39
Fig 3-2: Wind speed map of Danish seas (Denmark), retrieved from SAR image on 22 October 2004 ......................................................................................................................... 40
Fig 3-3: Reference energy system of supply chain system of solar and wind energy ............... 41
Fig 3-4: Grids of Germany with spatial resolution of 0.5×0.5 degrees ....................................... 45
Fig 4-1: Equilibrium theory with fixed demand ......................................................................... 48
Fig 4-2: Operation procedure of remote sensing-based constraints ........................................ 59
Fig 4-3: Diagram of the OFLR model ...................................................................................... 63
Fig 5-1: Spatial statistical model selection procedure ............................................................... 66
Fig 6-1: Spatial layout of five studied countries ................................................................------- 69
Fig 6-2: Spatial distribution of solar direct normal irradiation of five European countries ...... 70
Fig 6-3: Landuse of five studied countries ................................................................................ 71
Fig 6-4: Land slope of five studied countries ............................................................................ 72
Fig 6-5: Electricity supply chain of solar energy with photovoltaics ........................................ 74
Fig 6-6: Workflow of case study of OFLR model ...................................................................... 75
Fig 6-7: Studied area grids with spatial resolution of 0.5×0.5 degrees .................................... 76
Fig 6-8: Solar direct normal irradiation with spatial resolution of 0.5×0.5 degrees of studied area ................................................................................................................................. 77
Fig 6-9: Land slope of grids of studied area .............................................................................. 78
Fig 6-10: Potential photovoltaics locations and demand regions ........................................... 80
Fig 6-11: Optimal locations of photovoltaics and HVDC, Scenarios I and II ......................... 83
Fig 6-12: Optimal locations of photovoltaics and HVDC, Scenario III ..................................... 85
Fig 6-13: Optimal locations of photovoltaics and HVDC, Scenario IV .................................... 87
Fig 6-14: Spatial layout of 49 U.S. states .................................................................................... 88
Fig 6-15: Spatial distribution of biomass consumption, U.S. .................................................. 89
Fig 6-16: Spatial distribution of biomass price, U.S. ................................................................. 90
Fig 6-17: Spatial distribution of total disposable personal income, U.S. ................................. 90
Fig 6-18: Spatial locations of 49 U.S. states ............................................................................. 91
Fig 6-19: Workflow of case study of IST model ...................................................................... 93
Abbreviations

AIC  Akaike Information Criterion
BIGCC  Biomass Integrated Gasification Combined-Cycle
CO₂  Carbon dioxide
DLR  German Aerospace Center
DNI  Direct Normal Irradiation
DP  Dynamic Programming
EIA  Energy Information Administration
ESA  European Space Agency
EU  European Union
GHG  GreenHouse Gases
GIS  Geographic Information System
HVAC  High Voltage Alternating Current Transmission
HVDC  High Voltage Direct Current transmission
IEA  International Energy Agency
IPCC  Intergovernmental Panel on Climate Change
JREC  Johannesburg Renewable Energy Coalition
LM  Lagrange Multiplier
LP  Linear Programming
LR  Likelihood Ratio
MBR  Minimal Bounded Rectangular
MESSAGE  Model for Energy Supply Strategy Alternatives and their General Environmental Impact
MIP  Mixed Integer Programming
ML  Maximum Likelihood
MLP  Mixed integer Linear Programming
MODIS  Moderate Resolution Imaging Spectroradiometer
NASA  National Aeronautics and Space Administration
OECD  Organization for Economic Co-operation and Development
OFLR model  Optimization Facility Location and capacity size model for solar energy facilities and wind energy facilities in electricity supply chain system integrating Remote sensing data
OLS  Ordinary Least Square
RES  Reference Energy System
SAR  Synthetic Aperture Radar
SERS  Special Report on Emission Scenarios
SIC  Schwarz Information Criterion
TIMES model  an acronym for The Integrated MARKAL-EFOM System
Abstract

The location and capacity size of the renewable energy facility and the influencing factors of renewable energy consumption play an important role in planning renewable energy. Therefore, the solutions to determining the optimal locations and optimal capacity sizes of renewable energy facilities in renewable energy supply chain and identifying the spatial influencing factors of renewable energy consumption are essential in making scientific renewable energy planning strategy with optimal cost. The main focus of the study is to develop two renewable energy models to provide such solutions. The study also combines remotely-sensed data into the models. The study is composed of two main stages: the development of an optimization supply model integrating spatial information from remote sensing and a consumption model also integrating spatial information. Both are needed for optimizing the energy market but each of them has first to be developed individually.

In the first stage, the study developed an optimization supply model with remotely-sensed data to determine the optimal locations and optimal capacity sizes of renewable energy facilities in renewable energy supply chain. In order to develop this model, the first step was to restrict the studied renewable energy supply chain to electricity supply chain of solar and wind energy and then to simplify the supply chain. In the second step, the solar and wind energy potentials were evaluated with remote sensing data. A grid structure was established to integrate remote sensing data and remote sensing-based constraints. The spatial distributions of supply and demand were modeled taking into account remote sensing data. The by-products and pre-determined capacity sizes for solar and wind energy facilities were also developed and can be taken into account if needed. Then the equilibrium theory with fixed demand was used to formulate the objective function of the model such that the model will not be used to optimize the profit of one single plant, but rather to maximize the social welfare of the demands. The model was called Optimization Facility Location and capacity size model for solar energy facilities and wind energy facilities in electricity supply chain system integrating Remote sensing data (OFLR model). It was a mixed integer program. As renewable energy planning is important, the model is a valuable tool for decision makers in order to make spatially explicit location planning strategies for solar and/or wind energy.

In the second stage, the study aimed to develop a statistical model to identify the spatial
influencing factors of renewable energy consumption. Initially, the considered influencing factor of renewable energy consumption was primarily the locations of consumptions and therefore the spatial trend model was used to identify location factor. The spatial trend model did not meet the requirements and an Improved Spatial Trend model (IST model) was developed to further identify location factor and non-location spatial influencing factors of renewable energy consumption and solve the ill-conditioned least squares problem occurred in spatial trend model. Finally, a spatial statistical model selection procedure was provided to select the best IST model for renewable energy consumption, regarding the spatial autocorrelation among residuals.

The OFLR model was developed with GAMS programming language and the IST model was developed with Matlab programming language. These two models were illustrated by case studies.
Zusammenfassung


Das OFLR-Modell ist in der Programmiersprache GAMS entwickelt, das IST-Modell in Matlab. Die zwei Modelle werden durch Fallstudien illustriert.
1 Introduction

Energy, global warming and growing population are three major global issues of the 21st century. The growing population needs more energy to support its standard of living. The consumption of conventional energy, especially fossil fuel, is increasing the emission of GreenHouse Gases (GHG) into the atmosphere which will increase the temperature of the earth. The temperature increase of the earth may threaten humankind. As a result, the issue of energy has become one of the most important issues, especially in implementing sustainable development.

1.1 Research background

Renewable energy refers to the energy generated from natural resources such as sunlight, wind, forest and geothermal heat, which are renewable. Renewable energy cannot be used directly because it is not the end-use energy which can be used directly by residents, such as light, heat and gas, etc. Fig 1-1 illustrates the relationship between selected renewable energies and their renewable products.

The Intergovernmental Panel on Climate Change (IPCC) has developed a lot of future scenarios for sustainable development. Among these scenarios, the so-called Special Report on Emission Scenarios (SERS) stressed the importance of renewable energy and
expected a significant contribution of renewable energy (Nakicenovic 2000; Hoogwijk 2004). Later, in 2006, IPCC held the 25th plenary meeting in Mauritius. In this meeting, a decision was taken to develop a special report on renewable energy sources and climate change mitigation that acknowledged the importance of the issues related to the use of renewable energy sources. The structure of the special report was scoped at an IPCC meeting on renewable energy sources held at Lübeck, Germany, from January 21st to 25th 2008 by 120 leading world experts who provided a possible structure of the special report and an assessment of the availability of published scientific literature on the topic. The special report stated that emission trading schemes, energy efficiency and the rapid development of renewable energy were needed to combat the world's looming environmental problems (IPCC 2008). The meeting and the special report highlighted the importance of developing research towards renewable energy.

The Organization for Economic Co-operation and Development (OECD) pointed out that the exploitation of renewable energy resources and the technologies played a key role in sustainable development (OECD 1995). The sub department of OECD, the International Energy Agency (IEA), developed an alternative policy scenario in which the share of renewables in global energy consumption by 2030 will remain largely at 14% (IEA 2006). In order to support renewable energy research, IEA developed a global renewable energy policies and measures database which provided information on the renewable energy policies and the measures taken or planned to encourage the uptake of renewable energy. The database covered the measures in IEA member countries, together with the members of the Johannesburg Renewable Energy Coalition (JREC), and Brazil, China, the European Union, India, Mexico, Russia and South Africa. Comprising more than 1,000 records dating back to 2000 and some even earlier, the database provided an excellent source of information on renewable energy policy developments for decision makers, policy experts and researchers, and provided the practical information on renewable energy to the business community and the broader public (IEA database 2008).

The European Union (EU) has released a lot of plans on renewable energy. The EU has been working towards a renewable energy supply equivalent to 12% of the European Union's total energy consumption by 2010. The European council has proposed that 15% of total energy consumption should be produced from renewable resources by 2015. In the European conference for renewable energy held in Berlin in 2004, the EU defined the
ambitious goals of its own that the EU would seek to obtain 20% of its total energy consumption requirements with renewable energy sources in 2020. In 2006, the European Parliament called for 25% target for renewable energy in European Union's energy consumption by 2020. In order to implement this target, the European council presented the renewable energy roadmap (EU 2007).

In order to motivate research on renewable energy, the European Union has launched a lot of renewable energy projects. For example, under the sixth framework program, the EU launched the European Sustainable Electricity: Comprehensive analysis of Future European Demand and Generation of European Electricity and its Security of Supply (EUSUSTEL) project (Contract No.: 006602). The aims of this project were to provide a fully consistent framework for a secure electricity provision which was at the same time environmentally friendly and affordable. The project investigated the possibility of generating electricity from renewable energy and concluded that renewable energy was a promising energy to generate electricity for 25 EU countries (EUSUSTEL 2009). Under the scientific support to policies priority of the sixth RTD framework program, the EU partially funded the CAse Study Comparisons And Development of Energy Models for INtegrated Technology Systems (CASCADE MINTS) project (Contract No.: 502445). The CASCADE MINTS project can be split into two distinct parts: Part 1 focused on modeling, scenario evaluation and detailed analysis of the prospects of the hydrogen economy. The ultimate aim of this part was to enable perspective analysis of the conditions under which a transition to an energy system dominated by hydrogen was possible. Part 2 was to use a wide range of existing operational energy and energy/economy models in order to build analytical consensus concerning the impacts of policies aimed at sustainable energy systems. In this project, a lot of attention concentrated on the renewable energy research. The project concluded that renewable energy was expected to be a robust way of addressing climate issues by decreasing the share of fossil fuels in Europe’s energy mix (CASCADE MINTS 2008).

1.2 Research significance

1.2.1 Energy and population

Energy consumption has a close relation with population. The total population of the world is expected to reach 8.05 billions in 2030, 8.75 billions in 2050, 8.87 billions in
2060, 8.62 billions in 2090 and 8.39 billions in 2100 (Lutz et al. 2008). The large increasing population will need more energy if consumption stays the same. Compared with that in 1950-1990, the world primary energy consumption in 2000-2050 is expected to more than double (EU 2006). However, the world final energy growth rate is only 1.4%/year\(^1\) for 2000-2030. As a consequence, the fast growing population and the slow energy growth rate shape a conflict. The conflict suggests that we must control population and look for new energy sources. Otherwise, an energy crisis will occur and the stability of the world will be endangered.

### 1.2.2 Energy and climate

Energy consumption, especially conventional energy consumption, has an intimate relationship with climate issues. When we burn fossil fuels such as oil, coal and gas, etc., to make electricity, heat our buildings, and power our transportation tools, greenhouse gases, especially the carbon dioxide (CO\(_2\)), are released. These greenhouse gases warm the earth and influence the climate. Climate plays a dominant role in determining where and how people live. Climate change impacts our lives and is expected to destroy many natural environments in the coming years. A rise of 2°C minimum of the global mean temperature above pre-industrial levels will damage the ecosystems and make people at risk. Consequently, protecting the climate is extremely important for people’s lives.

### 1.2.3 Chance of renewable energy

Nowadays, the main end-use energy is still conventional energy (e.g. fossil energy). According to the U.S. Energy Information Administration's 2006 estimate (EIA 2006), the estimated 15TW total energy consumption in 2004 was divided as Table 1-1, among which fossil fuels supplied 86% of the world’s energy.

<table>
<thead>
<tr>
<th>Energy source</th>
<th>Energy power (TW)</th>
<th>Energy consumption (EJ)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Oil</td>
<td>5.6</td>
<td>180</td>
</tr>
<tr>
<td>Gas</td>
<td>3.5</td>
<td>110</td>
</tr>
<tr>
<td>Coal</td>
<td>3.8</td>
<td>120</td>
</tr>
<tr>
<td>Hydroelectric</td>
<td>0.9</td>
<td>30</td>
</tr>
<tr>
<td>Nuclear</td>
<td>0.9</td>
<td>30</td>
</tr>
<tr>
<td>Geothermal, wind, solar, wood</td>
<td>0.13</td>
<td>4</td>
</tr>
</tbody>
</table>

\(^1\) The figure is calculated based on IIASA Low Scenario (Häfele (1981) and Häfele and Rogner (1984)).
The share of renewable energy in total energy consumption is still small. Fig 1-2 illustrates the share of renewable energy in final global energy consumption in 2006. In terms of Fig 1-2, the main source of energy consumed is still fossil fuels, representing 79%, whereas renewable energy accounts for only 18% in which the traditional biomass owns 13%. Since the consumption of conventional energy will give off greenhouse gases which will change the climate and thus hit sustainable development, the large share of conventional energy in the final total energy consumption will cause big damage to the atmosphere. As a consequence, we must find a new way to effectively utilize conventional energy or look for new clean and green energy sources. Moreover, for conventional energy the recycle period is considerably long. For example, the recycle period of oil needs million years. The long recycle period makes it difficult for conventional energy to be reused in relatively short time. Furthermore, conventional energy is finite and can be exhausted in the coming years if continually consumed.

![Figure 1-2: Renewable energy share in final global energy consumption in 2006 (REN21 2008).](image)

Renewable energy holds a bright promise to complement conventional energy. The potential of renewable energy is extremely huge. According to BP statistics (BP 2002), the global technical potential of solar energy is about 23 times of current global electricity consumption. Meanwhile, the advantages of renewable energy are numerous which will be discussed in the following section. The contribution of renewable energy to the total energy consumption is, however, very small. As a result, increasing the contribution of renewable energy to the total energy consumption is necessary and important. It is an effective solution to implement sustainable development.
1.2.4 Characteristics of renewable energy

The advantages of renewable energy are numerous. The following text summarizes some of them:

1). Renewable energy, like other energy, can be satisfied with the demand of energy. For example, biomass and solar energy can be converted into electricity for residents.

2). It will contribute to the diversity of the energy supply portfolio and reduce risks of continued/expanded use of primary conventional energy. For example, electricity can be generated from solar energy through photovoltaics rather than from fossil fuels.

3). It is clean and green energy source. Renewable energy is an environmentally friend and healthy energy. Its consumption will give off little greenhouse gases. For example, the estimated global carbon emissions from biomass burning vary from about one fifth to one third of the carbon released from fossil fuel combustion (Soares Neto et al. 2009) and the CO2 released by biomass combustion is equivalent to about 12% of the CO2 emitted by fossil fuel use (Barker et al. 2007). The Biomass Integrated Gasification Combined-Cycle (BIGCC) generating plants can significantly reduce particulate emissions by a factor of 4.5 in comparison with coal-based electricity generation processes. NOx emissions can be reduced by a factor of about 6 for dedicated BIGCC plants compared with average pulverized coal-fired plants (Mann and Spath 2002).

4). Its potential is very huge. The global technical potential of solar energy is about 23 times of the current global electricity consumption (BP 2002).

5). Finally, it aids economic development, including employment and investment opportunities. For example, in 2006, around 235,000 people in Germany were employed in the renewable energy sector, especially in small and medium sized companies. Over half of these jobs were attributed to the renewable energy sources act (BMU 2007).

1.2.5 Development of renewable energy

The consumption of renewable energy is growing rapidly. In 2006, about 18% of global final energy consumption came from renewable energy and the ratios were increasing rapidly (REN21 2007). The share of renewable energy in electricity generation was around 18%, with 15% of global electricity coming from hydroelectricity. From the end of 2004 to the end of 2008, solar photovoltaics capacity increased sixfold to more than 16 gigawatts (GW), wind power capacity increased 250% to 121 GW, and total power capacity from new renewable energy increased 75% to 280 GW. During the same period,
the solar heating capacity doubled to 145 gigawatts-thermal (GWth), while biodiesel production increased sixfold to 12 billion liters per year and ethanol production doubled to 67 billion liters per year (REN21 2009). Table 1-2 illustrates the selected renewable energy indicators. On the basis of Table 1-2, investment in new renewable capacity is growing annually, from 63 billion US dollars in 2006 to 120 billion US dollars in 2008. It reflects that renewable energy is receiving more and more attention. Among all renewable energies, the development of biomass heating is the fastest, reaching around 250 GWth in 2008. We also know from Table 1-2 that renewable energy policy is being highlighted on a yearly basis. The number of countries with renewable energy policy increased from 66 in 2007 to 73 in 2008, reflecting that more and more countries realize the importance of renewable energy policy. Renewable energy policy plays a greatly important role in renewable energy system. Scientific ambitions renewable energy policy can motivate the development of renewable energy and thus can solve a variety of energy issues caused by conventional energy. Bad and ambiguous renewable energy policy will, however, hinder the development of renewable energy. As a consequence, we need to carefully create reasonable and scientific renewable energy policy.

Table 1-2. Selected renewable energy indicators (REN21 2009; Martinot and Sawin 2009)

<table>
<thead>
<tr>
<th>Selected global indicator</th>
<th>2006</th>
<th>2007</th>
<th>2008</th>
</tr>
</thead>
<tbody>
<tr>
<td>Geothermal heating</td>
<td>~50 GWth</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Biomass heating</td>
<td>~250 GWth</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Solar hot water/ Space heating</td>
<td>145 GWth</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Countries with policy targets for renewable energy use</td>
<td>66</td>
<td>73</td>
<td></td>
</tr>
<tr>
<td>Ethanol production (annual)</td>
<td>39</td>
<td>50</td>
<td>67 billion liters</td>
</tr>
<tr>
<td>Investment in new renewable capacity (annual)</td>
<td>63</td>
<td>104</td>
<td>120 billion USD</td>
</tr>
<tr>
<td>Wind power capacity (existing)</td>
<td>74</td>
<td>94</td>
<td>121 GWe</td>
</tr>
<tr>
<td>Existing Renewables power capacity, excluding large-scale hydro</td>
<td>207</td>
<td>240</td>
<td>280 GWe</td>
</tr>
<tr>
<td>Existing Renewables power capacity, including large-scale hydro</td>
<td>1,020</td>
<td>1,070</td>
<td>1,140 GWe</td>
</tr>
</tbody>
</table>

USD is the US dollar; GWe is Gigawatts-electric.

Although the development of renewable energy is rapid, the ratio of renewable energy in global energy consumption is still small, accounting for only 18% of the total energy consumption (REN21 2008). Consequently, we need to devise reasonable and scientific renewable energy policy to increase the share of renewable energy in the total energy system and thus to implement sustainable development. The renewable energy policy will guide and motivate the investment to renewable energy such that the share of renewable
energy will increase in the coming years. Modeling renewable energy is one of the good sources which can provide valuable suggestions to renewable energy policy. As a result, we should develop renewable energy models.

1.3 Research motivation

The characteristics of renewable energy discussed above and sustainable development as well as the shortcomings of primary conventional energy (refer to section 1.2.3 chance of renewable energy) motivate us utilize renewable energy. It is very urgent. In order to utilize renewable energy, we should first supply renewable energy to energy consumers. Therefore, it is worth investigating that how to economically supply renewable energy. Many attempts have been made to date to simulate and optimize renewable energy supply chain (Rentizelas et al. 2009). In renewable energy supply chain, the first importance is the determination of the optimal locations and optimal capacity sizes of renewable energy facilities because they directly determine the whole supply chain and its economic profits (Rentizelas et al. 2009). It is therefore worth investigating how to determine the optimal locations and optimal capacity sizes of renewable energy facilities in renewable energy supply chain (Wang and Koch 2010).

Renewable energy supply is directly related with renewable energy consumption. Without consumption, it makes no sense to supply renewable energy. Different consumptions will need different supplies. Even for the same consumptions, if consumptions are situated in different locations, supplies will be changed accordingly. Therefore, to create an optimal long-term supply strategy, we should also simultaneously investigate the spatial variations of renewable energy consumption and identify the spatial influencing factors of renewable energy consumption. When identifying the spatial influencing factors of renewable energy consumptions, the first factor we should consider is the locations of consumptions. In addition to locations, we should also identify any other spatial influencing factors of renewable energy consumption. Once we know the spatial influencing factors of renewable energy consumption, we can effectively manage consumption and generate the consumption data for supply chain to create an optimal supply strategy. In this sense, it is worth investigating the spatial influencing factors of renewable energy consumption, especially the location factor.

On the other hand, renewable energy consumption will be also influenced by renewable energy supply. The more supply available, the more consumption there will be. As a result,
renewable energy supply and consumption should be investigated together (Belyaev et al. 1976; Basile 1980; LEAP 2002).

Data is a vitally important component of renewable energy models. Without data, the renewable energy model can not work, just like fish without water. As before, it was hard to obtain large scale data such as global land cover data. Therefore almost all renewable energy models were only applied to local region rather than country level or larger areas. The launch of remote sensing satellites can quickly provide the large sources of remote sensing data, especially the large scale remote sensing data like global land cover remote sensing data, global land slope remote sensing data, global forest remote sensing data and country climate remote sensing data, etc. The data is real-time and economic. For example, we can obtain the global land cover data from Landsat ETM quickly and conveniently. The large scale remote sensing data can extend the geographic scope of renewable energy model to country level or larger regions. For example, with the help of country level land cover remote sensing data, we can determine potential transportation routes within country and then can extend the geographic scope of renewable energy model to the country level. Moreover, with the help of remote sensing data, the assessment to renewable energies potentials becomes easier than before. With remote sensing data we can evaluate the contemporary potentials of renewable energies with the specified spatial resolution (Wang et al. 2009). In the supply chain system of renewable energy, the potential of renewable energy plays a very important role (Wang et al. 2009; Wang and Koch 2010). In some cases, the small variation of its potential will influence the whole supply chain system. Additionally, by virtue of remote sensing data we can map the spatial distribution of renewable energy. This spatial distribution will also influence the whole supply chain system of renewable energy since the supply chain of renewable energy is directly related to the spatial distribution of renewable energy (Dunnet et al. 2008). As a result, it is interesting to integrate remote sensing data into the renewable energy models, especially the models focusing on the renewable energy supply chain.

1.4 Research questions and objectives

With respect to the above research motivation and the framework of GEOBENE (Global Earth Observation - Benefit Estimation: Now, Next and Emerging) project funding this dissertation, the following research questions are addressed:

- How to integrate remote sensing data to model electricity supply chain system of solar and wind energy?
• What is the improvement to model electricity supply chain system of solar and wind energy through integrating remote sensing data?
• Where are the optimal locations of solar and wind energy facilities in electricity supply chain of solar and wind energy?
• How much are the optimal capacity sizes (also called size in most literatures) of solar and wind energy facilities which should be set up in electricity supply chain of solar and wind energy?
• What is the relation between renewable energy consumptions and the locations of consumptions?
• In addition to the locations of consumptions, are there any other spatial factors which will influence renewable energy consumptions?

In order to tackle these research questions in a novel way, this dissertation carefully deals with the following research objectives:
• develop an Optimization Facility Location and capacity size model for solar energy facilities and wind energy facilities in electricity supply chain system integrating Remote sensing data (OFLR model) to determine the optimal locations and optimal capacity sizes of solar and wind energy facilities in electricity supply chain of solar and wind energy.
  1) the OFLR model will integrate remote sensing and Geographical Information System (GIS).
  2) the geographic scope of the OFLR model will focus on country level or eventually larger regions.
  3) the main input data of the OFLR model will be the economic data sets related with solar energy facilities and/or wind energy facilities, the remote sensing data sets, the spatial distribution of solar energy and/or wind energy and the spatial distribution of electricity demand.
• develop an Improved Spatial Trend model (IST model) to identify the spatial influencing factors of renewable energy consumptions, especially the location factor.
• present a spatial statistical model selection procedure to select the best IST model for renewable energy consumption.
1.5 Organization of the dissertation

This dissertation consists of eight chapters. The logical arrangement and the brief summary for each chapter are as follows:

**Chapter 2, Literature Review.** This chapter reviews the methods and theory related to modeling renewable energy. It also reviews the recent developments of the optimization renewable energy supply model and the statistical consumption model.

**Chapter 3, Design of OFLR Optimization Model for Supply and Chapter 4, Description of OFLR Model.** These two chapters focus on the development of the OFLR model. Chapter 3 mainly demonstrates the design philosophy of the OFLR model. In chapter 4, the hidden theory, the symbol system, the subscripts and sets, the variables and parameters of the OFLR model are described at length. Chapter 4 also describes the constraints, including the remote sensing-based constraints, and the objective function of the OFLR model. It furthermore introduces the way to tackle the pre-determined capacity sizes of renewable energy facilities in the OFLR model and illustrates the diagram of the OFLR model.

**Chapter 5, Description of IST Model for Consumption.** This chapter outlines the IST model as well as the spatial statistical model selection procedure.

**Chapter 6, Case Studies.** Two case studies are carried out in this chapter. The materials and results of these two case studies are presented in this chapter.

**Chapter 7, Discussion.** All relevant issues associated with the OFLR model and IST model have been expounded in this chapter. The results of two case studies are also discussed in this chapter.

**Chapter 8, Conclusions and Future Work.** The main findings obtained during the entire research period are summarized in this chapter. The chapter also points out future work.
2 Literature Review

2.1 Overview of modeling renewable energy

Since Krzhizhanovsky put forward the idea that energy economy should be approached comprehensively from production of energy sources to consumers (Belyaev et al. 1976), a number of methods of modeling renewable energy were developed which can found in, among others, Beaujean and Charpentier (1976), APAD(1985), Grubb et al. (1993), IIASA (1995), Kleinpeter (1995) and Hourcade et al. (1996). Beeck (1999) summarized these methods as eight kinds: 1) econometric; 2) macro-economic; 3) economic equilibrium; 4) optimization; 5) simulation; 6) spreadsheet; 7) backcasting, and 8) multi-criteria. In this section, we only review the optimization method and the statistical (i.e. econometric) method because these two methods will be used in this dissertation, and refer to the resulted models as the optimization model and the statistical model accordingly.

The optimization models employ the optimization techniques to optimize renewable energy system. In general, the optimized objects are renewable energy investment decisions. Through inputting data, the optimal investment decisions can be directly made. Because taking the optimization techniques, the outcome represents the best solution for given variables while meeting the specified constraints. Typically, the concrete optimized objects rely upon the taken economic theory. If the renewable energy models adopt the equilibrium theory, then the optimized objects reduce to the total surplus. By optimizing the total surplus, we can create the optimal renewable energy investment strategies. For example, an acronym for The Integrated MARKAL-EFOM System (TIMES) creates the energy investment strategies by maximizing the total surplus (Loulou et al. 2005).

Moreover, the optimization model can be used by utilities or municipalities to derive their optimal renewable energy investment strategies. In national energy planning, the optimization model is used to analyze the future of renewable energy system. In this case, the underlying assumption of model is that for given constraints, every acting agent is optimized. The mathematical knowledge behind the renewable energy optimization models is the linear programming, non-linear programming and mixed integer programming. The disadvantages of the optimization renewable energy models are (DHV
1984; Beeck 1999): 1). They require a relatively high level of mathematical knowledge and 2). The included processes must be analytically defined.

The statistical models take the statistical theory to predict the future renewable energy behaviors, based on its past behaviors. Therefore, the main merit of the statistical renewable energy models is the prediction which projects the past behavior into the future behavior. The statistical renewable energy models rely upon the aggregated data which had been measured in the past to predict the short- or medium-term future in terms of labor, capital, or other inputs and are frequently used to analyze the energy-economy interactions. As a result, the statistical renewable energy models always try to capture the main factors to predict the future renewable energy behavior as accurately as possible.

The statistical renewable energy models can be used to solve the issues on the energy demand. By analyzing the impact factors on the past renewable energy demand, we can develop the statistical renewable energy model for this demand. Then we can use this model to predict the future renewable energy demand. After that, we can input the future renewable energy demand into the optimization renewable energy models and obtain the future renewable energy investment strategies. Therefore, the statistical renewable energy model can be integrated into the optimization renewable energy models. For instance, the statistical method is employed by the LEAP model to model the renewable energy demand of the LEAP model (LEAP 2002). In addition, we can model the trend of renewable energy behaviors through the statistical method. By extrapolating the past trends of energy-economic activity and the renewable energy per capita ratios, we can predict the future energy-economic activity and the renewable energy per capita ratios.

However, the statistical models have certain disadvantages: 1). they have no way to represent a set of technologies, especially specific technologies (Beeck 1999). 2). since the main object of the statistical models is to project the past renewable energy behaviors into the future renewable energy behaviors, a reasonable stability of the economic behaviors is required. That is, no big fluctuations over time can be permitted (APDC 1985; Munasinghe 1988). 3). finally, the models are developed by experts rather than general people because developing these models needs lots of experience.
2.2 Mathematical approaches of renewable energy optimization models

The main mathematical approaches to develop renewable energy optimization model are linear programming, mixed integer programming and non-linear programming. A number of energy systems can be modeled with Linear Programming (LP). In mathematics, LP is a technique to optimize a linear objective function, subjecting to the linear equality and linear inequality constraints. In other words, LP determines the way to achieve the best outcomes (such as the maximum profit or minimum total cost) in a given mathematical model with some given list of requirements represented as the linear equations. LP is a practical technique which can be used to configure the energy system. LP can be used, for instance, to find the most profitable investment strategies that can be produced with the given inputs and the given output prices in the energy system. Linear programs can be expressed in formal mathematical terms as

\[
\begin{align*}
\text{Max/Min} & \quad C^T X \\
\text{s.t.} & \quad AX \leq b
\end{align*}
\]  

where \(X\) represents the vector of variables (to be determined), while \(C\) and \(b\) are vectors of coefficients (known) and \(A\) is a known matrix of coefficients. \(C^T X\) is the linear objective function and \(AX \leq b\) is the linear constraints.

LP is a relatively simple technique which quickly gives the results in a limited time and requires little mathematical knowledge for the users. We only formulate the energy system as the linear format, whether it is based on the input-out table or other linear methods, and then solve it through LP. LP is always used to almost all the optimization models and is applied to the national energy planning as well as the technology-related long-term energy researches. However, when using LP to model the energy system, we should be aware that the objective function and the constraints of the energy system must be translated into the linear format. That is, all coefficients must be constant. Because not all energy phenomena in reality can be modeled with linear format, we must look for some other mathematic methods to transform the non-linear energy system to the linear energy system. In addition, LP models can be greatly sensitive to the input parameter variations. A little variation of the input parameters will cause the big different results.

In renewable energy systems, one situation we often encounter is whether we should take
decision or not for a renewable energy problem. For example, we will determine whether we should install solar photovoltaics in Freiburg, Germany or not. The other situations always encountered are such as we have to choose one value from certain pre-determined values for renewable energy facilities. For example, we can only install solar photovoltaics with one of 20 MW, 100 MW and 200 MW capacities. For these two kinds of questions, we can take advantage of Mixed Integer Programming (MIP) to model renewable energy system. In MIP, the decisions such as Yes/No or 0/1 or the discrete decision problems are expressed as the dummy variable with value of 0/1. By using MIP, the variables which cannot reasonably assume any arbitrary value (e.g. small) –such as unit sizes of power plants– can be properly reflected in an otherwise linear model (World Bank 1991; Beeck 1999). Among MIP, the mixed integer linear programming is the most extensively researched. Mixed integer linear programming is actually an extension of LP in which a greater detail in formulating the technical properties and the relations is allowed during modeling renewable energy systems. In mathematical terms, the mixed integer linear programming can be expressed as

\[
Max/Min \quad C_1^T X_1 + uC_2^T X_2 \\
\text{s.t.} \quad A_1 X_1 \leq b_1 \\
\text{and } uA_2 X_2 \leq b_2
\]  

where \( u \) is the dummy variable with value of 0/1.

Some renewable energy systems can be modeled with Non-Linear Programming (NLP). In mathematics, NLP is to optimal an objective function, subjecting to the equality and inequality constraints where the objective function or some of constraints are nonlinear. Because the objective function or some of constraints are nonlinear, the global solution of NLP is always not assured. Also because the objective function or some of constraints are nonlinear, NLP should take longer time than LP to find out the solution. Due to these shortcomings, NLP is out-of-date in modeling renewable energy.

### 2.3 Energy scenario

Scenarios are images of possible alternatives for the future. They are more a consistent description of some general conditions assumed created than an attempt to forecast any actual picture of the precise technologies and the activities prevailing at the time (Sørensen 1999). It is also important to stress that the scenarios are not the predictions of the future. They should be presented as the policy options that may come true only if a
prescribed number of political actions are indeed carried out (Sørensen 1999). The first global energy scenario was constructed by the International Institute of Applied System Analysis (IIASA) during the late 1970s. Since then, a large number of global energy scenarios had been developed. Typically, the complete scenario is composed of four main ingredients, namely the population dynamics, the economic dynamics, the technological changes and the social dynamics respectively. The population dynamics reflect the demand dynamics of the energy system. The economic dynamics influence both the demand and the supply of the energy system because they are the measure of the consumption capability and the investment in the energy supply. The technological changes affect the demand side and supply side of the energy system.

A lot of energy scenarios are addressed in recent decades. These energy scenarios especially aim only at coping with one particle problem. For example, because of the considerations of energy security issues and the climate change impacts, the European Council has proposed that 15% of the total energy consumption should be produced from renewable energy resource in 2015. The European Parliament even suggested that the share of the renewable resources on total energy consumption should be 25% in 2020. To fulfill the objectives, the European Commission has constructed a renewable energy scenario which supports an obligatory share of 20% in 2020. Energy scenarios can be performed by using energy models which are widely used by national governments and international agencies to provide the scientific proof.

2.4 Equilibrium theory

The equilibrium theory, also called the supply-demand theory, is one of the extremely useful theories widely applied to the renewable energy models. Before describing the theory, some fundamental concepts should be introduced. The equilibrium price is defined as the price where the quantity demanded is equal to the quantity supplied. The equilibrium quantity is the quantity at which the quantity demanded and the quantity supplied is equal at a certain price. The surplus is the benefit quantity and can be distinct as the consumer surplus, the producer surplus and the total surplus. The consumer surplus is the amount that the consumers benefit by being able to purchase a product for a price which is less than they would be willing to pay. The producer surplus is the amount that the producers benefit by selling a product at a market price that is higher than they would be willing to sell for. The total surplus is the amount of the consumer surplus plus the
producer surplus. Fig 2-1 illustrates these concepts. In Fig 2-1, \( P_E \) is the equilibrium price and \( Q_E \) is the equilibrium quantity.

![Diagram of supply and demand curves with labeled surplus areas]

Intuitively, if the quantity of the commodities supplied in market is larger than the quantity demanded, according to the market mechanism, the market price of the commodities will decrease and then the consumers will buy more commodities. If the quantity of the commodities demanded in market, however, is larger than the quantity supplied, in terms of the market mechanism, the market price of the commodities will increase and then more commodities will of course be produced. When the quantity supplied is equal to the quantity demanded, the price and the quantity will be stable until the balance is broken. The equilibrium theory is formulated from the phenomenon as that in a competitive market, the price will function to equalize the quantity demanded by consumers and the quantity supplied by producers, resulting in an economic equilibrium of price and quantity. At this time, the total surplus is maximized. We can understand the equilibrium theory from the surplus point of view as well. On the basis of Fig 2-1, it is clearly known that if the price supplied is equal to the price demanded, the total surplus is the biggest. Otherwise the total surplus is not optimized. In practice, the essence of the produce as well as the incentive of the market will make the total surplus toward the
optimization such that the consumers and the producers will obtain the maximum profit.

In general, the equilibrium theories widely applied to the energy models are the general equilibrium theory and the partial equilibrium theory respectively. The general equilibrium theory assumes that the equilibrium prices for commodities exist and that all prices are at equilibrium with several or many markets. The partial equilibrium theory, however, assumes that the clearance on the market of some specific commodities is obtained independently from the prices and the quantities demanded and supplied in other markets.

2.5 Reference energy system

The introduction of Reference Energy System (RES) to energy models makes the mathematical formulation of energy models easier. Moreover, by the aid of RES we can conveniently know which data set we need for energy models. At present, RES becomes a pretty popular method of developing the energy models because it can easily express the relation among the capital flow, the commodity flow and technology. As a matter of fact, the essence of RES is to dedicate to the expression of technology. Therefore, it is better fitted for the bottom-up energy models. The famous energy models, such as Model for Energy Supply Strategy Alternatives and their General Environmental Impact (MESSAGE model) (Messner and Strubegger 1995) and an acronym for The Integrated MARKAL-EFOM System (TIMES model) (Loulou et al. 2005), etc., all employ this method.

There have, from energy system point of view, common components for energy system although the compositions of energy systems are different. Simply speaking, energy system consists of energy plants, materials and commodities. Energy plant will process, convert and distribute energy or commodities. Materials can be put into energy plants and be processed to commodities by energy plants. Commodity is the production of energy plant. For example, the solar energy system with photovoltaics consists of photovoltaics (energy plant), sun light (material) and electricity (commodity). The energy plants can be further distinguished as the process technologies, the conversion technologies and the end-use technologies. RES categorizes energy plants, materials and commodities into different types with respect to their characters. In general, the complete RES consists of five components: process technologies, conversion technologies, end-use technologies,
commodities and the demands for energy services. Typically, the term “PRC” is often used to represent a set of process technologies and the term “CON” denotes a set of conversion technologies. The terms “DMD” and “DM” are used to represent a set of end-use technologies and a set of demands for energy services respectively. The term “COM” is used to represent commodities. “PRC”, “CON” “DMD” and “DM” are bridged by “COM”. By this way the energy system is formed.

In order to deal with the greenhouse gases (GHG), we thought of GHG as a normal commodity which can be transacted. However, the “PRC” for GHG is a special process. It only processes GHG and can not generate any other commodities. Fig 2-2 illustrates a simple regional RES. The gray color represents the “PRC” and the green color represents the “CON”. The orange color, however, represents the “DMD” and the red color represents the “COM”. The demanded quantities for commodities are not listed in Fig 2-2. The middle commodities in RES are often not necessarily colored. For example, the raw gas in Fig 2-2 is the middle commodity and is not colored. Nevertheless, if unambiguous, it is also not necessary to distinguish “PRC”, “CON” and “DMD”. On the basis of Fig 2-2, we can clearly know the inputs and outputs of the whole system. We can also know how
many kinds of energy plants are used and what are the inputs and outputs for each kind of energy plant.

2.6 Spatial weight

The spatial weight is the reflection of spatial interactions and plays a greatly major role in spatial statistic. It can not deny that if the spatial weights are not right the statistical results are invalid. If the spatial weights, however, are correctly determined, the results are always interpreted. Anselin (1988, 2002) argued that the validity of estimates is pre-conditioned by the extent to which the spatial structure is correctly reflected in the spatial weights and it thus makes the deficiency for the interpretation of the results. The determination of spatial weights influences the following analysis, including the following spatial inference. In particular, it could potentially lead to the inference of spurious relationships and thus can not be incorporated into the spatial statistic models in advance. Since the spatial weights are pretty important for the spatial inference, we should discuss the spatial weights at length. This section restricts the attention to the discussion of spatial weight matrices.

The simple and widely applied spatial weight matrix is the binary spatial contiguity weight matrix. The first order spatial contiguity weight matrix \( W \) is defined as: if two regions \( i \) and \( j \) are connected each other, the element \( w_{ij} \) of the spatial weight matrix \( W \) is one, otherwise it is zero. The first order spatial contiguity weight matrix can be easily extended into the higher order spatial contiguity weight matrix through the matrix operator. However, the higher order spatial contiguity weight will have redundant and circular paths. Blommestein and Koper(1992) and Anselin and Smirnov (1996) provided the effective solutions for this problem. In addition, in reality not all regions are physically connected each other. As a consequence, it has to be very careful to determine the contiguity criterion.

Dacey (1968) presented a spatial weight matrix that took into account the relative area of the spatial units. The spatial weight matrix is defined

\[
w_{ij} = u_{ij} \alpha_i \beta_{ij}
\]  

(2-3)

where \( u_{ij} \) is a binary contiguity factor. When the unit \( i \) is connected with unit \( j \), \( u_{ij} \) is one. Otherwise it is zero; \( \alpha_i \) is the share of unit \( i \) in the total area of all spatial units in the system; \( \beta_{ij} \) is the proportion of the interior boundary of unit \( i \) which is in contact with unit
In the similar manner, Cliff and Ord (1973, 1981) suggested a spatial weight matrix that not only considered the distance but also took into account the relative length of the common border between two spatial units. Formally, the resulted spatial weights can be expressed as

$$w_{ij} = d_{ij}^{-a} \cdot \beta_{ij}^b$$  \hspace{1cm} (2-4)

where $d_{ij}$ is the distance between spatial unit $i$ and $j$; $\beta_{ij}$ is the proportion of the interior boundary of unit $i$ which is in contact with unit $j$; $a$ and $b$ are the parameters.

Other spatial weight matrices are often related with the particular phenomenon under studying. For example, Bodson and Peeters (1975) introduced a general accessibility spatial weight which took into account the influence of several channels of communication between regions, such as roads, railways and other communication links. Formally, the spatial weight can be described as

$$w_{ij} = \sum k_j \{ a/(1 + b \cdot \exp(-c_j d_{ij})) \}$$  \hspace{1cm} (2-5)

where $k_j$ stands for the relative importance of the means of communication $j$ and can be understood as the weight of communication $j$; $d_{ij}$ is the distance; $a$, $b$ and $c_j$ are parameters.

For the spatial weight matrices with parameters, the parameters and the spatial weight matrices should be determined in advance, or in a step separating from the rest of the spatial analysis.

**2.7 Spatial autocorrelation**

Spatial autocorrelation refers to the fact that similarity of locations tends to have similar attributes. If high values of attribute in one location are associated with high values of that attribute in neighboring locations, the spatial autocorrelation is positive, whereas when high and low value alternate, the spatial autocorrelation is negative. In time series analysis, in order to identify the autocorrelation between the current status and the former status or latter status, the time series autocorrelation coefficient is presented, defined as

$$\rho_{x_t, x_s} = \frac{\text{Cov}(x_t, x_s)}{\sqrt{\text{var}(x_t)} \cdot \sqrt{\text{var}(x_s)}}$$  \hspace{1cm} (2-6)
where $x$ is the random variable; the subscripts $t$ and $s$ represent the time points.

The magnitude of $\rho_{x,t}$ reflects the linear relation or non-linear relation between the current status and the former or latter status. More specifically, if $\rho_{x,t} = 1$, the two statuses have strongly positive linear relation. If, however, $\rho_{x,t} = -1$, the two statuses have strongly negative linear relation. If $\rho_{x,t} = 0$, the two statuses have no any relation. If $0 < |\rho_{x,t}| < 1$, the two statuses have some linear or non-linear relation. With respect to the values of $\rho_{x,t}$, we can present the Autocorrelation model (AR model).

In the context of spatial statistic, the spatial autocorrelation is, however, not so straightforward. The autocorrelation coefficient $\rho_{x,t}$ can not be used any more because the space is multi-direction whereas the time is one-direction. In order to measure the spatial autocorrelation in spatially autocorrelated phenomenon, we need new index.

Moran (1950) developed an index, later called Moran’s I coefficient, to measure the spatial autocorrelation. Let $x$ represent the interesting attribute, $\bar{x}$ denote the mean of the $x$ values, $z_{ij}$ denote the attribute similarity of $x$, $w_{ij}$ be the $i$-th row and $j$-th column entry of spatial weight matrices $W$, $N$ is the number of spatial units indexed by $i$ and $j$, then the Moran’s I coefficient can be expressed in mathematical terms as

$$I = \sum_{i} \sum_{j} w_{ij} z_{ij} / (s^2 \sum_{i} \sum_{j} w_{ij}) \quad (2.7)$$

where $z_{ij} = (x_i - \bar{x})(x_j - \bar{x})$; $s^2 = \sum (x_i - \bar{x})^2 / N$ denotes the sample variance.

The expected value of the Moran's I under the hypothesis of no spatial autocorrelation is

$$E(I) = \frac{-1}{N-1} \quad (2.8)$$

Its variance equals to

$$Var(I) = \frac{N^2 (N-1)S_1 - N(N-1)S_2 - 2S_0^2}{(N+1)(N-1)S_0^2} \quad (2.9)$$

where

$$S_0 = \sum_{i} \sum_{j} w_{ij}, i \neq j$$

$$S_1 = \frac{1}{2} \sum_{i} \sum_{j} (w_{ij} + w_{ji}), i \neq j$$
The magnitude of Moran’s I reflect the similarity (positive spatial autocorrelation) or dissimilarity (negative spatial autocorrelation) of spatial configuration. If 0<I<1, it indicates that the spatial configuration demonstrates the cluster and is similar. However, if 0>I>-1, it indicates that the spatial configuration is checkerboard and dissimilar. If I=0, it suggests that the spatial configuration is random (no spatial autocorrelation). For statistical hypothesis test, Moran’s I values can be transformed to Z-scores in which the values greater than 1.96 or smaller than −1.96 indicate that the spatial autocorrelation is significant at the significance level of 5%.

Geary (1954) developed another index to measure the spatial autocorrelation. The index is now called Geary’s C coefficient which is also known as Geary's Contiguity Ratio, Geary's Ratio, or the Geary Index. Formally, the Geary’s C coefficient can be defined as

\[ C = \frac{\sum_i \sum_j w_{ij} z_{ij}}{\sum_i \sum_j w_{ij} \sigma^2} \]  

(2-10)

where \( z_{ij} = (x_i - x_j)^2 \), \( \sigma^2 \) denotes the variance of the \( x \) values, or \( \sigma^2 = \sum_i (x_i - \bar{x})^2 / (N - 1) \).

The value of Geary’s C lies between 0 and 2. If 1<C<2, it indicates that the spatial configuration is dissimilar (negative spatial autocorrelation). The larger than 1 the Geary’s C, the more dissimilar the spatial configuration is. However, if 0<C<1, it suggests that the spatial configuration is cluster (positive spatial autocorrelation). The smaller than 1 the Geary’s C, the more cluster the spatial configuration is. If C=1, it shows no spatial autocorrelation.

Table 2-1 summarizes the values of Moran’s I and Geary’s C as well as the indication. On the basis of Table 2-1, we clearly know that Moran’s I is inversely related to Geary’s C, but it is not identical.

<table>
<thead>
<tr>
<th>Moran’s I</th>
<th>Geary’s C</th>
<th>Indication</th>
</tr>
</thead>
<tbody>
<tr>
<td>1&gt;I &gt; 0*</td>
<td>0&lt;C&lt;1</td>
<td>Positive spatial autocorrelation</td>
</tr>
<tr>
<td>I = 0*</td>
<td>C=1</td>
<td>No spatial autocorrelation</td>
</tr>
<tr>
<td>-1&lt;I &lt; 0*</td>
<td>2&gt;C&gt;1</td>
<td>Negative spatial autocorrelation</td>
</tr>
</tbody>
</table>
Cliff and Ord (1983) tested the performance of Moran’s I and Geary’s C and argued that Moran’s I was a measure for the global spatial autocorrelation while Geary’s C was more sensitive to the local spatial autocorrelation.

### 2.8 Spatial model specification test

To diagnose the spatial model specifications, we should present the appropriate hypothesis and then test the hypothesis. By this way, we can judge which specification is right. For spatial model specification, there are typically four test methods, which are Moran’s I test, Wald test, likelihood ratio test, and Lagrange multiplier test respectively. Suppose that the tested spatial hypothesis is

\[ H_0: g(\theta) = 0 \]
\[ H_A: g(\theta) \neq 0 \]

where \( g \) is a \( q \)-dimensional linear or nonlinear matrix function in the elements of the parameter vector \( \theta \).

We first introduce the Moran’s I test. Moran’s I test is greatly popular in spatial autocorrelation test for errors. Formally, this test can be expressed in matrix notation as

\[ I = \frac{e'W e}{S_0} \]

where \( e \) is an \( N \times 1 \) vector of Ordinary Least Square (OLS) residual \( Y - X \hat{\beta} \);
\( S_0 = \sum_{i} \sum_{j} w_{ij} \) is a normalizing factor.

The Moran’s I test distributed asymptotically as the normal distribution. Cliff and Ord (1973) deduced the first two moments for Moran’s I test under the normal assumption. The first two moments were given respectively as

\[ E[I] = (N / S_0) * tr(MW) / (N - K) \]

and

\[ Var[I] = \frac{(N / S_0)^2 [tr(MWMW') + tr(MWMW) + (tr(MW))^2]}{(N - K)(N - K + 2)} - (E[I])^2 \]
constant term 1); \( M = I - X(X'X)^{-1}X' \).

Given that the Maximum Likelihood (ML) estimates of the parameters vector of \( \theta \) under the null hypothesis are \( \theta_R \), the Wald test can then be formally expressed as

\[
W = g(\theta_R)'[G(\theta_R)V(\theta)G(\theta_R)]^{-1}g(\theta_R)
\]

(2-14)

where \( g(\theta_R) \) as a \( q \) by 1 vector of the values which result when the constraints \( (H_0) \) are evaluated for the ML parameter estimates; \( G \) as a \((3+K)\) by \( q \) matrix of partial derivations \( \partial g(\theta)/\partial \theta \), evaluated at the null for parameter estimates; \( K \) is the explanatory variables’ number (including constant 1); \( V \) is the estimated asymptotic variance matrix, which is the inverse of the Fisher information matrix \( I(\theta) \) \((I(\theta) = E(-\partial^2 L(\theta)/\partial \theta \partial \theta'))\)), \( L(\theta) \) is the log-likelihood function; \( L(\theta_R) \) is the log-likelihood function for the null.

In general, the Wald test follows the chi-squares distribution.

Suppose that the probability density functions or the probability mass functions of statistical models is \( f(x;\theta) \), the likelihood function is \( L(\theta) = L(\theta \mid x) = p(x \mid \theta) = f_{\theta}(x) \) which is a function of the parameter \( \theta \) with \( x \) held fixed at the observed value \( (i.e.\ the\ data) \), the LikelihoodRatio (LR) test is then defined as

\[
LR = \frac{f(x;H_A)}{f(x;H_0)}
\]

(2-15)

The LR test will reject the null hypothesis \( H_0 \) if the ratio \( LR \) exceeds a critical value \( c \). That is, the decision rule has the form

If \( LR \geq c \), reject \( H_0 \).

If \( LR < c \), accept (or don’t reject) \( H_0 \).

The critical value \( c \) is chosen usually based on a specified significance level \( \alpha \) through the relation: \( P_0(LR \geq c) = \alpha \) (if \( x \) is discrete, some randomization on the boundary may be needed). The Neyman-Pearson lemma states that the LR test is the most powerful among all level-\( \alpha \) tests.

In the environment of spatial statistics, the appearance of LR test has, however, a slight
difference. Formally, the LR test can in general be expressed as follows

\[ LR = 2[L(\theta) - L(\theta_R)] \]  

(2-16)

The decision rule is the same with the preceding. The LR test follows the chi-squares distribution.

The Lagrange Multiplier (LM) test, in recent decades, receives fairly much attention in spatial hypothesis tests (Anselin 1988; Bera and Yoon 1993; Anselin et al. 1996; Baltagi and Li 2001; Anselin 2006). The LM tests or the robust LM tests use an asymptotic variance matrix which is based on the Fisher information matrix to construct the known probability distribution. In mathematical terms, the LM test can be defined as

\[ LM = d_R^\top I(\theta_R)^{-1} d_R \]  

(2-17)

where \( d_R \) as the Rao’s score vector for \( L(\theta)(d(\theta) = \partial L/\partial \theta) \), evaluated at the null hypothesis \( H_0 \). \( I(\theta_R) \) is the information matrix \( I(\theta) = E(-\partial^2 L/\partial \theta \partial \theta^\top) \), evaluated at the null hypothesis \( H_0 \).

The LM test follows the chi-squares distribution.

### 2.9 Spatial trend model

The trend model in time series is used to analyze the change of the interesting feature with time. In spatial statistic, the trend model, however, is used to analyze the change of the interesting feature with the location. Moreover, the spatial trend model can also be used to interpolate the missing value and capture coarse scale pattern in the data. Formally, the spatial trend model can be expressed as

\[ y = A\theta + \varepsilon \]  

(2-18)

where the error terms, \( \varepsilon \), are always assumed to be independently normally distributed with zero mean and variances, \( \sigma^2 I \). \( I \) is identity matrix; \( A \) is a matrix of location coordinates of \( N \) sites (i.e. location factor); \( \theta \) is a vector of spatial trend model parameters; \( (a_{i1}, a_{i2}) \) is the coordinate of site \( i \), such that
The sum \((p+q)\) is called the order of the spatial trend model. Three widely familiar spatial trend models, which are linear spatial trend model, quadratic spatial trend model and cubic spatial trend model, result if \(p=q=1\), \(p=q=2\) and \(p=q=3\) respectively. The higher order models can be obtained by the same way.

When the orders of spatial trend models become bigger and bigger, and the dataset is clustered, the spatial trend models have, however, an ill-conditioned least squares problem (i.e. \(A'A\) is nearly singular) which makes the parameter estimates both difficult to find and unstable (Unwin 1975; Mather 1977; Shaw 1977; Unwin and Wrigley 1986; Haining 2003). On the other hand, if the orders are too low, then the goodness of fit of the models is sometimes rather bad and the residuals from the models tend to be spatially autocorrelated (Haining 2003). Unwin(1973) argued that the ill-conditioned least squares problem was caused by the excessive influence of solitary points away from any clusters. In order to solve the problem, Unwin and Wrigley (1986) made use of the concept of the leverage. The leverage of site \(i\) was defined as the \(i\)th diagonal element, \(h_{ii}\), of projection matrix \(H\)

\[
H = A(A'A)^{-1}A'
\]  

Unwin and Wrigley pointed out that the leverage was bounded between 0 and 1 (i.e. \(0 \leq h_{ii} \leq 1\)). The leverage of 0 indicated that an observation had no influence on the fit, whereas a maximal value of 1 indicated that a degree of freedom had been devoted to fitting the model to that observation.

Unwin and Wrigley’s method was simple and the least squares estimation methods could be used again. However, for Unwin and Wrigley’s method, the residuals still have chance of spatially autocorrelated. Because the main target of model intends to capture the main factors which reveal the studied phenomenon, we always assume the errors can not provide information any more. Since the spatially autocorrelated residuals can still provide information about the studied phenomenon, Unwin and Wrigley’s method is not...
always a good solution for ill-conditioned least squares problem.

Haining (2003) thus argued that it was safer to assume the spatially autocorrelated errors in order to solve the ill-conditioned least squares problem. Then he presented a spatial trend model with autocorrelated errors which was given by

\[ \begin{align*}
  y &= A\theta + \varepsilon \\
  \varepsilon &= \lambda W\varepsilon + \mu
\end{align*} \tag{2-20} \]

where \( \mu \) are assumed to be independently normally distributed with zero mean and variances, \( \sigma^2 I \).

Actually, Haining’s model specification is based on the interaction theory and can solve the ill-conditioned least squares problem. However, the parameter estimators for Haining’s model are involved in a highly nonlinear maximum likelihood procedure since the logarithm likelihood function will contain a Jacobian term as well as a sum of squares term.

### 2.10 Statistical model selection criterions

In order to understand the nature and then manage and control it, modeling the nature is always an effective way. Because the experience is different from one person to another, the understanding to the same phenomenon will be different. Consequently, the developed models are different. Therefore, we should answer such question: which model(s) do(es) capture the intrinsic factors of the underlying phenomenon? This question is essentially the model selection problem which depends upon the model selection criteria. In the context of statistical models, the model selection criteria are various.

The most popular model selection criterion is the R square. R square is used to measure the proportion of variability in a data set that is accounted for by a statistical model. In this definition, the term "variability" is defined as the sum of squares. We can decompose the variability in a data set as two parts: the explained variability and the unexplained variability. If we denote the total variability in a data set as TSS, the explained part as ESS, and the unexplained part as RSS, then the R square can be defined as

\[ R^2 = \frac{RSS}{TSS} \tag{2-21} \]
Because TSS=ESS+RSS, we can re-express R square as

$$R^2 = 1 - \frac{ESS}{TSS}$$  \hspace{1cm} (2-22)

So the value of R square is bounded within 0 and 1. The bigger the R square, the better the statistical model is. However, if we add more terms to the statistical model, the R square will always increase. In order to only consider the real factors which influence the researched phenomenon, we need adjust the R square. The adjusted R square only takes into account the terms which improve the model more than would be expected by chance. The adjusted R square is defined as

$$R_a^2 = 1 - (1 - R^2) \frac{N-1}{N-K-1}$$  \hspace{1cm} (2-23)

where \(N\) is the sample size; \(K\) is the number of explanatory variable (terms), excluding the constant term.

Akaike (1974), from the information perspective, presented a model selection criterion which was later called Akaike Information Criterion (AIC). In order to only evaluate the effective explanatory variables, AIC includes a penalty that is an increasing function of the number of estimated parameters. This penalty discourages overfitting. As a result, AIC not only rewards goodness of fit but also considers the penalty. Formally, the AIC can be written as

$$AIC = 2K - 2\ln(L)$$  \hspace{1cm} (2-24)

where \(N\) is the sample size; \(K\) is the number of explanatory variable (terms), excluding the constant term; \(L\) is the maximized value of the likelihood function for the estimated model.

Because the least squares methods are often used to estimate the parameters, we should translate the likelihood function in AIC into the least square function for convenience. If we denote the residual sum of square as RSS, the AIC can be described as

$$AIC = 2K + N[\ln(2\pi * RSS/N) + 1]$$  \hspace{1cm} (2-25)

where \(RSS = \tilde{e}'\tilde{e} \); \(\tilde{e}\) is residuals.

The selection criterion for AIC is: the smaller the AIC, the better the model is.
In the similar manner, Schwarz (1978) developed a later called Schwarz Information Criterion (SIC). In general, SIC penalizes free parameters pretty strongly whereas AIC attempts to find the model best explaining the data with a minimum number of the estimated parameters. In mathematical terms, the SIC can be defined as

$$SIC = K \ln N - 2\ln(L)$$

(2-26)

where \(N\) is the sample size; \(K\) is the number of explanatory variable (terms), excluding the constant term; \(L\) is the maximized value of the likelihood function for the estimated model.

Similar transformation for AIC can be applied to SIC. The transformed SIC is

$$SIC = K \ln N + N\ln(RSS/N)$$

(2-27)

Actually, the above model selection criteria reviewed are presented mainly for time series models. For spatial statistical models, due to the special character of spatial datasets (they always show spatial autocorrelation), we should modify the above model selection criteria such that they can be used to spatial statistical model selection (Geary 1954; Cliff and Ord 1981).

2.11 Optimization models for supply

A number of studies have been carried out on the determination of optimal locations and optimal capacity sizes of solar energy facilities and wind energy facilities in the supply chain of solar and wind energy. These studies all involve the optimization method. Some of them will involve remote sensing and geographical information system.

Azadeh et al. (2008a) studied the optimal locations of solar plants in solar energy supply chain in Iran. To determine the optimal geographical locations of solar plants in Iran, Azadeh et al. developed an integrated hierarchical Data Envelopment Analysis (DEA) model based on linear program. In this model, a number of local and geographical parameters had been defined and were considered in two levels: level 1 for location of solar plants in a number of nominal cities, and in level 2 for locating the position of solar plant in each city. This model would enable the energy policy makers to select the best-possible location for construction of a solar power plant with lowest possible costs. For identifying the optimal locations of wind plants in wind supply chain in Iran, Azadeh
et al. (2008b) also developed a DEA-based multivariable linear program model. In their model, the wind speed, the quantity of proper geological areas and the quantity of proper topographical areas along with two non-geographical parameters: the distance of power distribution network and the cost of wind devices were considered. The model was a country level model.

Milligan and Arig (1998) studied the optimal capacity size and optimal location for wind plant in wind supply chain. In order to determine the optimal capacity size and optimal location for wind plant, they developed an optimization model with load-duration curve. The model was applied to six U.S. cities.

Burke and O’Malley (2008) investigated the optimal wind power location on transmission systems. In order to find the optimal locations of wind power on the existing transmission system network to reach desired renewable energy penetration targets in a secure, least-cost manner, they presented a linear programming optimization model which was accomplished by using the probabilistic load flow technique based on direct-currency load flow and recorded load. In this model, the geographical statistical dependencies of individual bus load and wind power outputs as well as the temporal dependencies of the conventional plant unit-commitment process on total system load and wind patterns were accounted for.

In recent years, extensive researches have been conducted which are based on remote sensing/Geographic Information System (GIS) to identify the optimization locations for solar and wind plants in supply chain. Broesamle et al. (2001) studied the site of the solar thermal power plants in Northern Africa via an Evaluation system for Solar Thermal Power Stations (STEPS) with remote sensing data and GIS. Their research was composed of four stages: 1) the first stage was to rank and evaluate the potential sites of solar thermal power plants. In this stage, the land slope data, land cover data, distance from cooling water resources data were involved. 2) the second stage was to calculate the solar direct normal irradiation at potential sites of solar thermal power plants since they thought that the most important parameter for the site selection of solar thermal power plants was the solar direct normal irradiation. In this stage, the cloud index data from visible and infrared image channel of the METEOSAT weather satellite, aerosols transmittance data from Global Aerosol Data Set by Kopke et al. (1997), water vapor and ozone data sets
taken from NASA Vapour Project (NVAP) and NASA Ozone Mapping Spectrometer project (TOMS) were used. 3) in the third stage, the electricity cost of solar thermal power station operating in solar only mode was calculated with respect to the investment cost, the infrastructure cost for connecting the plant to the roads and the public grid, the annual running expense of operation and maintenance, the economic lifetime, the mean capital interest rate and the net annual solar electricity yield. 4) the fourth stage was to rank all potential sites and select the optimal one in terms of the solar direct normal irradiation and the electricity cost at potential sites of solar thermal power plant. Their model can be used for the assessment, evaluation and ranking of sites of solar thermal power plant projects, giving project developers, governments, intergovernmental institutions and other decision makers a well founded basis for planning and designing the build out of solar power capacity world wide. However, their model did not take into account the solar electricity demand and thus can not determine the optimal capacity size of solar thermal power.

Wang et al. (2007) studied the location choice for the solar thermal plant in China. They developed a location choice model integrated remote sensing data and GIS which was based on the technology-resource-economy principle and took into account the availability of data in China. Their model took account of the solar direct normal irradiation, land use, the transportation road, the salary of residents, the population of region, water resource, solvable salt, gas, maintenance expense, the spatial distribution of human and economic activities, the environmental and ecological influence, the investment and the rewarding period. In their model, in order to calculate the solar direct normal irradiation, they acquired the water vapor data and aerosols transmittance data by using the Moderate Resolution Imaging Spectroradiometer (MOD IS) terra/Aqua satellite, and acquired the cloud data from the NOAA. After these data sets were collected, they can calculate the solar direct normal irradiation. For some region where these data sets were not available, they used CoKriging method to interpolate the solar direct normal irradiation data for these regions. Their model can provide the spatially explicit suggestion for the policy decision makers. However, their model was only a conceptual framework and it needed a lot of work to translate the conceptual model into the mathematical model. In addition, their model did not consider the demand for the solar electricity and thus can not determine the optimal capacity size of the solar thermal plant. Their model was applied to China.
Johannessen and Korsbakken (2000) studied the potential windmill locations in wind supply chain in coastal zones. They argued that in the planning for wind mill park installations it was of fundamental importance to have sufficient information about the wind characteristics for different seasons. Because the standard wind measurements were very local and would not properly resolve the spatial variations at a nearby site of the wind field, they took use of the Synthetic Aperture Radar (SAR) to calculate the surface wind speed since the accuracy of SAR image was normally within ±2ms⁻¹ and SAR was independent of daylight and clouds. Once acquiring the wind speed, they can calculate the wind energy potential and then ranked the wind energy potential for each site to determine the optimal location for windmill in coastal zones. Their model can be used to large and regional scale. However, their model only considered the wind potential and should be complementary to the mathematical models which considered more factors, such as the energy demand, the economic factors and the environmental factors, etc.

In order to understand the factors necessary to determine site suitability for wind farm in the wind supply chain in the UK, Baban and Parry (2001) developed the simplified GIS-assisted wind farm locating criteria for the UK. They took advantage of the questionnaires to develop the wind farm location criteria. Their location criteria for wind farm consisted of a number of constraint factors which included topography, wind speed and direction, land use/cover, population, access, hydrology, ecology and resources. Then they converted these constraints into the map layers and assigned the scores to each map layer with respect to their sensitivity. The scores ranged from 0 to 10 in which a score of 0 indicated no constraint and a score of 10 indicated total constraint. In order to produce a single index of evaluation, each map layer was assigned a weight. There were two ways to assign the weights to the map layers. One way was to assign the same weight to all map layers and another way was to assign the different weights to the map layers based on the perceived importance. For the first way, all scores and all weights were combined using the “ADD" OVERLAY operation of GIS and then the suitability map was calculated. In this suitability map, 0 represents ideal location for wind farm whereas 10 represent unsuitable locations. For the second way, the weights for map layers were allocated based on a pairwise comparison for the relative importance of the two layers by rating rows relative to columns and entering the ratings into a matrix. The procedure required that the principle eigenvector of the pairwise matrix was computed to produce a best-fit set of weights assigned to map layers. Then the scores and the weights were combined to
generate a suitability map. The suitability map had a slight increase in the geographical extent for the most suitable sites, compared with the suitability map with same weight.

In the above studies, the models developed respectively by Azadeh et al. (2008a), Azadeh et al. (2008b), Milligan and Arig (1998) and Burke and O’Malley (2008) take advantage of the simple objective functions to determine the optimal locations for solar and wind energy facilities in supply chain. They are easily implemented and can be applied to different special cases through modifying the objective functions. However, these models do not consider remote sensing data. Remote sensing satellites can provide high resolution and global real-time remote sensing data. These data sets are pretty useful in estimating the potential of renewable energy which plays a major role in optimizing the locations and capacity sizes of renewable energy facilities (Wang et al. 2009; Wang and Koch 2010). Moreover, these models mainly focus on special parts of supply chain rather than the whole supply chain. As a consequence, they can not exactly determine the optimal capacity sizes for solar and wind energy facilities. Furthermore, these models are often the local region level models.

The models developed respectively by Broesamle et al. (2001), Wang et al. (2007), Johannessen and Korsbakken (2000) and Baban and Parry (2001) mainly take use of remote sensing data and GIS to determine the optimal locations of solar and wind energy facilities in supply chain. These models can be implemented easily. However, these models do not consider the demands for the products of solar and wind energy. Therefore they have no way to determine the optimal capacity sizes of solar and wind energy facilities. In addition, the geographic scopes of these models are always the local region level.

In order to maximally take advantage of remote sensing data and GIS and to investigate the large scale supply chain system of solar and wind energy, we should develop country level or larger regions level model in which the optimal locations and optimal capacity sizes of solar and wind energy facilities will be determined and remote sensing data and GIS will be integrated.

**2.12 Statistical models for consumption**

For renewable energy consumption, the related researches are hard to find. This is partly
because of two reasons: 1) The researches for renewable energy are advocated in recent years (refer to 1.1 research background), which makes datasets hard to be acquired. 2) Although renewable energy has good promise, the acceptance to renewable energy needs a long time, which also makes datasets unavailable. However some exceptions can be found in biomass consumption. This is probably biomass is the most accepted renewable energy in human life. As a result, this section only reviews biomass consumption models.

Lefevre et al. (1997) developed a model to study the growth of fuelwood consumption. The main idea of the model was based on the 3 different scenarios for future fuelwood consumption in Asian countries: (1) business-as-usual, (2) “green” energy, and (3) massive fossil fuel promotion. In this model, the future growth in fuelwood consumption was determined by its past and present rate of growth. The responsiveness of fuelwood consumption was measured by the economic and population growth.

Lambert D’Apote (1998) analyzed biomass consumption in Africa, Latin America, China, East Asia and South Asia and developed a biomass consumption model for these continents. In his model, Lambert D’Apote presumed that the level of per capital biomass consumption would be influenced by GDP per capita, the availability of supply of biomass fuels, the prices of final biomass fuels, the level of conventional energy use, the price of alternative fuels and the share of urban population. Using the available historical data from a number of Latin American and Asian countries as well as one or two Africa countries, Lambert D’Apote regressed the model and concluded that the most important factor influencing per capital biomass consumption was per capital income whereas the smallest influence on the per capital biomass consumption was the fuel prices. Lambert D’Apote also pointed out that the income elasticities were negative but smaller than 1 and the price elasticities for the competing fuel were positive (when significantly different from zero). That is, per capital use of biomass decreased as per capital income increased and grew if the price of the competing fuel increased (Victor and Victor 2002).

Victor and Victor (2002) analyzed the factors influencing biomass consumption in developing countries and in U.S.. They concluded that four factors largely explained the level of traditional biomass consumption. The four factors were respectively the income, the availability of fuels, the urbanization and the industrialization. They argued that in countries with higher per capital income, industrialization and urbanization, the share of
biomass in energy consumption was smaller. In countries with abundant forests close to population centers, firewood played a large role in the traditional energy system. Then Victor and Victor developed a statistical model to evaluate the biomass intensity variation against the growth in income. The model was defined as

\[ Y = aX^b \]  

(2-28)

where \( Y \) is the biomass intensity of GDP; \( b \) is the progress ratio; \( X \) is the GDP per capital; \( a \) is a scale coefficient that reflects the biomass and GDP ration of the first observation.

The model was estimated by non-linear estimation method and applied to almost all African countries except the Central African Republic.

The above researches investigated the influencing factors of biomass consumption and constituted the basis of the research of this dissertation.
3 Design of OFLR Optimization Model for Supply

Model is the simplification of the complex reality and only extracts the main elements which are necessary to reveal the underlying complex reality. The main aims of model design are to try to simplify the original complex reality while capturing the main elements which can at most represent the original complex reality. Model design directly influences the flexibility, extendibility and mathematical formulation of model. Good model design can develop a flexible and extendible model while the objective(s) of model can be implemented successfully. However, bad model design will make the mathematical formulation of model extremely difficult and the resulted model is hard to be extended. The chapter will introduce the design philosophy of the OFLR model. In this chapter, renewable energy refers to solar energy and/or wind energy.

3.1 Model objectives and scope

The main targets of the OFLR model are to determine the most appropriate locations and capacity sizes for solar energy facilities and wind energy facilities in electricity supply chain system of solar and wind energy through integrating remote sensing data. The complete electricity supply chain of solar and wind energy consists of supply regions, solar and wind energy potentials, solar and wind energy facilities, transportation facilities, electricity demand and demand regions (see Section 3.3). Additionally, the model should be able to

1) integrate remote sensing and GIS,

2) perform in large geographical areas, such as at country level or larger regions,

3) explicitly take into account the spatial information about supply regions, demand regions and facilities, etc.,

4) take into account the spatial variation of demand,

5) take into account the spatial variation of supply,

6) include policy tools, such as setting carbon emission tax on emissions from production and transportation,

7) take into account the environmental constraints,

8) take into account new advanced solar and wind energy facilities,

9) take into account not only the continual capacity sizes of solar and wind energy
facilities but also the pre-determined discrete capacity sizes of solar and wind energy facilities,
10) take into account electricity import and export,
11) determine the prices of electricity for demand regions,
12) take into account the by-products,
13) be applied to the supply chain system with solar plants and/or wind plants.

The model will not be used to optimize the profit of one single plant, but rather to maximize the social welfare of the demands.

### 3.2 Solar and wind energy potentials with remote sensing data

In renewable energy supply chain, the availability of renewable energy directly influences the locations and capacity sizes of renewable energy facilities and thus always influences the structure of supply chain. Broesamle et al. (2001) and Wang et al. (2007) argued that solar potential was the most important parameter for the site selection of solar thermal power plant. Then they used the solar Direct Normal Irradiation (DNI) on ground to evaluate the solar potential. The solar DNI on ground can be defined as

$$E_{Di} = E_{ot} \tau_{or} \tau_{ol} \tau_{cg} \tau_{wv} \tau_{ak} \tau_{ck}$$  \hspace{1cm} (3.1)$$

where $E_{Di}$ is the solar direct normal irradiation (Wm$^{-2}$); $E_{ot}$ is the normal-incidence extraterrestrial solar spectral irradiance (0.29-4.00μm) at the mean earth-sun distance (or modified by standard orbital correction factors if simulations are to be done for a specific time of the year) (Wm$^{-2}$); $\tau_{or}$ is the transmittance due to molecular (Rayleigh) scattering; $\tau_{ol}$ is the transmittance due to ozone absorption; $\tau_{cg}$ is the transmittance due to the uniformly-mixed gases (CO$_2$, O$_2$, CH$_4$, and N$_2$O); $\tau_{wv}$ is the transmittance due to water vapor; $\tau_{ak}$ is the transmittance due to aerosols combining the processes of scattering and absorption; $\tau_{ck}$ is the transmittance due to the attenuation of irradiation by cloud.

In order to take into account the influence of the solar zenith angle, the solar direct normal irradiation was modified as
\[ E_{D,\lambda} = E_{\cos} \mu_0 \tau_{\lambda \lambda} \tau_{\omega \lambda} \tau_{\nu \lambda} \tau_{\sigma \lambda} \tau_{\epsilon \lambda} \]  

(3-2)

where \( \mu_0 \) is the cosine of the solar zenith angle \( Z_0 \).

In general, in order to acquire the above parameters, the normal traditional methods have lots of difficulties or have no way to obtain these data sets. Moreover, the data sets for these parameters are not real-time, simultaneous and global. Remote sensing satellite, however, can provide these data sets. The \( \tau_{\nu \lambda} \) and \( \tau_{\omega \lambda} \) can be acquired from the MODIS Terra/Aqua satellite with spatial resolution of 500×500 \( \text{m}^2 \) (Wang et al. 2007). The \( \tau_{\epsilon \lambda} \) can be acquired from NOAA with spatial resolution of 1×1 \( \text{km}^2 \) (Wang et al. 2007). The \( \tau_{\nu \lambda} \) can be acquired from LandSat satellite. Fig 3-1 shows the world 22-year (July 1983 - June 2005) annual average solar direct normal irradiation with spatial resolution of 0.5×0.5 degrees. The data source is acquired from NASA (NASA 2008) with spatial resolution of 1×1 degrees and is interpolated via ordinary Kriging to generate the solar direct normal irradiation with spatial resolution of 0.5×0.5 degrees.

![Solar direct normal irradiation (kWh/m^2/day)](image)

**Value**

- **High**: 9.003943
- **Low**: 1.736197

*Fig 3-1: World 22-year (July 1983 - June 2005) annual average solar direct normal irradiation.*

For wind energy, its potential can be defined as
\[ E = \frac{1}{2} \rho v^3 \]  

(3-3)

where \( E \) is wind energy (Wm\(^2\)); \( \rho \) is air density (kgm\(^{-3}\)); \( v \) is wind speed (ms\(^{-1}\)).

In the similar manner, the data sets for the parameters of wind energy potential can be more easily acquired from remote sensing satellites. For example, the air density can be acquired from Landsat. The wind speed can be acquired from the Synthetic Aperture Radar (SAR) (Vachon et al. 1996; Scoon et al. 1996; Korsbakken et al. 1997). Fig 3-2 shows the wind speed map of Danish seas (Denmark), which is retrieved from SAR image with spatial resolution of 150×150 m\(^2\) on 22 October 2004 (Christiansen 2006). In Fig 3-2, the arrows represent the wind speed direction.

In summary, using remote sensing data to evaluate solar and wind energy potentials is not only possible but also economic because not only we can acquire the real-time and simultaneously global data from remote sensing satellites but also the data is
labor-released. Further, we can take use of remote sensing satellites to obtain the remote sensing data for regions which are no way to be accessed by human.

### 3.3 Simplification of supply chain

The electricity supply chain system of solar and wind energy considered by the OFLR model is composed of the electricity supply chain of solar energy and the electricity supply chain of wind energy, described by Fig 3-3. This supply chain system is complicated. In order to formulate the OFLR model mathematically, we should simplify the supply chain system.

![Reference energy system of supply chain system of solar and wind energy.](image)

For biomass, the locations of biomass plants are not necessarily the same with the regions of biomass supply. Biomass has to be always transported to biomass plants for biomass processing. However, for solar and wind energy, solar plants and wind plants should be located in the same area with the region of the usable solar and wind energy. In addition, the region must have a reasonable area. That is, the area is not too big or too small. We assume that all supply regions of solar and wind energy have reasonable areas and the solar and wind plants are situated at the centroids of supply regions. That is, there is no chance to transport solar and wind energy from supply region to solar plants and wind
In electricity supply chain system of solar and wind energy (i.e. Fig 3-3), electricity can be stored as battery and then battery can be used to provide electricity. In the OFLR model, no storage is permitted.

Hydrogen is suggested as energy carrier to transport electricity and heat. However, Bossel (2005) pointed out that the solar electricity transferred by hydrogen would force three quarters of the solar power generators to work exclusively for covering the energy losses of the hydrogen conversion chain. Under this perspective, the use of hydrogen has been discarded by the TRANS-CSP project (DLR 2006). In the supply chain system of the OFLR model, the use of hydrogen to transport electricity is not considered as well.

The second method to transport electricity is High Voltage Alternating Current Transmission (HVAC) in which the high voltage alternating current is used. The whole HVAC is composed of lines and stations. The alternating current is produced in a power plant by a generator, whose magnet is driven mechanically and passes three 120°-shifted coils during one rotation. Accordingly, the induced alternating currents are also 120°-phase-shifted. The main advantage of alternating current is the simple regulation of voltage and frequency. In addition, the engines driven by alternating current can be produced small, compact and cheap (Leuschner 2005). HVAC will create an unusable reactive power and reduce the effective power capacity which will result in a drop of voltage along the line. However, because the maximum transferable load and the transmission length are more limited by the drop of voltage along the line, in practice the installation for the compensation are used every 600 kilometers (Rudervall et al. 2000). The HVAC system will also have the voltage-dependent losses in the form of gas discharge in areas of heavy curved surface and high field strength (DLR 2006). The total losses in the HVAC system come to 15%/1000 km (380 kV) and 8%/1000 km(750 kV) respectively (DLR 2006).

The third method to transport electricity is High Voltage Direct Current transmission (HVDC) which uses the direct current. In the same with HVAC, HVDC is also composed of lines and stations. Compare with HVAC, HVDC has diverse advantages. First of all, the transmission length is only limited by ohmic resistance. The cheaper the power input
is, the less important the heat losses are. Moreover, there are no capacitive, inductive or dielectric losses which would result in a drop along the line. Furthermore, the lines costs and the lines losses for HVDC are lower than those of HVAC. So it is especially fitted for long distance transmission. May (2005) and DLR (2006) compared the losses and investment cost of HVAC and HVDC transmission lines at comparable voltage levels for a transmission of 5GW which were showed in Table 3-1. In addition, HVDC can contribute to the stability and the controllability. No short-circuit current can be transmitted.

Table 3-1. Cost and performance parameters of HVAC and HVDC with transmission of 5GW (May 2005; DLR 2006)

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Unit</th>
<th>HVAC</th>
<th>HVDC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Operation Voltage</td>
<td>kV</td>
<td>750</td>
<td>±600</td>
</tr>
<tr>
<td>Overhead line losses</td>
<td>%/1000km</td>
<td>8%</td>
<td>6%</td>
</tr>
<tr>
<td>Sea cable losses</td>
<td>%/100km</td>
<td>60%</td>
<td>50%</td>
</tr>
<tr>
<td>Terminal losses</td>
<td>%/station</td>
<td>0.2%</td>
<td>0.2%</td>
</tr>
<tr>
<td>Overhead line cost</td>
<td>M€/1000 km</td>
<td>400-750</td>
<td>1000</td>
</tr>
<tr>
<td>Sea cable cost</td>
<td>M€/1000 km</td>
<td>3200</td>
<td>5900</td>
</tr>
<tr>
<td>Terminal cost</td>
<td>M€/station</td>
<td>80</td>
<td>80</td>
</tr>
</tbody>
</table>

HVDC has the disadvantage of not being directly transformable to another voltage, by which the erection of networks with different voltage levels becomes difficult (DLR 2006). In addition, it is not easy to switch off the current with conventional switches at a high network voltage.

Nowadays, a power of about 75,000 MW is transmitted by HVDC lines in more than 90 projects all over the world (DLR 2006). Table 3-2 lists the countries adopting HVDC.

Table 3-2. Countries adopting HVDC (DLR 2006)

<table>
<thead>
<tr>
<th>HVDC/country</th>
<th>Start of operation</th>
<th>Power [MW]</th>
<th>Voltage ±[kV]</th>
<th>Length [km]</th>
</tr>
</thead>
<tbody>
<tr>
<td>SACOL/Sardinia-Corsica-Italy</td>
<td>1967</td>
<td>300</td>
<td>200</td>
<td>423</td>
</tr>
<tr>
<td>Cahora-Bassa/Mozambique-South Africa</td>
<td>1977-1979</td>
<td>1930</td>
<td>533</td>
<td>1420</td>
</tr>
<tr>
<td>Inga-Shaba/Congo</td>
<td>1982</td>
<td>560</td>
<td>500</td>
<td>1700</td>
</tr>
<tr>
<td>Itaipu/Brasilia</td>
<td>1984-1987</td>
<td>6300</td>
<td>600</td>
<td>800</td>
</tr>
<tr>
<td>BalticCable/Sweden-Germany</td>
<td>1994</td>
<td>600</td>
<td>450</td>
<td>250</td>
</tr>
<tr>
<td>SwedPol/Sweden-Poland</td>
<td>2000</td>
<td>600</td>
<td>450</td>
<td>260</td>
</tr>
<tr>
<td>Italy-Greece</td>
<td>2001</td>
<td>500</td>
<td>400</td>
<td>310</td>
</tr>
<tr>
<td>Murraylink/Australia</td>
<td>2002</td>
<td>220</td>
<td>150</td>
<td>177</td>
</tr>
<tr>
<td>NorNed/Norway-Netherland</td>
<td>2007</td>
<td>700</td>
<td>450</td>
<td>580</td>
</tr>
</tbody>
</table>
In the OFLR model, we only consider the HVDC system and the future higher efficiency transmission systems, and assume that the HVDC system is composed of lines and stations regardless of the transportation distance. In order to better mathematically formulate the model, we regard the HVDC system as a normal plant in which the input and output are the same (i.e. electricity) and further assume that one HVDC system connects one supply region and one demand region. Although in the OFLR model, we only consider HVDC system to transport electricity, if any other higher efficiency transportation systems for electricity are possible we can easily add the systems to the OFLR model without big change.

3.4 Regular grid of supply and demand regions

In order to naturally integrate remote sensing data into the OFLR model, we use the square grids with appropriate specified spatial resolution to cover the studied area because remote sensing images are grid-based. Thus we can obtain the grid network for studied area. In the grid network, we can identify the supply regions and the demand regions to the grids. Fig 3-4 depicts the grids of Germany with spatial resolution of 0.5×0.5 degrees. In Fig 3-4, the green grids represent the demand regions and the red grids represent the supply regions.

In order to naturally express the demand regions, we can aggregate the neighbor demand grids into a polygon which approaches the demand region in reality at most. Then we take the centroids of the polygon as the centroids of the aggregated demand grids. For example, we can aggregate the two demand grids of the left bottom of Fig 3-4 into one big demand grid if necessary. For supply grids, we can also aggregate the supply grids. We first decide the reasonable buffer zone for each supply grid with renewable energy potential. Then we choose the supply grids with the strongest renewable energy potential and aggregated the neighbor supply grids of them which fall into the buffer zone of the grids into polygons. The centroids of polygons are regarded as the centroids of aggregated supply grids and also the locations of renewable energy facilities. However, we must cautiously choose the reasonable buffer zone for supply grids. Normally, we do not recommend taking this way to aggregate supply grids.
The advantage of grid structure of the studied area is that it is easier to integrate the remote sensing-based constraints with remote sensing data into the OFLR model. We can create the buffer zone for remote sensing-based constraints and generate a grid map for each constraint. Then we only judge the grid maps. Through this way, a lot of remote sensing-based constraints with remote sensing data can be integrated into the OFLR model.

3.5 Spatially explicit

In the OFLR model, the spatial distribution of solar and wind energy supply, the spatial distribution of solar and wind energy demand, the spatial information about the locations of solar and wind plants, the spatial information about landuse and roads, etc., are
explicitly expressed. The spatial information is pretty important in determining the optimal locations of renewable energy facilities as well as the optimal capacity sizes of renewable energy facilities and should be explicitly considered (Dunnet et al. 2008; Wang and Koch 2010). For example, the spatial distribution of solar energy supply, the spatial distribution of solar energy demand, the spatial information about the potential locations of solar plants and the spatial information about roads can determine the distances from the potential locations of solar energy facilities to demand regions. These distances will influence the optimal locations and optimal capacity sizes of solar energy facilities. As a result, the spatial distribution of renewable energy supply, the spatial distribution of renewable energy demand and the spatial information about the locations of renewable energy plants play a major role in determining the optimal locations and optimal capacity sizes of renewable energy facilities and should be explicitly considered (Dunnet et al. 2008; Wang and Koch 2010). This is consistent with our insight that for the same renewable energy supplies or the same renewable energy demands, if the spatial distributions of them are different, the optimal locations and optimal capacity sizes of renewable energy facilities are always different. As a consequence, the OFLR model explicitly takes into account the spatial distribution of solar and wind energy supply, the spatial distribution of solar and wind energy demand and the spatial information about the locations of solar and wind energy facilities, etc.

### 3.6 Flexibility

The OFLR model not only can model the supply chain system of solar and wind energy with solar plants and wind plants such as solar photovoltaics, solar thermal plants and wind farm, but also can model the supply chain system of solar energy with only one or more kinds of solar plant such as solar photovoltaics or/and solar thermal plants or model the supply chain system of wind energy with only one or more kinds of wind plant such as wind farm. This character can make the model maximal flexibility. In the case that we need only consider one kind of solar energy facility or one kind of wind energy facility, this merit is greatly appreciated. Sometimes, because of the availability of the data sets, we have to take into account one kind of solar energy facility or wind energy facility. In this case, we can still use the OFLR model.

Moreover, the OFLR model can be extended to include other modules. The OFLR model has set the interface to connect other modules. These modules can be coupled into the
OFLR model with loose coupling (Wang et al. 2009). For example, we can use the BIRD model (BIRD 1984) to calculate solar energy potential and then input the potential into the OFLR model. On the other hand, we can embed the OFLR model into other energy models, such as the TIMES model (Loulou et al. 2005).
4 Description of OFLR Model

After designing the OFLR model, we should develop a computational OFLR model. However, it is still a long way to walk before we develop a computational OFLR model with respect to the model design. We should formulate the OFLR model in mathematical terms. In order to mathematically formulate the OFLR model, we should reveal the theory behind the OFLR model. The chapter will introduce the hidden economic theory of the OFLR model and will demonstrate how to formulate the OFLR model in mathematical terms. Finally the chapter will illustrate the diagram of the OFLR model. In this chapter, renewable energy refers to solar energy and/or wind energy.

4.1 Economic theory

The economic theory employed by the OFLR model is the equilibrium theory with fixed demand (i.e. the supply-demand theory with fixed demand), illustrated by Fig 4-1.

According to the equilibrium theory, when the price supplied is equal to the price demanded, the total surplus is the biggest and then the social welfare is maximized. When the demand is fixed, we only maximize the producer surplus with respect to Fig 4-1 and in this case the social welfare is also maximized. Because the social welfare is the most important factor that the renewable energy policy decision-makers should consider when
making renewable energy policy and also because the final aim for most of all energy systems is to obtain the maximal social welfare, it is reasonable for the OFLR model to adopt the equilibrium theory with fixed demand and maximize the social welfare. In fact, the TIMES model also takes the equilibrium theory with fixed demand and maximizes the social welfare (Loulou et al. 2005).

For the supply chain system of solar and wind energy considered by the OFLR model, the demands for electricity are fixed. In this case what we need do for modeling this supply chain system is to maximize the producer surplus. According to Fig 4-1, maximizing the producer surplus in this case is equal to minimize the total cost of the supply chain system. As a result, the OFLR model will minimize the total cost of the electricity supply chain system of solar and wind energy. When the total cost of the supply chain system is minimized, the price of electricity of demand region is equal to its marginal value.

4.2 Symbol system

The symbols of variables and parameters play a greatly important role in reading, understanding, extending and managing the model. Good symbol system can make people quickly understand the representation of symbol and thus conveniently read the model formulation. Moreover, good symbol will help and motivate people extend the model. Bad symbol system will, however, make the model formulation unread and hard to be understood. It is one of the sources that make the extendibility of the model worse. For the OFLR model, a concise symbol system is elaborately designed. The symbol system can make people quickly understand the original ideas of the OFLR model, the constraints and objective function of the OFLR model. Using the symbol system, people can extend the OFLR model or borrow the symbol system to other models. The symbol system of the OFLR model distinguishes the subscripts, the variables and the parameters. Moreover, the symbol system of the OFLR model also distinguishes renewable energies, renewable energy facilities, electricity, transportation facilities, by-products and the operation time. In addition, the symbol system of the OFLR model distinguishes the studied area and non-studied areas as well as the communications between the studied area and non-studied areas.

In the symbol system of the OFLR model, the subscript is used to indicate what it is. Normally, if it has unambiguous, the subscript is always the first small letter of the name
of the thing indicated by the subscript. Otherwise, the subscript is the abbreviation of the
name of the thing which can be understood by most people. For example we use “e” to
indicate renewable energy. All possible values of subscript are called set and accordingly
we can use the capital letter of the subscript to indicate the largest possible values and use
the capital letter of the subscript with “~” to indicate the set. So, 
\[ E = (\ldots, E) \]
indicates all
renewable energies.

In the symbol system of the OFLR model, the variables are distinguished into two kinds.
One is the binary variables and another is the normal variables. Binary variables can only
take the value of 0 or 1. Under no any other limitations, normal variables, however, can
take any values. In the symbol system of the OFLR model, we use “u” indicate the binary
variable. For normal variables, the situation is a little complicated. We take use of two
parts to represent the normal variables. We make use of “VAR” as the first part to indicate
they are the normal variables and make use of the second part to clearly indicate what the
normal variables are. The second part is always organized with the order of from rough
detail to more detail. We take respectively the letters from rough detail and more detail to
form the second part via “_”. For example, if we want to determine the capacity size of
solar photovoltaics, we will use VAR_P_SIZE.

In the symbol system of the OFLR model, the parameters will be distinguished as two
kinds. One is the normal parameter and another is the professional parameter. In the
similar manner with the normal variables, the normal parameter is always organized with
the order of from rough detail to more detail. We take respectively the letters from rough
detail and more detail to form the parameter name via “_”. For example, the cost of solar
photovoltaics can be denoted as P_COST. The professional parameter always takes the
special symbol that is commonly recognized by this field. For example, we can use \( \rho \)
denote the efficiency of renewable energy plants.

In all, the OFLR model will use the following symbols for variables and parameters.
“ENY” represents renewable energy. “P” represents renewable energy plant. “SIZE” is
the capacity size of the facility. “PROD” is the production. “TP” is the transportation
plant. “TRS” is the transport. “D” is the demand regions. “AVB” is the available. “DMD”
is the demand. “COM” is the commodity. “VAR” is the variable.
4.3 Subscripts and sets

In the OFLR model, we should indicate supply regions, renewable energies (i.e. solar energy and wind energy), renewable energy facilities (i.e. solar plants and wind plants), transportation facilities, by-products, demand regions and the operation time, etc. So we use “s” to indicate supply region, use “e” to indicate renewable energy, use “p” to indicate renewable energy facility, use “tp” to indicate transportation facility, use “c” to indicate by-product (since by-product is commodity), use “d” to indicate demand region and use “y” to indicate the operation time, etc. Accordingly, \( \tilde{S} = \{1, \ldots, S\} \) indicates all supply regions. \( \tilde{E} = \{1, \ldots, E\} \) indicates all renewable energies. \( \tilde{P} = \{1, \ldots, P\} \) indicates all renewable energy facilities. \( \tilde{TP} = \{1, \ldots, TP\} \) indicates all transportation facilities. \( \tilde{C} = \{1, \ldots, C\} \) indicates all by-products. \( \tilde{D} = \{1, \ldots, D\} \) indicates all demand regions. \( \tilde{Y} = \{1, \ldots, Y\} \) indicates all studied times, etc.

4.4 Variables description

In the OFLR model, we will use the following variables to describe the mathematical formulation of the OFLR model.

- **VAR__ENY__P** \(_{s,e,p,y}\) is the amount of the available renewable energy \( e \) at supply region \( s \) which can be used by renewable energy plant \( p \) at time \( y \).
- **VAR__P__SIZE** \(_{s,p,y}\) is the capacity size of renewable energy plant \( p \) at supply region \( s \) at time \( y \).
- **VAR__P__NSIZE** \(_{s,p,y}\) is the new capacity size of renewable energy plant \( p \) to be set up at supply region \( s \) at time \( y \).
- \( u_{s,p,y} \) is the binary variable. If \( u_{s,p,y} = 1 \), the renewable energy plant \( p \) will be set up at supply region \( s \) at time \( y \). Otherwise if \( u_{s,p,y} = 0 \), the renewable energy plant \( p \) will not be set up at supply region \( s \) at time \( y \).
- **VAR__P__PROD** \(_{s,p,y}\) is the electricity amount produced by renewable energy plant \( p \) at supply region \( s \) at time \( y \).
- **VAR__TRS__S** \(_{s,tp,y}\) is the electricity amount transported by transportation facility \( tp \) from supply regions \( s \) at time \( y \).
- **VAR__COM__OTHER** \(_{s,c,y}\) is the amount of by-product \( c \) transported from other sources around the supply region \( s \) to this supply region \( s \) at time \( y \).
- **VAR__TP__SIZE** \(_{tp,y}\) is the capacity size of transportation facility \( tp \) at time \( y \).
- **VAR__TP__NSIZE** \(_{tp,y}\) is the new capacity size of transportation facility \( tp \) to be set up at...
time $y$.
$u_{y,p,y}$ is the binary variable. If $u_{y,p,y} = 1$, the transportation facility $tp$ will be set up at time $y$. Otherwise if $u_{y,p,y} = 0$, the transportation facility $tp$ will not be set up at time $y$.
$VAR\_TRS\_D_{y,p,d,y}$ is the electricity amount transported by transportation facility $tp$ to demand region $d$ at time $y$.
$VAR\_IMPORT_{d,y}$ is the amount of electricity imported into demand region $d$ at time $y$.
$VAR\_EXPORT_{d,y}$ is the amount of electricity exported from demand region $d$ at time $y$.

### 4.5 Parameters description

In the OFLR model, we will use the following parameters to describe the mathematical formulation of the OFLR model.

$MAP(s,p)$ is the map of supply regions and renewable energy plants. It used to indicate that the renewable energy plant $p$ locates at supply region $s$.

$MAP(e,p)$ is the map of renewable energy and renewable energy plants. It used to indicate that renewable energy $e$ can be processed by renewable energy plant $p$.

$MAP(s,tp)$ is the map of supply regions and transportation facilities. It used to indicate that electricity resulting from supply region $s$ can be transported by transportation facility $tp$.

$MAP(tp,d)$ is the map of transportation facilities and demand regions. It used to indicate that electricity can be transported by transportation facility $tp$ to demand region $d$.

$ENY\_AVB\_S\_e,y$ is the amount of the available renewable energy $e$ at supply region $s$ at time $y$.

$P\_PRICE_{s,p,y}$ is the price of setting up renewable energy plant $p$ at supply region $s$ at time $y$.

$P\_PROD\_PRICE_{s,p,y}$ is the price of producing electricity via renewable energy plant $p$ at supply region $s$ at time $y$.

$\rho_{p,y}$ is the efficiency of renewable energy plant $p$ converting renewable energy to electricity at time $y$.

$\gamma_{p,c,y}$ is the efficiency of renewable energy plant $p$ converting renewable energy to by-product $c$ at time $y$.

$COM\_DMD_{s,c,y}$ is the demanded amount of by-product $c$ around supply region $s$ at time $y$.

$COM\_PRICE_{s,c,y}$ is the price of by-product $c$ transported from other source to supply region $s$ at time $y$. 
\( TP\_{PRICE}_{ip,y} \) is the price of setting up transportation facility \( tp \) at time \( y \).

\( TP\_{TRS\_PRICE}_{ip,y} \) is the price of transporting electricity via transportation facility \( tp \) at time \( y \).

\( \lambda_{ip,y} \) is the loss ratio of transportation facility \( tp \) at time \( y \).

\( IMP\_PRICE_{d,y} \) is the imported price of electricity for demand region \( d \) at time \( y \).

\( EXP\_PRICE_{d,y} \) is the exported price of electricity for demand region \( d \) at time \( y \).

\( DMD_{d,y} \) is the demanded amount of electricity in demand region \( d \) at time \( y \).

### 4.6 Constraint equations

#### 4.6.1 Economics

Typically, we buy the facilities in one time. However, the capital costs for facilities are annualized through a periodic payment of total installed capital cost (\( TC \)) as an annual cost (\( ANCOST \)). If the economic lifetime for facility is \( t_e \) and suppose that the interesting rate is \( IR \) (typically the moderate interesting rate is assumed as 8%), representing the risk associated with the return on investment relative to an alternative allocation of capital, then we have the following equation

\[
ANCOST = TC \left( \frac{IR}{1 - (1 + IR)^{-t_e}} \right) 
\]  
(4-1)

#### 4.6.2 Renewable energy potential

The theoretical potential of renewable energy is the theoretical limit for renewable energy. As a matter of fact, not all the theoretical amount of renewable energy is used to renewable energy plants. In many cases, there are several different reasons which make it impossible to collect 100% of renewable energy for plant production. The consideration of the available and suitable area for energy purpose is greatly important. Some areas should be left to residents and ecosystem, etc., and others are not fitted for renewable energy plants. Consequently, the geographical potential of renewable energy is presented, defined as the theoretical potential of renewable energy timing the suitable area and then dividing the total area. Actually, calculating the geographical potential of renewable energy is pretty complicated and many factors should be considered. This dissertation only takes into account the impaction of landuse. Similar with Sørensen (1999), the dissertation adopts the suitable factors for land categories to figure out the suitable area
sometimes the suitable area is also called the available area). The suitable factor is the value assigned to the considered land category which reflects how much ratio of the land category can be used to renewable energy. Sørensen (1999) proposed some suitable factors for land categories illustrated by Table 4-1. However, quantification of the suitable factors is arbitrary and relies upon the experience. The land categories data can often be acquired from remote sensing data.

<table>
<thead>
<tr>
<th>Land category</th>
<th>Suitable factor</th>
<th>Land category</th>
<th>Suitable factor</th>
</tr>
</thead>
<tbody>
<tr>
<td>Urban area</td>
<td>0.00</td>
<td>Tundra</td>
<td>0.01</td>
</tr>
<tr>
<td>Bioreserve</td>
<td>0.00</td>
<td>Grassland</td>
<td>0.01</td>
</tr>
<tr>
<td>Forest</td>
<td>0.00</td>
<td>Extensive grassland</td>
<td>0.05</td>
</tr>
<tr>
<td>Agriculture</td>
<td>0.01</td>
<td>Desert</td>
<td>0.05</td>
</tr>
<tr>
<td>Shrubland</td>
<td>0.01</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

For renewable energy $e$, the total availability of renewable energy $e$ at supply region $s$ must be larger than the amounts of renewable energy $e$ processed by renewable energy plants locating at this supply region $s$, which is described by following equation

$$
ENY_{AVB} s,e,y \geq \sum_{p=1}^{P} VAR_{ENY} P s,e,p,y \quad s \in \tilde{S}, e \in \tilde{E}, y \in \tilde{Y} \quad (4-2)
$$

### 4.6.3 Renewable energy plants

We describe renewable energy plants (i.e. solar plants and wind plants) from the inputs and outputs of renewable energy plants point of view. For renewable energy plant $p$ at supply region $s$, the electricity production of renewable energy plant $p$ is limited by the amount of renewable energy $e$ for this renewable energy plant $p$. That is,

$$
VAR_{P PROD} s,e,p,y \leq \rho_{p,y} \cdot VAR_{ENY} P s,e,p,y \quad MAP(e,p), s \in \tilde{S}, e \in \tilde{E}, p \in \tilde{P}, y \in \tilde{Y} \quad (4-3)
$$

Electricity produced by renewable energy plant $p$ should be smaller than or equal to the capacity size of renewable energy plant $p$ timing the efficiency of renewable energy plant $p$ for electricity.

$$
VAR_{P SIZE} s,p,y \geq VAR_{P PROD} s,p,y / \rho_{p,y} \quad s \in \tilde{S}, p \in \tilde{P}, y \in \tilde{Y} \quad (4-4)
$$
The capacity size of renewable energy plant \( p \) at time \( y \) is the sum of all renewable energy plant \( p \) which are set up in the current and previous time and still available at current time \( y \). Provided that \([t', y]\) is such an interval that all renewable energy plant \( p \) which are set up during the interval will be still available at time \( y \), then

\[
VAR_{p\_SIZE}^{y} = \sum_{t=t'}^{y} (VAR_{p\_NSIZE}^{t} * u_{t\_p}) \quad s \in S, p \in \tilde{P}, y \in \tilde{Y} \tag{4-5}
\]

The total produced electricity by renewable energy plants at supply region \( s \) will be transported by transportation facilities. That is,

\[
\sum_{p=1}^{P} VAR_{p\_PROD}^{t} \geq \sum_{t=1}^{TP} VAR_{p\_TRS} S^{t} \quad s \in S, t \in \tilde{S}, y \in \tilde{Y} \tag{4-6}
\]

In some cases, the renewable energy plant will produce some by-products, such as heat. Around the renewable energy plant, there will be the demand for these by-products. Suppose that the efficiency of renewable energy plant \( p \) for by-product \( c \) is \( \gamma_{p,c,y} \), the amount of by-product \( c \) provided by the other sources to supply region \( s \) is \( VAR_{COM\_OTHER}^{s,c,y} \) and the demand for by-product \( c \) around supply region \( s \) is \( COM\_DMD^{s,c,y} \), then the following equation is defined

\[
\sum_{c=1}^{C} \sum_{p=1}^{P} (\gamma_{p,c,y} * VAR_{ENY}^{p,s,c,y}) + VAR_{COM\_OTHER}^{s,c,y} \geq COM\_DMD^{s,c,y} \quad s \in S, c \in \tilde{C}, y \in \tilde{Y} \tag{4-7}
\]

For renewable energy plants, the unit cost per plant capacity and the unit cost per plant production will be strongly influenced by scales. They decrease with larger plants or equipments. For example, the unit cost of 80GW solar plant can be expected to be less expensive per GW set up than that of 8GW solar plant, even though both plants are made with the same materials and technology. So does the unit cost per plant production. To adjust the difference, the scale effect is presented and can be modeled by

\[
\frac{Cost_a}{Cost_0} = \left( \frac{Size_a}{Size_0} \right)^{SF} \tag{4-8}
\]

where \( SF \) is the scaling factor; \( Cost_a \) is the cost of plant with \( Size_a \); \( Cost_0 \) is the reference cost of plant with reference \( Size_0 \).
Broesamle et al. (2001) provided the reference cost for solar thermal plant with 200MW, which was detailed by Table 4-2.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Parabolic trough</td>
<td>LS-3</td>
</tr>
<tr>
<td>Solar collector aperture area</td>
<td></td>
</tr>
<tr>
<td>using a dry cooling system</td>
<td>1.228 km</td>
</tr>
<tr>
<td>using an evaporation cooling tower</td>
<td>1.124 km</td>
</tr>
<tr>
<td>using a once-through cooling system</td>
<td>1.075 km</td>
</tr>
<tr>
<td>Required land area (cost free)</td>
<td>3 times solar collector area</td>
</tr>
<tr>
<td>Plant investment</td>
<td></td>
</tr>
<tr>
<td>using a dry cooling system</td>
<td>460 Mio USD</td>
</tr>
<tr>
<td>using an evaporation cooling system</td>
<td>420 Mio USD</td>
</tr>
<tr>
<td>using a once-through cooling system</td>
<td>405 Mio USD</td>
</tr>
<tr>
<td>Infrastructure costs</td>
<td></td>
</tr>
<tr>
<td>per km road</td>
<td>185,000 USD</td>
</tr>
<tr>
<td>per km high tension grid</td>
<td>125,000 USD</td>
</tr>
<tr>
<td>per km pipeline (once-through)</td>
<td>2 Mio USD</td>
</tr>
<tr>
<td>per km pipeline (evaporation)</td>
<td>305,000 USD</td>
</tr>
<tr>
<td>Operating costs</td>
<td></td>
</tr>
<tr>
<td>personnel</td>
<td>2.7 Mio USD per year</td>
</tr>
<tr>
<td>operation and maintenance</td>
<td>1% of investment per year</td>
</tr>
<tr>
<td>insurance</td>
<td>1% of investment per year</td>
</tr>
<tr>
<td>Economic lifetime</td>
<td>25 years</td>
</tr>
</tbody>
</table>

### 4.6.4 Transportation plants

In the OFLR model, electricity is transported by transportation facilities (i.e. HVDC or higher efficiency plants). As a consequence, the capability of transportation plant $tp$ must be able to handle the total amounts of electricity transported by this transportation plant $tp$ from supply region $s$. Equation (4-9) describes the case.

$$
VAR_{TP \_ SIZE}_{p,y} \geq \text{VAR}_{TRS \_ S_{s,\tilde{p},y}} \_ MAP(s, tp), s \in \tilde{s}, tp \in \tilde{T}, y \in \tilde{Y}
$$

(4-9)

The capacity size of transportation plant $tp$ at time $y$ is the sum of all transportation plant $tp$ which are set up in the current and previous time and still available at current time $y$. Provided that $[t', y]$ is such an interval that all transportation plant $tp$ which are set up during the interval will be still available at time $y$, then
\[
\text{VAR\_TP\_SIZE}_{p,y} = \sum_{t=1}^{y} (\text{VAR\_TP\_NSIZE}_{p,t} \times u_{p,t}) \quad tp \in TP, \; y \in \tilde{Y} 
\]

(4-10)

For transportation plant \(tp\), the total amount of electricity transported by this transportation plant \(tp\) from supply region \(s\) is composed of two parts: one is the amount transported by transportation plant \(tp\) to demand region \(d\) and another is the loss. The following equation describes the electricity transportation and the loss with loss ratio \(\lambda_{p,y}\):

\[
\text{VAR\_TRS\_S}_{s, tp, y} = \text{VAR\_TRS\_D}_{d, tp, y} l(1 - \lambda_{p,y}) \\
\tilde{\text{MAP}}(s, tp), \text{MAP}(tp, d), tp \in \tilde{TP}, d \in \tilde{D}, y \in \tilde{Y} 
\]

(4-11)

The loss is a greatly important factor for transportation plants. It consists of line loss and station loss. Table 4-3 details the cost for HVDC with voltage \(\pm 800\)kV, unit capacity 5GW and the loss for each part of HVDC.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overhead line investment</td>
<td>350M€/1000 km</td>
</tr>
<tr>
<td>Sea Cable Investment</td>
<td>2500M€/1000km</td>
</tr>
<tr>
<td>Converter stations investment</td>
<td>350 M€/station</td>
</tr>
<tr>
<td>Overhead line and cable losses</td>
<td>2.5%/1000km</td>
</tr>
<tr>
<td>Station loss</td>
<td>0.9%/station</td>
</tr>
<tr>
<td>Economic lifetime</td>
<td>40 years</td>
</tr>
<tr>
<td>Operation&amp;Maintenance cost</td>
<td>1% of investment per year</td>
</tr>
</tbody>
</table>

In determining the loss of transportation facilities and the cost of transportation facilities, for example, in determining the loss and the cost for HVDC, we should calculate the transportation distance of electricity. Leduc (2009) studied the methods used to calculate the distance between two points in the supply chain system and summarized them as two ways. This dissertation adopts one of these two ways for studied grid area. The way can be described as: The average actual distance (in km or m), \(d\), is defined as the average direct distance multiplied by a dimensionless factor accounting for irregularities in the transportation network called the tortuosity factor (Overend 1982; Heller et al. 2004). The ratio is calculated by the straight distances from grid to grid with the real distances. Normally, this ratio can vary from 1 for straight roads to 3 for mountainous terrain (Heller et al. 2004). Rentizelas et al. (2009) and Leduc (2009) recommended the value of 1.4.
4.6.5 Demand regions

For demand region $d$, the transported electricity from supply regions adding the imported electricity subtracting the exported electricity has to be equal to or larger than the demanded amount of electricity at demand region $d$. This will be described by equation (4-12).

\[
\sum_{tp=1}^{TP} \text{VAR\_TRS\_D}_{tp,d,y} + \text{VAR\_IMPORT}_{d,y} - \text{VAR\_EXPORT}_{d,y} \geq \text{DM}\_D_{d,y}
\]

(4-12)

\[d \in \tilde{D}, y \in \tilde{Y}\]

The marginal value of the equation (4-12) is the marginal electricity price of demand region $d$.

4.6.6 Remote sensing-based constraints

When determining the optimal locations and optimal capacity sizes for renewable energy plants, certain factors should be considered. These factors might be the environmental constraints, the technical constraints, the social constraints, the economic constraints and the politic constraints, and could be regarded as the restrictive areas or constrained areas where the renewable energy plants are not permitted. For example, Perpina et al. (2009) took into account the wetlands and lakes, airports and residential areas, etc., for the site of the bioenergy plant. These factors might be expressed as the remote sensing-based constraints and the corresponding data can be acquired from remote sensing satellites. In the OFLR model, these remote sensing-based constraints with remote sensing data are also considered.

In the OFLR model, two kinds of remote sensing-based constraints are considered. One is that the remote sensing-based constraints can be translated into “0” or “1”. “0” means that the constraint is not satisfied whereas “1” means that the constraint is satisfied. Another is the quantity remote sensing-based constraints, such as the total available area of one potential location has to be enough to set up the renewable energy plants. For the first kind of remote sensing-based constraints, we create a buffer zone with specified distance for each selected geographic entity/feature of the remote sensing-based constraint. Thus different remote sensing-based constraints correspond to different widths of buffer zones.
Then a binary grid map for each remote sensing-based constraint is created in which if the cells fall within the constrained areas (i.e. the buffer zone) of this constraint, then the cells are assigned “0” and if the cells fall outside the buffer zone, the cells are assigned “1”. For example, for land slope we can take the buffer zone with distance of 0 for each land slope grid and then judge if the land slope of the grid is less than 5 degree, then the grid is assigned the value of “1”. Otherwise, the grid is assigned the value of “0”. By multiplying all grid maps of constraints, a final constraint map is calculated. This final constraint map can be regarded as a final remote sensing-based constraint of the OFLR model. The procedure is illustrated by Fig 4-2.

![Diagram showing the operation procedure of remote sensing-based constraints.](image)

**Fig 4-2: Operation procedure of remote sensing-based constraints.**

For the second kind of remote sensing-based constraints, we only calculate the quantities of the considered factor of the remote sensing-based constraint for the grids and then input them into the model to judge whether they are satisfied with the constraint or not. For example, the required area for renewable energy plant must be larger than a fixed number. So we can calculate the areas of the grids and then input them into the OFLR model to determine whether the required area for renewable energy plant with specified capacity size is larger than the area of grid or not.
4.7 Objective function

So far, we can define the objective function of OFLR model. According to the economic theory employed by the OFLR model, the objective function is to minimize the total system cost of the whole supply chain system. The whole supply chain system is mainly composed of supply regions, renewable energies, renewable energy plants, electricity, transportation plants and demand regions. As a consequence, the system cost is composed of the cost for renewable energy plants and transportation plants as well as the related cost. Hence, the objective function is defined as

\[ g = \sum_{y=1}^{Y} \sum_{s=1}^{S} \sum_{p=1}^{P} P_{\text{COST}}_{s,p,y} + \sum_{y=1}^{Y} \sum_{p=1}^{P} P_{\text{PROD \_COST}}_{s,p,y} + \sum_{y=1}^{Y} \sum_{p=1}^{P} TP_{\text{COST}}_{y,p} + \sum_{y=1}^{Y} \sum_{p=1}^{P} TP_{\text{TRAS \_COST}}_{y,p} + \sum_{y=1}^{Y} \sum_{d=1}^{D} IMP_{\text{COST}}_{d,y} - \sum_{y=1}^{Y} \sum_{d=1}^{D} EXP_{\text{COST}}_{d,y} + \sum_{y=1}^{Y} \sum_{x=1}^{X} \sum_{c=1}^{C} COM_{\text{COST}}_{x,c,y} + f(\text{emission}) \]

The different summands are:

(1) cost for setting up renewable energy plants at supply regions. It is the function of the capacity sizes of renewable energy plants set up (variables) and the corresponding binary variables.

(2) cost for producing electricity via renewable energy plants at supply regions. It is the function of the electricity production of renewable energy plants (variables).

(3) cost for setting up transportation plants. It is the function of the distance of the lines of transportation plants and the capacity sizes of transportation plants set up (variables) as well as the corresponding binary variables.

(4) cost for transporting electricity. It is the function of the transported electricity amount (variables).

(5) cost of importing electricity from non-studied regions. It is the function of the electricity amount imported into demand regions (variables).

(6) cost of exporting electricity to non-studied regions. It is the function of the electricity amount exported from demand regions (variables).

(7) cost of transporting by-products from other sources. It is the function of the by-products’ amounts transported from other sources (variables).

(8) cost for abating the emission resulting from production and transportation. It is the
function of the produced and transported electricity amount (variables).

\[
P_{\text{COST}}_{s,p,y} = \sum_{t=1}^{y} (P_{\text{PRICE}}_{s,p,t} \times VAR_{P_{\text{NSIZE}}_{s,p,t}} \times u_{s,p,t})
\]

\[
P_{\text{PROD\_COST}}_{s,p,y} = P_{\text{PROD\_PRICE}}_{s,p,y} \times VAR_{P_{\text{PROD}}_{s,p,y}}
\]

\[
TP_{\text{COST}}_{q,p,y} = \sum_{i=1}^{y} (TP_{\text{PRICE}}_{q,p,i} \times VAR_{TP_{\text{NSIZE}}_{q,p,i}} \times u_{q,p,i})
\]

\[
TP_{\text{TRS\_COST}}_{q,p,y} = TP_{\text{TRS\_PRICE}}_{q,p,y} \times VAR_{TP_{\text{S}}_{q,p,y}} \times \tilde{\text{MAP}}(s, tp)
\]

\[
IMP_{\text{COST}}_{d,y} = IMP_{\text{PRICE}}_{d,y} \times VAR_{IMP_{\text{IMPORT}}_{d,y}}
\]

\[
EXP_{\text{COST}}_{d,y} = EXP_{\text{PRICE}}_{d,y} \times VAR_{EXP_{\text{EXPORT}}_{d,y}}
\]

\[
COM_{\text{COST}}_{s,c,y} = COM_{\text{PRICE}}_{s,c,y} \times VAR_{COM_{\text{OTHER}}_{s,c,y}}
\]

Finally, the OFLR model is defined as

\[
\text{Min } g
\]

\[
s.t.
\]

\[
(4-2), (4-3), (4-4), (4-5), (4-6), (4-7), (4-9), (4-10), (4-11), (4-12)
\]

\[
\text{all remote sensing-based constraints}
\]

\[
\text{all variables except binary variables } \geq 0
\]

\[
\text{all binary variables } \in [0,1]
\]

The model is an ordinary Mixed Integer Program (MIP) and can be solved using standard MIP techniques (Wolsey 1998). The model was developed in the commercial software GAMS using the solver BARON (Sahinidis and Tawarmalani 2009).

The OFLR model also takes into account the pre-determined discrete capacity sizes for renewable energy facilities. Provided that the pre-determined discrete capacity sizes for renewable energy facilities are \(AS = (0, ..., AS)\) and a set of costs \(P_{\text{COST\_SIZE}}\) for capacity sizes and a set of binary variables \(v\) are defined accordingly, then for this case, the following equations should be added to the OFLR model.

\[
P_{\text{COST}}_{s,p,y} = \sum_{t=1}^{y} \sum_{as=0}^{AS} (P_{\text{COST\_SIZE}}_{s,p,as,t} \times v_{s,p,as,t}) \quad (4-13)
\]
In the same manner, we can predetermine the discrete capacity sizes for transportation facilities and add the similar equations into the OFLR model. Then the model becomes the mixed integer linear program and can be solved in the commercial software GAMS using the solver CPLEX (McCarl et al. 2008).

In the OFLR model, the conversion among units is not considered. When necessary, the appropriate conversion factors of units should be added to the OFLR model.

### 4.8 Diagram

The diagram of the OFLR model is illustrated by Fig 4-3. Remote sensing data, the spatial distribution of renewable energy potential\(^2\), the spatial distribution of electricity demand, the geographically explicit potential locations of renewable energy plants and some other parameters such as cost of renewable energy plants and cost of transportation plants, etc., are first processed into the right units and formats before they are input into the OFLR model. Then they are input into the OFLR model. When the running of the OFLR model is finished, an optimization will be reached. As a consequence, the locations and capacity sizes of renewable energy plants, the locations and amount of renewable energy supplied, the system cost of the whole electricity supply chain system of renewable energy, the electricity prices of demand regions, and the emission tax, etc. will be optimally obtained.

---

\(^2\) The renewable energy potential can be obtained from remote sensing data.
Fig 4-3: Diagram of the OFLR model.
5 Description of IST Model for Consumption

5.1 IST model

Although Unwin and Wrigley’s method (1986) can to some extent solve the ill-conditioned least squares problem of spatial trend model, a chance remains that the residuals are autocorrelated. By assuming that the residuals are autocorrelated, Haining’s method (2003) solves the ill-conditioned least squares problem of spatial trend model and can avoid the autocorrelated residuals problem. However, the estimation for the parameters of Haining’s model involves a non-linear program and it therefore must take longer time to estimate the parameters. Furthermore, in the case when the dataset is clustered and the goodness of fit is so low that increasing order of spatial trend model can not improve the $R^2$ and eliminate the spatial autocorrelation among residuals, the Haining’s method, however, is not preferred. Different from the time series analysis, the spatial trend model does not aim at capturing the whole variability of the considered feature but rather attempts to ascribe it to trend. Hence, the variation of the considered feature can not be simulated completely by the spatial trend model and some parts should be interpreted by other spatial factors. In this sense, the spatial trend model becomes the combined model (here called the Improved Spatial Trend model (IST model)) and the OLS can be employed again. Formally, the IST model can be written

$$y = A\theta + f(X, \beta) + \varepsilon$$ (5-1)

where $f(X, \beta)$ stands for the other model which functions upon $X$ and $\beta$. $X$ is the matrix of non-location spatial variables (i.e. non-location spatial factors), excluding the locations. $\beta$ is the associated coefficient vector.

Because the non-linear function can be transformed into liner polynomial function, without losing generality we set $f(X, \beta) = X\beta$ and the OLS estimators for $\theta$ and $\beta$ hence are

$$\begin{pmatrix} \theta \\ \beta \end{pmatrix} = \begin{pmatrix} (A'X)(A'X)^{-1} & (A'X) \end{pmatrix} y$$ (5-2)

The advantages of the IST model are as follows:

1) it can solve the ill-conditioned least squares problem while the residuals are not
spatially autocorrelated,
2) the estimation for its parameters is simple.

The IST model can be applied to identify the spatial influencing factors of renewable energy consumption, especially the location factor. In general, the considered spatial influencing factors of renewable energy consumption are location factor, economic factors such as population, income and energy price (Lambert D’Apote 1998; Victor and Victor 2002; LEAP 2002) and/or GDP (Nakićenović et al. 1998; IEA 2002; World Bank 2002), and any other spatial factors. These factors excluding location factor can be regarded as the non-location spatial factor $X$ of IST model and location factor as the $A$ of IST model. Then the IST model will assess whether all these factors or some of them will influence renewable energy consumption through regression.

5.2 Spatial statistical model selection procedure

The main aim of modeling the natural phenomenon is to capture the most main impaction factors as far as possible. Therefore, if the modification to the model can provide more information, the modification is appropriate. If the modification to the model, however, can not provide information any more, the modification is not necessary and the present model is better than the modified model. So the definition of “information” is pretty important and becomes the essence of the model selection criterion. The philosophy of the $R^2$ or $R_a^2$ is to evaluate how much the variation of dependent variable is explained by a series of explanatory variables while the philosophy of the AIC or SIC is to evaluate how much information the model does not capture.

The spatial statistical models also try to capture the main factors influencing the natural phenomenon. Hence, the $R^2$/$R_a^2$ or AIC/SIC can also be used to select the spatial statistical models. However, the spatial statistical models have their own character: the interesting feature often demonstrates the spatial autocorrelation. Ignoring the spatial autocorrelation will make the result invalid (Anselin and Griffith 1988). As a consequence, when developing the spatial statistical models, we should test the spatial autocorrelation among errors. Geary (1954) and Cliff and Ord (1981) argued that for spatial data the spatial autocorrelation for residuals should be taken into account in selecting the spatial statistical models. In fact, model selection is an extremely complicated thing. Lots of things must be considered. In this dissertation, we present a simple and easily
implemented spatial statistical model selection procedure to select a spatial statistical model. The procedure is as follows: the first step is to employ Moran’s I test to test the spatial autocorrelation among residuals resulting from the spatial statistical models. If the spatial autocorrelation does exist among residuals, it suggests that some information is not captured by the spatial statistical model. However, the information can be hopefully captured if we modify the model specification. Then we modify the model specification and use Moran’s I test to test the spatial autocorrelation among residuals again. The preceding steps can be repeated until the spatial autocorrelation among residuals disappears. However, if the residuals resulting from the spatial statistical models are not spatially autocorrelated, then we have already captured most information of the underlying phenomenon and thus we have no necessity to modify the model specifications. After that, we take use of the $R^2$, $R_s^2$, AIC and SIC to select the best spatial statistical model. This step is the same with the classical statistical model selection procedure presented in time series. Fig 5-1 illustrates the whole procedure.

![Diagram](image)

Fig 5-1: Spatial statistical model selection procedure.

The spatial statistical model selection procedure can be applied to select the best IST model for renewable energy consumption because the IST model is also the spatial statistical model.
6 Case Studies

Because of the poor availability of data sets, we can not obtain the needed input data sets which are the same kind of renewable energy of the same region for the two models developed in this dissertation. Nevertheless, we have being obtained the needed input data sets respectively for the OFLR model and the IST model. These input data sets are enough to demonstrate how these models deal with the reality renewable energy issues and present the scientific policy suggestions for renewable energy planning.

The study area of the case study of OFLR model consists of five European countries and the studied renewable energy is solar energy. This case study aims at mainly determining the optimal locations and optimal capacity sizes of photovoltaics (one kind of solar energy facility) in five European countries via OFLR model because photovoltaics is recently advocated (IEA-PVPS 2007). The study area of the case study of IST model is the U.S. which is composed of 50 states and a federal district and the studied renewable energy is biomass. A comparable set for a large area was not available in Europe and it was not available for solar and wind energy neither in Europe nor U.S.. This case study aims to identify the spatial influencing factors of biomass consumption in U.S. via IST model.

6.1 Case study of OFLR model

6.1.1 Materials

6.1.1.1 Study region

In this case study, the study region consists of five European countries, which are Austria, France, Germany, Italy, and Switzerland respectively. Austria is a landlocked country of roughly 8.3 million people in Central Europe. It borders both Germany and the Czech Republic to the north, Slovakia and Hungary to the east, Slovenia and Italy to the south, and Switzerland and Liechtenstein to the west. The territory of Austria covers 83,872 square kilometers and is influenced by a temperate and alpine climate. Austria's terrain is high mountainous due to the presence of the Alps: only 32% of the country is below 500 meters, and its highest point is 3,797 meters. Austria is one of the richest countries in the world, with a nominal per capita GDP of $43,570. The country has developed a high standard of living and in 2008 was ranked 14th in the world for its
human development index (Wikipedia 2010). France is a developed country and possesses the fifth largest economy by nominal GDP and eighth largest economy by purchasing power parity. The country locates in Western Europe, with several overseas islands and territories located on other continents. France is bordered (clockwise from the north) by Belgium, Luxembourg, Germany, Switzerland, Italy, Monaco, Andorra, and Spain and also share land borders with Brazil and Suriname, and the Netherlands Antilles. France is linked to the United Kingdom by the Channel Tunnel, which passes underneath the English Channel. France is the largest country in the European Union by area (Wikipedia 2010). Germany is a developed country and possesses the world's fourth largest economy by nominal GDP and the fifth largest in purchasing power parity. Germany is recognized as a scientific and technological leader in several fields. The country locates in Central Europe. It is bordered to the north by the North Sea, Denmark, and the Baltic Sea, to the east by Poland and the Czech Republic, to the south by Austria and Switzerland, and to the west by France, Luxembourg, Belgium, and the Netherlands. The territory of Germany covers 357,021 square kilometers and is influenced by a temperate seasonal climate (Wikipedia 2010). Switzerland is one of the richest countries in the world by per capita gross domestic product, with a nominal per capita GDP of $67,384. The country situates in Central Europe where it is bordered by Germany to the north, by France to the west, by Italy to the south, and by Austria and Liechtenstein to the east. The territory of Switzerland is geographically divided between the Jura, the Central Plateau and the Alps. Switzerland is a landlocked mountain country with total area of 41,285 km² (Wikipedia 2010). Italy is a developed country with the eighth-highest quality of life index rating in the world. The country locates on the Italian Peninsula in Southern Europe and on the two largest islands in the Mediterranean Sea, Sicily and Sardinia. Italy shares its northern, Alpine boundary with France, Switzerland, Austria and Slovenia. Its total area is 301,230 km², of which 294,020 km² is land and 7,210 km² is water. Including islands, Italy has a coastline and border of 7,600 km on the Adriatic, Ionian, Tyrrhenian seas (Wikipedia 2010). The spatial layout of these five European countries is depicted in Fig 6-1.
Among the five studied countries, Germany has invested lots of money in renewable energy technologies. According to IEA statistic, Germany has the third largest installed photovoltaics in the world and has the largest installed photovoltaics in Europe (IEA 2006). In recent years, the five countries pay growing attention on photovoltaics and have taken use of photovoltaics to provide electricity. Table 6-1 documents the installed photovoltaics for five studied countries in 2003.

Table 6-1. Installed capacity of photovoltaics in 2003 (IEA 2006)

<table>
<thead>
<tr>
<th>Country</th>
<th>Capacity (MW)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Austria</td>
<td>16.8</td>
</tr>
<tr>
<td>Switzerland</td>
<td>21.0</td>
</tr>
<tr>
<td>France</td>
<td>21.1</td>
</tr>
<tr>
<td>Italy</td>
<td>26.0</td>
</tr>
<tr>
<td>Germany</td>
<td>431.0</td>
</tr>
</tbody>
</table>

6.1.1.2 Data sources

The complete data sets and sources are listed by Table 6-2. The solar direct normal irradiation data is acquired from NASA (NASA 2008). This data is the latitude-longitude grid data with the spatial resolution of 1×1 degrees and is global level with the coordinate system of WGS_84. The data is 22-year Monthly & Annual Average (July 1983 - June 2005) and is in kWh/m²/day. Fig 6-2 depicts the spatial distribution of solar direct
normal irradiation of five studied countries which is obtained through the overlap function of ArcGIS 9.3 to the global solar direct normal irradiation data.

Table 6-2. Data list of case study of OFLR model

<table>
<thead>
<tr>
<th>Data</th>
<th>Data source</th>
</tr>
</thead>
<tbody>
<tr>
<td>Solar direct normal irradiation data</td>
<td>NASA</td>
</tr>
<tr>
<td>Landuse data</td>
<td>ESA Globcover</td>
</tr>
<tr>
<td>Land slope data</td>
<td>SRTM</td>
</tr>
<tr>
<td>Photovoltaics module price</td>
<td>IEA</td>
</tr>
<tr>
<td>Efficiency of photovoltaics</td>
<td>15%</td>
</tr>
<tr>
<td>Production cost of photovoltaics</td>
<td>1% of photovoltaics module price</td>
</tr>
<tr>
<td>Lifetime of photovoltaics</td>
<td>25 years</td>
</tr>
<tr>
<td>Cost of HVDC</td>
<td>See Table 4-3 (DLR 2006)</td>
</tr>
<tr>
<td>Transportation cost of HVDC</td>
<td>See Table 4-3 (DLR 2006)</td>
</tr>
<tr>
<td>Loss of HVDC</td>
<td>See Table 4-3 (DLR 2006)</td>
</tr>
<tr>
<td>Electricity demand</td>
<td>IEA</td>
</tr>
</tbody>
</table>

The landuse data is acquired from European Space Agency (ESA) and the ESA Globcover project led by MEDIAS-France Corporation, pertaining to 2006. The spatial resolution of
this landuse data is 300×300 m² (or 1/360×1/360 degrees) with the coordinate system of WGS_84. As the landuse data is global level, we use the spatial analysis of ArcGIS 9.3 to cut this global landuse data to obtain the landuse data for the considered area. The resulting map is illustrated by Fig 6-3.

The values of the suitable factors for land categories in most cases rely upon the experience. There is no agreement about how to determine the appropriate values for the suitable factors of land categories. In this dissertation, the following suitable factors for the preceding landuse data are adopted and are documented by Table 6-3.
Table 6-3. Suitable factors of land cover

<table>
<thead>
<tr>
<th>Land cover type</th>
<th>Suitable factor</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cultivated Terrestrial Areas and Managed Lands</td>
<td>0.00</td>
</tr>
<tr>
<td>Natural and Semi-natural Terrestrial Vegetation - Woody / Trees</td>
<td>0.01</td>
</tr>
<tr>
<td>Natural and Semi-natural Terrestrial Vegetation - Shrubs</td>
<td>0.01</td>
</tr>
<tr>
<td>Natural and Semi-natural Terrestrial Vegetation - Herbaceous</td>
<td>0.01</td>
</tr>
<tr>
<td>Natural and Semi-natural Terrestrial Vegetation</td>
<td>0.01</td>
</tr>
<tr>
<td>Natural and Seminatural Aquatic Vegetation</td>
<td>0.05</td>
</tr>
<tr>
<td>Artificial Surfaces and Associated Areas</td>
<td>0.05</td>
</tr>
<tr>
<td>Inland Waterbodies</td>
<td>0.00</td>
</tr>
<tr>
<td>Permanent Snow and Ice</td>
<td>0.00</td>
</tr>
</tbody>
</table>

The land slope data is acquired from IIASA. It belongs to Shuttle Radar Topography Mission (SRTM) project. The data is the global data with the spatial resolution of 0.5x0.5 degrees. The coordinate system for the data is WGS_84. Since the data is global data, we use the spatial analysis of ArcGIS 9.3 to cut the global land slope data in order to obtain the land slope data for the studied area. The resulting map is depicted in Fig 6-4.

![Land slope of five studied countries.](image)

The photovoltaics module price data is acquired from IEA (IEA-PVPS 2007), pertaining
to 2006. In this dissertation we simply refer to the photovoltaics module price as the photovoltaics price. Table 6-4 details the data. The efficiency of photovoltaics is set to 15% and the lifetime of photovoltaics is 25 years. The production cost of photovoltaics (here refer to the sum of the variable operation&maintenance cost and fix operation&maintenance cost) is 1% of photovoltaics price.

Table 6-4. Photovoltaics prices of five European countries, 2006

<table>
<thead>
<tr>
<th>Country</th>
<th>Price($/W)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Austria</td>
<td>6.45</td>
</tr>
<tr>
<td>France</td>
<td>6.30</td>
</tr>
<tr>
<td>Germany</td>
<td>5.50</td>
</tr>
<tr>
<td>Italy</td>
<td>6.90</td>
</tr>
<tr>
<td>Switzerland</td>
<td>7.20</td>
</tr>
</tbody>
</table>

The cost of HVDC, the loss of HVDC and the transportation cost of HVDC (i.e. operation&maintenance cost) are taken from Table 4-3(DLR 2006).

The electricity demand data recorded by Table 6-5 is acquired from IEA (IEA 2009), pertaining to 2006. The demanded electricity refers to electricity which comes from photovoltaics and is the total demanded electricity of country rather than the sector demand of economy department of country. We assume that the electricity demand of one country is concentrated on one grid of this country.

Table 6-5. Electricity demands of five European countries, 2006

<table>
<thead>
<tr>
<th>Country</th>
<th>Electricity(GWh/a)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Austria</td>
<td>15.0000</td>
</tr>
<tr>
<td>France</td>
<td>22.0000</td>
</tr>
<tr>
<td>Germany</td>
<td>2220.0000</td>
</tr>
<tr>
<td>Italy</td>
<td>35.0000</td>
</tr>
<tr>
<td>Switzerland</td>
<td>23.0000</td>
</tr>
</tbody>
</table>

### 6.1.2 Experimental design

The experiment aims at mainly analyzing solar energy via the OFLR model to provide the geographically explicit renewable energy planning strategy to renewable energy planning policy decision makers. The considered supply chain is composed of solar light, photovoltaics, electricity, HVDC and demand regions, described by Fig 6-5. In this supply chain, photovoltaics will collect solar light and convert it into electricity. Then HVDC will transport electricity to demand regions. In order to apply the OFLR model to the experiment, we should first grid the studied area. It is considerably important to
determine the appropriate spatial resolution for the grid to cover the studied area. The
determination to appropriate spatial resolution should take into account not only the
characteristics of the studied problem but also the spatial resolutions of all available data
sets.

![Diagram of Electricity supply chain of solar energy with photovoltaics.](image)

Scenario setting is greatly important for renewable energy models. In this experiment, we
take into account the different scenarios for analyzing the solar energy supply chain. We
first choose 5 demand grid regions for the studied area in which each country of the
studied area has one demand region. We also determine the possible photovoltaics
locations for the studied area with respect to the data and try to averagely assign the
possible photovoltaics locations to each country. We do not take into account the
by-products of photovoltaics and further assume that no any electricity imports and
exports between the demand regions and the other non-studied areas will occur. Then we
make the following four scenarios with respect to the situation widely encountered in
reality:

**Scenario I:**

The theoretical potential of solar energy is considered.

**Scenario II:**

The geographical potential of solar energy is considered.

**Scenario III:**

The geographical potential of solar energy is considered and one of the optimal
photovoltaics determined in scenario II is shut down (that is, the spatial distribution of
photovoltaics has been changed).

**Scenario IV:**

The geographical potential of solar energy is considered and one of the optimal
HVDC determined in scenario II is shut down (that is, the spatial distribution of HVDC
has been changed).

Scenarios I and II focus on investigating the influence of solar potential on the supply
chain. Scenario III focuses on investigating the influence of spatial distribution of
photovoltaics on the supply chain. Scenario IV focuses on investigating the influence of
spatial distribution of HVDC on the supply chain.
In order to perform these four scenarios, we should gather the appropriate data. Because of the unavailability of ozone data, gas data and water data, we only obtain the solar direct normal irradiation data. In order to calculate the theoretical potential of solar energy for regions, we multiply the solar direct normal irradiation by the total area of the regions. We acquire the landuse data. With respect to the landuse data we can calculate the available area through introducing the suitable factors. Then we can calculate the geographical potential of solar energy. We acquire the land slope data to impose the environmental constraint on the locations of photovoltaics. We can acquire the photovoltaics price data and the demanded electricity data. We only consider HVDC to transport electricity in this experiment and assume that HVDC consists of two stations and the overhead lines regardless of the geographical situation of transportation route. Then we can acquire the price data of the line and stations of HVDC. After processing all gathered data into the standard units and format, we input these data into the OFLR model. The model year is 2006. Fig 6-6 illustrates the workflow of experiment.
6.1.3 Results

We have to first determine the spatial resolution of grid for studied area with respect to the factors such as the landuse, the land slope, the solar irradiation and the reality requirements, etc. This is very important for the OFLR model. Since we only obtain the landuse data, the land slope data and the solar irradiation data, the spatial resolution of grid of studied area will mainly rely upon the spatial resolutions of these data sets. The spatial resolutions of the landuse data, the land slope data and the solar irradiation data are respectively $1/360 \times 1/360$ degrees, $0.5 \times 0.5$ degrees and $1 \times 1$ degrees. With respect to these spatial resolutions, we choose the spatial resolution of $0.5 \times 0.5$ degrees as the spatial resolution of grid of studied area. Of course, we can choose other spatial resolutions. However, the spatial resolution of $0.5 \times 0.5$ degrees is chosen because we will make use of the land slope data to determine the potential locations of photovoltaics such that we can simplify the data processing and decrease the running time of the OFLR model. The grids of studied area are showed by Fig 6-7.

Fig 6-7: Studied area grids with spatial resolution of $0.5 \times 0.5$ degrees.
Because the spatial resolution of the solar direct normal irradiation data is coarser than the spatial resolution of grid, we have to interpolate the solar direct normal irradiation data to obtain the proper spatial resolution. We use the ordinary Kriging to interpolate the data (Wang et al. 2007). In order to better interpolate the data and generate the raster data which matches the grid, we use the global solar direct normal irradiation data rather than the preceding five studied countries’ solar direct normal irradiation data obtained by overlay function of ArcGIS 9.3. Then we use the mask of ArcGIS 9.3 to extract the solar direct normal irradiation for studied area. The resulting map is illustrated by Fig 6-8.

In order to know clearly the land slope for each grid, we overlay the grids of studied area on the global land slope data to obtain the land slope for these grids which is illustrated by Fig 6-9.
Now we turn to determine the potential locations for photovoltaics. To determine the potential locations for photovoltaics, many factors should be considered. But the most importance is that the locations can be used to set up photovoltaics. The land slope is one of the most important factors which are used to determine the potential locations for photovoltaics. Normally, the land slope for photovoltaics is lower or equal to 5 degree. However, for this land slope data, because a lot of place has no data, the determination of the potential locations for photovoltaics must be careful. For this land slope data set, we adopt the land slope with lowering or equaling to 6 degree as the accepted land slope.
where photovoltaics can be set up. In order to decrease the running time of the OFLR model, we explicitly determine the potential locations of photovoltaics with respect to the land slope data instead imposing the land slope as a remote sensing-based constraint of the OFLR model.

On the basis of Fig 6-9, we can determine the potential locations of photovoltaics and can also determine the demand regions for studied area which are almost the central grids of the five studied countries. The potential locations of photovoltaics and the demand regions are depicted in Fig 6-10. According to Fig 6-10, the total number of the potential locations is 52 among which Italy owns 13, France owns 22 and Germany owns 17. Then the electricity demands of the five studied countries are projected to the corresponding demand regions. Table 6-6 shows the locations of demand regions and the corresponding electricity demands.

<table>
<thead>
<tr>
<th>Name</th>
<th>Longitude(°)</th>
<th>Latitude(°)</th>
<th>Country</th>
<th>Electricity demand(GWh/a)</th>
</tr>
</thead>
<tbody>
<tr>
<td>D1</td>
<td>44.25</td>
<td>11.25</td>
<td>Italy</td>
<td>35.0000</td>
</tr>
<tr>
<td>D2</td>
<td>46.75</td>
<td>2.25</td>
<td>France</td>
<td>22.0000</td>
</tr>
<tr>
<td>D3</td>
<td>46.75</td>
<td>8.25</td>
<td>Switzerland</td>
<td>23.0000</td>
</tr>
<tr>
<td>D4</td>
<td>47.75</td>
<td>14.75</td>
<td>Austria</td>
<td>15.0000</td>
</tr>
<tr>
<td>D5</td>
<td>51.25</td>
<td>10.75</td>
<td>Germany</td>
<td>2220.0000</td>
</tr>
</tbody>
</table>

In order to calculate the solar potentials for grids, including the theoretical potential and the geographic potential, we only multiply the solar direct normal irradiation by the total area of grid for the theoretical potential of solar energy and for the geographical potential of solar energy the calculation is a little difficult. We should first calculate the total area for each land category of grid and then multiply the total area of each land category of grid by the suitable factor of the land category to get the available area for each land category of grid. Summing all the available areas of land categories of grid, we can obtain the available area for grid. Then we multiply the available area of grid by the solar direct normal irradiation of grid to obtain the geographic potential of solar energy for grid. Table 6-7 details the longitude and latitude of the potential photovoltaics locations, the countries to which the potential locations belong, the theoretical potential of solar energy of the potential locations, and the geographical potential of solar energy of the potential locations.
Table 6-7. Potential photovoltaics locations and solar potential

<table>
<thead>
<tr>
<th>Name</th>
<th>Longitude (°)</th>
<th>Latitude (°)</th>
<th>Country</th>
<th>Theoretical solar potential (GWh/a)</th>
<th>Geographical solar potential (GWh/a)</th>
</tr>
</thead>
<tbody>
<tr>
<td>S1</td>
<td>40.75</td>
<td>15.25</td>
<td>Italy</td>
<td>4388987.636</td>
<td>24851.555</td>
</tr>
<tr>
<td>S2</td>
<td>40.75</td>
<td>15.75</td>
<td>Italy</td>
<td>4447668.308</td>
<td>22291.213</td>
</tr>
<tr>
<td>S3</td>
<td>41.25</td>
<td>14.75</td>
<td>Italy</td>
<td>4280156.840</td>
<td>21406.042</td>
</tr>
<tr>
<td>S4</td>
<td>41.25</td>
<td>15.25</td>
<td>Italy</td>
<td>4403639.506</td>
<td>21218.063</td>
</tr>
<tr>
<td>S5</td>
<td>41.75</td>
<td>14.75</td>
<td>Italy</td>
<td>4279911.516</td>
<td>21621.033</td>
</tr>
<tr>
<td>S6</td>
<td>42.25</td>
<td>14.25</td>
<td>Italy</td>
<td>4088075.987</td>
<td>24599.190</td>
</tr>
<tr>
<td>S7</td>
<td>42.75</td>
<td>11.75</td>
<td>Italy</td>
<td>4128825.237</td>
<td>21943.088</td>
</tr>
<tr>
<td>S8</td>
<td>42.75</td>
<td>12.25</td>
<td>Italy</td>
<td>4058463.129</td>
<td>23579.290</td>
</tr>
<tr>
<td>S9</td>
<td>43.25</td>
<td>10.75</td>
<td>Italy</td>
<td>4080987.635</td>
<td>34175.482</td>
</tr>
<tr>
<td>S10</td>
<td>43.25</td>
<td>11.25</td>
<td>Italy</td>
<td>3997499.887</td>
<td>34944.799</td>
</tr>
<tr>
<td>S11</td>
<td>43.25</td>
<td>13.25</td>
<td>Italy</td>
<td>3969168.236</td>
<td>18178.236</td>
</tr>
<tr>
<td>S12</td>
<td>43.75</td>
<td>11.25</td>
<td>Italy</td>
<td>3943228.193</td>
<td>53639.086</td>
</tr>
<tr>
<td>Case Study</td>
<td>Year</td>
<td>month</td>
<td>Country</td>
<td>Price</td>
<td>Efficiency</td>
</tr>
<tr>
<td>------------</td>
<td>------</td>
<td>-------</td>
<td>---------</td>
<td>--------</td>
<td>------------</td>
</tr>
<tr>
<td>S13</td>
<td>43.75</td>
<td>12.75</td>
<td>Italy</td>
<td>3958293.433</td>
<td>21364.054</td>
</tr>
<tr>
<td>S14</td>
<td>44.25</td>
<td>2.25</td>
<td>France</td>
<td>3317546.966</td>
<td>34523.013</td>
</tr>
<tr>
<td>S15</td>
<td>44.25</td>
<td>2.75</td>
<td>France</td>
<td>3357052.295</td>
<td>34699.121</td>
</tr>
<tr>
<td>S16</td>
<td>44.25</td>
<td>3.25</td>
<td>France</td>
<td>3464585.507</td>
<td>27259.713</td>
</tr>
<tr>
<td>S17</td>
<td>44.75</td>
<td>1.25</td>
<td>France</td>
<td>3189533.911</td>
<td>35371.605</td>
</tr>
<tr>
<td>S18</td>
<td>44.75</td>
<td>1.75</td>
<td>France</td>
<td>3167078.970</td>
<td>34350.067</td>
</tr>
<tr>
<td>S19</td>
<td>44.75</td>
<td>2.25</td>
<td>France</td>
<td>3199845.858</td>
<td>39251.525</td>
</tr>
<tr>
<td>S20</td>
<td>44.75</td>
<td>2.75</td>
<td>France</td>
<td>3217489.317</td>
<td>39221.075</td>
</tr>
<tr>
<td>S21</td>
<td>44.75</td>
<td>3.25</td>
<td>France</td>
<td>3271679.709</td>
<td>37469.068</td>
</tr>
<tr>
<td>S22</td>
<td>44.75</td>
<td>3.75</td>
<td>France</td>
<td>334062.229</td>
<td>38187.546</td>
</tr>
<tr>
<td>S23</td>
<td>45.25</td>
<td>0.75</td>
<td>France</td>
<td>3087932.024</td>
<td>30738.322</td>
</tr>
<tr>
<td>S24</td>
<td>45.25</td>
<td>1.25</td>
<td>France</td>
<td>3032843.241</td>
<td>35123.803</td>
</tr>
<tr>
<td>S25</td>
<td>45.25</td>
<td>1.75</td>
<td>France</td>
<td>3030566.765</td>
<td>39573.009</td>
</tr>
<tr>
<td>S26</td>
<td>45.25</td>
<td>2.25</td>
<td>France</td>
<td>3024306.253</td>
<td>38602.420</td>
</tr>
<tr>
<td>S27</td>
<td>45.25</td>
<td>2.75</td>
<td>France</td>
<td>3037452.762</td>
<td>37215.710</td>
</tr>
<tr>
<td>S28</td>
<td>45.25</td>
<td>3.25</td>
<td>France</td>
<td>3063744.970</td>
<td>34267.407</td>
</tr>
<tr>
<td>S29</td>
<td>45.25</td>
<td>3.75</td>
<td>France</td>
<td>3129475.896</td>
<td>36525.332</td>
</tr>
<tr>
<td>S30</td>
<td>45.25</td>
<td>4.25</td>
<td>France</td>
<td>3191450.677</td>
<td>36777.736</td>
</tr>
<tr>
<td>S31</td>
<td>45.75</td>
<td>1.25</td>
<td>France</td>
<td>2888666.035</td>
<td>38522.658</td>
</tr>
<tr>
<td>S32</td>
<td>45.75</td>
<td>1.75</td>
<td>France</td>
<td>2860164.309</td>
<td>36997.027</td>
</tr>
<tr>
<td>S33</td>
<td>45.75</td>
<td>2.25</td>
<td>France</td>
<td>2881979.720</td>
<td>37178.025</td>
</tr>
<tr>
<td>S34</td>
<td>45.75</td>
<td>2.75</td>
<td>France</td>
<td>2870761.032</td>
<td>36534.008</td>
</tr>
<tr>
<td>S35</td>
<td>45.75</td>
<td>3.75</td>
<td>France</td>
<td>2911954.689</td>
<td>35818.258</td>
</tr>
<tr>
<td>S36</td>
<td>48.75</td>
<td>13.25</td>
<td>Germany</td>
<td>2246788.027</td>
<td>24907.961</td>
</tr>
<tr>
<td>S37</td>
<td>48.75</td>
<td>13.75</td>
<td>Germany</td>
<td>2238128.836</td>
<td>29284.484</td>
</tr>
<tr>
<td>S38</td>
<td>49.25</td>
<td>7.25</td>
<td>Germany</td>
<td>2181709.048</td>
<td>29460.369</td>
</tr>
<tr>
<td>S39</td>
<td>49.25</td>
<td>12.75</td>
<td>Germany</td>
<td>2169303.376</td>
<td>25487.029</td>
</tr>
<tr>
<td>S40</td>
<td>49.25</td>
<td>13.25</td>
<td>Germany</td>
<td>2169303.376</td>
<td>24911.996</td>
</tr>
<tr>
<td>S41</td>
<td>49.75</td>
<td>6.75</td>
<td>Germany</td>
<td>2082833.202</td>
<td>25810.768</td>
</tr>
<tr>
<td>S42</td>
<td>49.75</td>
<td>7.25</td>
<td>Germany</td>
<td>2091848.359</td>
<td>24909.217</td>
</tr>
<tr>
<td>S43</td>
<td>49.75</td>
<td>7.75</td>
<td>Germany</td>
<td>2076932.231</td>
<td>20827.027</td>
</tr>
<tr>
<td>S44</td>
<td>50.25</td>
<td>6.75</td>
<td>Germany</td>
<td>2015182.308</td>
<td>25710.947</td>
</tr>
<tr>
<td>S45</td>
<td>50.25</td>
<td>7.75</td>
<td>Germany</td>
<td>2002776.935</td>
<td>25836.609</td>
</tr>
<tr>
<td>S46</td>
<td>50.25</td>
<td>9.75</td>
<td>Germany</td>
<td>1986795.190</td>
<td>23843.688</td>
</tr>
<tr>
<td>S47</td>
<td>50.25</td>
<td>11.75</td>
<td>Germany</td>
<td>2003372.356</td>
<td>20046.648</td>
</tr>
<tr>
<td>S48</td>
<td>50.75</td>
<td>7.75</td>
<td>Germany</td>
<td>1930611.220</td>
<td>26535.274</td>
</tr>
<tr>
<td>S49</td>
<td>50.75</td>
<td>8.25</td>
<td>Germany</td>
<td>1948389.417</td>
<td>27797.866</td>
</tr>
<tr>
<td>S50</td>
<td>51.25</td>
<td>7.75</td>
<td>Germany</td>
<td>1890235.000</td>
<td>29102.713</td>
</tr>
<tr>
<td>S51</td>
<td>51.25</td>
<td>8.25</td>
<td>Germany</td>
<td>1885521.203</td>
<td>25596.291</td>
</tr>
<tr>
<td>S52</td>
<td>51.25</td>
<td>9.75</td>
<td>Germany</td>
<td>1870683.048</td>
<td>21698.815</td>
</tr>
</tbody>
</table>

We assume that the photovoltaics price, the efficiency of photovoltaics, the lifetime of
photovoltaics and the production cost of photovoltaics for each grid are the same with those of country to which the grid belongs. In order to calculate the losses and line cost of HVDC, we first figure out the straight distances between the potential locations of photovoltaics and the demand regions. Then we assume that the real distance is the straight distance by 1.4. After calculating the distances, we can obtain the losses and line cost of HVDC.

Now it is time to perform the four scenarios. To perform the four scenarios we should first standard all the data sets and then input the standard data sets into the OFLR model. Table 6-8 shows the optimal locations and optimal capacity sizes of photovoltaics for scenarios I and II. According to Table 6-8, five optimal locations are chosen for photovoltaics and five optimal capacity sizes for photovoltaics are determined. These five optimal locations all fall into Germany.

Table 6-8. Optimal locations and capacity sizes of photovoltaics, Scenarios I and II

<table>
<thead>
<tr>
<th>Name</th>
<th>Year</th>
<th>Size(GW)</th>
<th>Annual production(GWh/a)</th>
</tr>
</thead>
<tbody>
<tr>
<td>S36</td>
<td>2006</td>
<td>0.0276</td>
<td>36.2740</td>
</tr>
<tr>
<td>S37</td>
<td>2006</td>
<td>0.0117</td>
<td>15.3531</td>
</tr>
<tr>
<td>S38</td>
<td>2006</td>
<td>0.0180</td>
<td>23.6605</td>
</tr>
<tr>
<td>S41</td>
<td>2006</td>
<td>0.0173</td>
<td>22.7956</td>
</tr>
<tr>
<td>S52</td>
<td>2006</td>
<td>1.7254</td>
<td>2267.1884</td>
</tr>
</tbody>
</table>

Table 6-9 shows the optimal capacity sizes of HVDC and the annual transported electricity by HVDC for scenarios I and II. In Table 6-9, the item “From” represents the supply regions of electricity and the item “To” represents the demand regions of electricity. So the total meanings of “From” and “To” are that electricity is transported by HVDC from supply regions to demand regions. In the following tables, the items of “From” and “To” have the same meanings with here. The annual transported electricity is equal to the optimal capacity size of HVDC according to Table 6-9 if a conversion factor of unit is considered.

Table 6-9. Optimal capacity sizes of HVDC, Scenarios I and II

<table>
<thead>
<tr>
<th>Name</th>
<th>From</th>
<th>To</th>
<th>Year</th>
<th>Size(GW)</th>
<th>Annual transported electricity(GWh/a)</th>
</tr>
</thead>
<tbody>
<tr>
<td>HVDC</td>
<td>S36</td>
<td>D1</td>
<td>2006</td>
<td>0.0041</td>
<td>36.2740</td>
</tr>
<tr>
<td>HVDC</td>
<td>S37</td>
<td>D4</td>
<td>2006</td>
<td>0.0018</td>
<td>15.3531</td>
</tr>
<tr>
<td>HVDC</td>
<td>S38</td>
<td>D3</td>
<td>2006</td>
<td>0.0027</td>
<td>23.6605</td>
</tr>
<tr>
<td>HVDC</td>
<td>S41</td>
<td>D2</td>
<td>2006</td>
<td>0.0026</td>
<td>22.7956</td>
</tr>
<tr>
<td>HVDC</td>
<td>S52</td>
<td>D5</td>
<td>2006</td>
<td>0.2588</td>
<td>2267.1884</td>
</tr>
</tbody>
</table>
Fig 6-11 depicts the optimal locations of photovoltaics and HVDC as well as the electricity transportation direction with HVDC for scenarios I and II. The blue line represents HVDC and the arrow direction represents the electricity transportation direction. In terms of Fig 6-11, one optimal location of photovoltaics will supply electricity via HVDC to one demand region.

Table 6-10 shows the electricity prices of demand regions for scenarios I and II. On the basis of Table 6-10, we can evidently know that the electricity price of demand region D5 (which lies within Germany) is the lowest whereas the electricity price of demand D1 (which lies within Italy) is the highest. The electricity prices are decreasing with the order of D1, D2, D3, D4 and D5.
Table 6-10. Electricity prices of demand regions, Scenarios I and II

<table>
<thead>
<tr>
<th>Demand region</th>
<th>Year</th>
<th>PRICE($/GWh)</th>
</tr>
</thead>
<tbody>
<tr>
<td>D1</td>
<td>2006</td>
<td>5.9776E+6</td>
</tr>
<tr>
<td>D2</td>
<td>2006</td>
<td>5.9754E+6</td>
</tr>
<tr>
<td>D3</td>
<td>2006</td>
<td>5.9086E+6</td>
</tr>
<tr>
<td>D4</td>
<td>2006</td>
<td>5.8622E+6</td>
</tr>
<tr>
<td>D5</td>
<td>2006</td>
<td>5.8416E+6</td>
</tr>
</tbody>
</table>

In order to perform the scenario III, we randomly choose one of the optimal locations of photovoltaics determined in scenario II and shut down photovoltaics in this location. Then we run the model again for scenario III. Table 6-11 shows the optimal locations and optimal capacity sizes of photovoltaics for scenario III. Four optimal locations of photovoltaics and four optimal capacity sizes of photovoltaics are determined according to Table 6-11. These four optimal locations lie within Germany.

Table 6-11. Optimal locations and capacity sizes of photovoltaics, Scenario III

<table>
<thead>
<tr>
<th>Name</th>
<th>Year</th>
<th>Size(GW)</th>
<th>Annual production(GWh/a)</th>
</tr>
</thead>
<tbody>
<tr>
<td>S37</td>
<td>2006</td>
<td>0.0393</td>
<td>51.6419</td>
</tr>
<tr>
<td>S38</td>
<td>2006</td>
<td>0.0180</td>
<td>23.6605</td>
</tr>
<tr>
<td>S41</td>
<td>2006</td>
<td>0.0173</td>
<td>22.7956</td>
</tr>
<tr>
<td>S52</td>
<td>2006</td>
<td>1.7254</td>
<td>2267.1884</td>
</tr>
</tbody>
</table>

Table 6-12 shows the optimal capacity sizes of HVDC and the annual transported electricity by HVDC for scenario III. The optimal capacity sizes of HVDC are equal to the annual transported electricity if a conversion factor of unit is considered and there are two HVDC to transport electricity from S37 respectively to D1 and D4.

Table 6-12. Optimal capacity sizes of HVDC, Scenario III

<table>
<thead>
<tr>
<th>Name</th>
<th>From</th>
<th>To</th>
<th>Year</th>
<th>Size(GW)</th>
<th>Annual transported electricity(GWh/a)</th>
</tr>
</thead>
<tbody>
<tr>
<td>HVDC</td>
<td>S37</td>
<td>D1</td>
<td>2006</td>
<td>0.0041</td>
<td>36.2887</td>
</tr>
<tr>
<td>HVDC</td>
<td>S37</td>
<td>D4</td>
<td>2006</td>
<td>0.0018</td>
<td>15.3531</td>
</tr>
<tr>
<td>HVDC</td>
<td>S38</td>
<td>D3</td>
<td>2006</td>
<td>0.0027</td>
<td>23.6605</td>
</tr>
<tr>
<td>HVDC</td>
<td>S41</td>
<td>D2</td>
<td>2006</td>
<td>0.0026</td>
<td>22.7956</td>
</tr>
<tr>
<td>HVDC</td>
<td>S52</td>
<td>D5</td>
<td>2006</td>
<td>0.2588</td>
<td>2267.1884</td>
</tr>
</tbody>
</table>

Fig 6-12 depicts the optimal locations of photovoltaics and HVDC as well as the electricity transportation direction with HVDC for scenario III. The blue line represents HVDC and the arrow direction represents the electricity transportation direction. Based on Fig 6-12, supply region S37 will supply electricity respectively to demand regions D1 and D4.
Table 6-13 shows the electricity prices of demand regions for scenario III. On the basis of Table 6-13, the electricity price of D5 is the lowest while the electricity price of D1 is the highest. The electricity prices are increasing with the order of D5, D4, D3, D2 and D1.

<table>
<thead>
<tr>
<th>Demand region</th>
<th>Year</th>
<th>PRICE($/GWh)</th>
</tr>
</thead>
<tbody>
<tr>
<td>D1</td>
<td>2006</td>
<td>5.9813E+6</td>
</tr>
<tr>
<td>D2</td>
<td>2006</td>
<td>5.9754E+6</td>
</tr>
<tr>
<td>D3</td>
<td>2006</td>
<td>5.9086E+6</td>
</tr>
<tr>
<td>D4</td>
<td>2006</td>
<td>5.8622E+6</td>
</tr>
<tr>
<td>D5</td>
<td>2006</td>
<td>5.8416E+6</td>
</tr>
</tbody>
</table>

In order to perform the scenario IV, we randomly choose one of the optimal HVDC determined in scenario II and shut down this HVDC. Then we run the model for scenario
IV. Table 6-14 shows the optimal locations and optimal capacity sizes of photovoltaics for scenario IV. In terms of Table 6-14, five optimal locations of photovoltaics are determined and accordingly, five optimal capacity sizes of photovoltaics are also determined. These five optimal locations all lie within Germany.

Table 6-14. Optimal locations and capacity sizes of photovoltaics, Scenario IV

<table>
<thead>
<tr>
<th>Name</th>
<th>Year</th>
<th>Size (GW)</th>
<th>Annual production (GWh/a)</th>
</tr>
</thead>
<tbody>
<tr>
<td>S36</td>
<td>2006</td>
<td>0.0276</td>
<td>36.2740</td>
</tr>
<tr>
<td>S37</td>
<td>2006</td>
<td>0.0117</td>
<td>15.3531</td>
</tr>
<tr>
<td>S41</td>
<td>2006</td>
<td>0.0173</td>
<td>22.7956</td>
</tr>
<tr>
<td>S43</td>
<td>2006</td>
<td>0.0180</td>
<td>23.6909</td>
</tr>
<tr>
<td>S52</td>
<td>2006</td>
<td>1.7254</td>
<td>2267.1884</td>
</tr>
</tbody>
</table>

Table 6-15 shows the optimal capacity sizes of HVDC and the annual transported electricity by HVDC for scenario IV. According to Table 6-15, the optimal capacity size of HVDC is equal to the annual transported electricity by HVDC if a conversion factor of unit is considered.

<table>
<thead>
<tr>
<th>Name</th>
<th>From</th>
<th>To</th>
<th>Year</th>
<th>Size (GW)</th>
<th>Annual transported electricity (GWh/a)</th>
</tr>
</thead>
<tbody>
<tr>
<td>HVDC</td>
<td>S36</td>
<td>D1</td>
<td>2006</td>
<td>0.0041</td>
<td>36.2740</td>
</tr>
<tr>
<td>HVDC</td>
<td>S37</td>
<td>D4</td>
<td>2006</td>
<td>0.0018</td>
<td>15.3531</td>
</tr>
<tr>
<td>HVDC</td>
<td>S41</td>
<td>D2</td>
<td>2006</td>
<td>0.0026</td>
<td>22.7956</td>
</tr>
<tr>
<td>HVDC</td>
<td>S43</td>
<td>D3</td>
<td>2006</td>
<td>0.0027</td>
<td>23.6909</td>
</tr>
<tr>
<td>HVDC</td>
<td>S52</td>
<td>D5</td>
<td>2006</td>
<td>0.2588</td>
<td>2267.1884</td>
</tr>
</tbody>
</table>

Fig 6-13 depicts the optimal locations of photovoltaics and HVDC as well as the electricity transportation direction for scenario IV. The blue line represents HVDC and the arrow direction represents the electricity transportation direction. According to Fig 6-13, one optimal location of photovoltaics will supply electricity through HVDC to one demand region.
Table 6-16 shows the electricity prices of demand regions for scenario IV. The electricity price of demand region D1 is still the highest whereas the electricity price of demand region D5 is still the lowest. The electricity prices are decreasing with the order of D1, D2, D3, D4 and D5.

Table 6-16. Electricity prices of demand regions, Scenario IV

<table>
<thead>
<tr>
<th>Demand region</th>
<th>Year</th>
<th>PRICE($/GWh)</th>
</tr>
</thead>
<tbody>
<tr>
<td>D1</td>
<td>2006</td>
<td>5.9776E+6</td>
</tr>
<tr>
<td>D2</td>
<td>2006</td>
<td>5.9754E+6</td>
</tr>
<tr>
<td>D3</td>
<td>2006</td>
<td>5.9204E+6</td>
</tr>
<tr>
<td>D4</td>
<td>2006</td>
<td>5.8622E+6</td>
</tr>
<tr>
<td>D5</td>
<td>2006</td>
<td>5.8416E+6</td>
</tr>
</tbody>
</table>
6.2 Case study of IST model

6.2.1 Materials

6.2.1.1 Study area

In this case study, the study area is the U.S. which is composed of 50 states and a federal district. Because Alaska and Hawaii are not directly related with the other states of U.S., we exclude these two states and merely consider the 48 states and a federal district. Such processing has an advantage for the spatial statistical model selection which will be illustrated in the discussion (refer to 7.3 Discussion of case study of IST model). In this dissertation, the federal district of U.S. is referred to as a state of U.S. in order to better illustration. The spatial layout of the 49 U.S. states is illustrated by Fig 6-14. The numbers in Fig 6-14 are the codes of 49 states. The codes of 49 U.S. states and the related name are listed in Appendix I.

Fig 6-14: Spatial layout of 49 U.S. states.

6.2.1.2 Datasets

The complete datasets of this experiment are composed of Biomass Consumption (BC), the coordinates of states, the biomass PRICE and the Total Disposable personal INCOME (TDINCOME). In this experiment, biomass refers to the wood and waste. The consumption refers to the total consumption by the residential, commercial, industrial and transportation sectors of the economy of the studied states. The biomass price refers to the biomass price itself not the biomass product price. BC is in trillion Btu, PRICE is in dollar per million Btu (=$10^{-4}$ billion dollars per trillion Btu) and TDINCOME is in billion dollars. Table 6-17 lists the source of each dataset.
### Table 6-17. Sources of datasets of case study of IST model

<table>
<thead>
<tr>
<th>Dataset</th>
<th>Source</th>
<th>Year</th>
</tr>
</thead>
<tbody>
<tr>
<td>Biomass consumption</td>
<td>Energy Information Administration (EIA)</td>
<td>2005</td>
</tr>
<tr>
<td>Biomass price</td>
<td>Energy Information Administration (EIA)</td>
<td>2005</td>
</tr>
<tr>
<td>Total disposable personal income</td>
<td>Regional Economic Information System, Bureau of Economic Analysis, U.S. Department of Commerce</td>
<td>2005</td>
</tr>
</tbody>
</table>

Fig 6-15 demonstrates the spatial distribution of biomass consumption. According to Fig 6-15, most of large biomass consumptions are around the border states of U.S., especially around the south-west border states of U.S. For example, based on Fig 6-15, Alabama, Florida and Georgia have large biomass consumptions.

![Biomass consumption map](image)

**Biomass consumption (Trillion Btu)**

- **0.900000 - 11.900000**
- **11.900001 - 37.700000**
- **37.700001 - 60.100000**
- **60.100001 - 119.400000**
- **119.400001 - 180.700000**

Fig 6-15: Spatial distribution of biomass consumption, U.S.

Fig 6-16 depicts the spatial distribution of biomass price. Based on Fig 6-16, the high biomass prices cluster around the middle part of U.S.. Interestingly, the biomass price has an inverse tendency to biomass consumption. Take Arizona State for example, the biomass price of Arizona State is rather high whereas biomass consumption of Arizona State is pretty small.
Fig 6-17 illustrates the spatial distribution of total disposable personal income. According to Fig 6-17, the low total disposable personal incomes locate around the north part of U.S.. It is interesting that the total disposable personal income has a positive relation with biomass consumption. Take California State for instance, the total disposable personal income of California State is greatly high while biomass consumption of California State is rather large.
In spatial trend model and IST model which will be used in this experiment to investigate the spatial influencing factors of biomass consumption in U.S., the definition of location plays a pretty major role. It will influence the following spatial statistical analysis as well as the inference. The right definition of location can obtain the right results and thus the subsequent inference is valuable, whereas the wrong definition of location will make the results invalid and thus the following inference is spurious. The requirement to the definition of location is in general that the location can exactly represent the studied area.
Normally, the location must lie within the studied area. The geometric centroids of area are always thought of as the representative of this area. However, for large scale area-based data, the geometric centroids of area do not always lie within this area. In order to solve the problem, Cressie (1991) defined the capital location as the location of city in order that the location lies within the considered city. In this experiment, we use the Minimal Bounded Rectangular (MBR) of the 49 U.S. states to obtain the locations of these states. Specially, we take the centroids of the MBR of these 49 U.S. states as the locations of states. If the centroids of the MBR of these states lie out the states, we take the most representative location of these states, such as the capital, the church, etc., as the locations of states. Fig 6-18 shows the locations of 49 U.S. states. According to Fig 6-18, all centroids of the MBR of these states lie within the states.

6.2.1.3 Spatial Econometric toolbox

The spatial econometric toolbox is developed mainly by James P. LeSage, Department of Economics, the University of Toledo. The toolbox is based on MATLAB and mainly deals with the spatial econometric problem. Rather than developing advance spatial econometric routines for econometric analysis and inference, the toolbox provides a set of basic routines and functions to support econometric analysis and the following inference, and tries to provide a relatively complete set of basic econometric analysis tools such that the users do not do a lot of work to solve their problem. The toolbox also includes a number of functions to mimic those available in Gauss, which should make converting existing Gauss functions and applications easier (LeSage 1999).

6.2.2 Experimental workflow

The aims of the experiment are to investigate the spatial influencing factors of biomass consumption in U.S., especially the location factor, through spatial trend model and IST model, and then select the best model for biomass consumption in U.S. via the spatial statistical model selection procedure proposed in this dissertation.
In order to perform the experiment, a workflow is presented and is as follows: First, the spatial trend model will be employed to analyze the pure spatial trend of biomass consumption in U.S.. The centroids of the MBR of the 49 U.S. states and the biomass consumption will be input into the spatial trend model. Then we check the ill-conditioned least squares problem. If the problem does not happen, then we can check the spatial autocorrelation for regressive residuals. If the residuals are spatially autocorrelated, we have to modify the model specification and repeat the preceding steps. However, if the residuals are not spatially autocorrelated, we can calculate the $R^2$, $R_a^2$, AIC and SIC. When the ill-conditioned least squares problem occurs, the IST model workflow starts and then the centroids of the MBR of the 49 U.S. states, biomass consumption, the biomass price
and the total disposable personal income will be input into the IST model. In this case, we need only check the spatial autocorrelation for residuals. This step is the same with the preceding step for checking the spatial autocorrelation among residuals in spatial trend model. After all the preceding steps are finished, we obtain lots of models as well as the associated $R^2$, $R_a^2$, AIC and SIC. Then we can choose the best model for biomass consumption in the 49 U.S. states with respect to $R^2$, $R_a^2$, AIC and SIC. In this experiment, we only consider the IST models which have the same orders with the spatial trend models. The whole workflow is illustrated by Fig 6-19. In Fig 6-19, the black solid lines represent the workflow of spatial trend model and the blue dashed lines represent the workflow of IST model. The green dashed line triggers the workflow of IST model. The arrows represent the directions of workflows.

In this experiment, the employed test for spatial autocorrelation among residuals is the Moran’s I test. According to the definition of Moran’s I test, the essential component of Moran’s test is the definition of spatial weight. The widely applied spatial weight matrices are the first-order spatial contiguity weight matrix and the second-order spatial contiguity weight matrix. In this experiment, we only consider these two kinds of spatial contiguity weight matrices.

### 6.2.3 Results

We first take use of the spatial trend models to analyze the spatial trend of biomass consumption in U.S.. The estimation method for the parameters of spatial trend models is the ordinary least square. Table 6-18 gives the estimated values of the parameters of spatial trend models. In order to evaluate the models and measure the significance of the parameters, Table 6-18 also contains the $R^2$, $R_a^2$, AIC and SIC as well as the t-value of T-test for these parameters. In Table 6-18, $a_1$ is the parameter related to $x$ coordinate, $a_2$ is the parameter related to $y$ coordinate. The combinations of $a_1$ and $a_2$ represent the parameter related to the combinations of $x$ coordinate and $y$ coordinate. For example, $a_1a_2$ is the parameter related to $xy$.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Linear model</th>
<th>Quadratic model</th>
<th>Cubic model*</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>coefficient</td>
<td>coefficient</td>
<td>coefficient</td>
</tr>
<tr>
<td>constant</td>
<td>260.2942</td>
<td>2,519.2176</td>
<td>-5,470.7958</td>
</tr>
<tr>
<td>$a_1$</td>
<td>0.4361</td>
<td>13.8878</td>
<td>-262.0526</td>
</tr>
<tr>
<td>$a_2$</td>
<td>-4.2321</td>
<td>-87.7923</td>
<td>-148.2800</td>
</tr>
</tbody>
</table>
According to Table 6-18, when the spatial trend model is cubic, the ill-conditioned least squares problem occurs. On the basis of Table 6-18, the $R^2$ and $R_a^2$ of the linear spatial trend model are rather low, which are 0.1604 and 0.1239 respectively. The $R^2$ and $R_a^2$ of the quadratic spatial trend model are relatively high, which are 0.4142 and 0.3461 respectively. For linear spatial trend model, the t-values show that the constant and y coordinate are significant at the significance level of 1%. However, the x coordinate is not significant at the significance level of 1%. For quadratic spatial trend model, the t-values show that the constant and all pure y coordinates (i.e. only composed of y coordinates) are significant at the significance level of 1%. The x coordinate and terms containing x coordinate (e.g. $x^2$ and $xy$ terms) are, however, not significant at the significance level of 1%.

Since the ill-conditioned least squares problem happens when the order of the spatial trend model is six, in this case we should modify the spatial trend model specification. The IST models proposed in this dissertation are adopted. In the IST models, we should determine $f(X, \beta)$ which is the essential component of the IST model. We take advantage of classical economic theory to determine $f(X, \beta)$.

According to the classical economic theory, the goods consumption has an intimate relation with the goods price and total disposable personal income. If the goods price is low, people will buy more goods than expected without changing the expenditure schedule. As a consequence, the goods is growingly consumed. However, if the goods price is high, people will not buy more goods than expected due to the limited expenditure

<table>
<thead>
<tr>
<th>Expression</th>
<th>$a_1^2$</th>
<th>$a_{12}$</th>
<th>$a_2^2$</th>
<th>$a_3^3$</th>
<th>$a^2_{12}$</th>
<th>$a_{12}^2$</th>
<th>$a_3^3$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$a_1^2$</td>
<td>0.0577</td>
<td>1.8165</td>
<td>-3.3723</td>
<td>-4.3852</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$a_{12}$</td>
<td>-0.0557</td>
<td>-0.4859</td>
<td>-2.7686</td>
<td>-1.3825</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$a_2^2$</td>
<td>0.9961</td>
<td>3.5586</td>
<td>-0.6342</td>
<td>-0.1273</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$a_3^3$</td>
<td>-0.0110</td>
<td>-4.6724</td>
<td>0.9961</td>
<td>3.5586</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$a^2_{12}$</td>
<td>0.0068</td>
<td>0.9636</td>
<td>0.0540</td>
<td>2.3337</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$a_{12}^2$</td>
<td>0.0586</td>
<td>1.1623</td>
<td>0.0586</td>
<td>1.1623</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.1604</td>
<td>0.4142</td>
<td>0.6547</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$R_a^2$</td>
<td>0.1239</td>
<td>0.3461</td>
<td>0.5750</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AIC</td>
<td>10.6992</td>
<td>10.4616</td>
<td>10.0965</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SIC</td>
<td>10.8150</td>
<td>10.6932</td>
<td>10.4826</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Matrix is close to singular or badly scaled. Results may be inaccurate (ill-conditioned least squares).
schedule and thus the goods consumption will be decreased. If the total disposable personal income increases, people will spend more money to the goods and thus more goods will be consumed. If the total disposable personal income, however, decreases, people will not spend more money to the goods and thus the goods consumption will be decreased. Briefly, the goods consumption has a negative relation with the goods price and has a positive relation with the total disposable personal income. The classical economic theory can be applied to biomass. Based on Fig 6-15, Fig 6-16, and Fig 6-17, we can identify that in the state where the biomass price is high, biomass consumption is low whereas in the state where the biomass price is low, biomass consumption is high. The Arizona State is one of the good examples. In terms of Fig 6-15, Fig 6-16, and Fig 6-17, we can also identify that in the state where the total disposable personal income is high, biomass consumption is high while in the state where the total disposable personal income is low, biomass consumption is low. The California State is one of the good instances. The analysis is consistent with the conclusions of Lambert D’Apote (1998), Nakićenović et al. (1998), IEA (2002) and World Bank (2002) that biomass consumption has a positive relation with GDP. If we further analyze Fig 6-15, Fig 6-16, and Fig 6-17, we can find that Connecticut State has high biomass consumption, middle biomass price and high total disposable personal income. The phenomenon is consistent with the classical economic theory. Therefore, we take the biomass PRICE and the Total Disposable personal INCOME (TDINCOME) as the explanatory variables X of \( f(X, \beta) \) and take the linear model of \( f(X, \beta) = \beta_1 PRICE + \beta_2 TDINCOME \).

We take the ordinary least square method to estimate the parameters of the IST models. Because the cubic spatial trend model has caused the ill-conditioned least squares problem, we only consider the IST model with linear spatial trend and the IST model with quadratic spatial trend. Accordingly, we call the resulted models as the Linear IST model (LIST model) and the Quadratic IST model (QIST model). Table 6-19 reports the estimated values of the parameters of IST models. Table 6-19 also documents the \( R^2 \), \( R_a^2 \), AIC and SIC as well as the t-value of the T-test for these estimated parameters.

<table>
<thead>
<tr>
<th>Variable</th>
<th>LIST model</th>
<th>QIST model</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>coefficient</td>
<td>t-value</td>
</tr>
<tr>
<td>constant</td>
<td>195.8049</td>
<td>3.6559</td>
</tr>
<tr>
<td>PRICE</td>
<td>-13.4305</td>
<td>-4.3811</td>
</tr>
<tr>
<td>TDINCOME</td>
<td>0.0958</td>
<td>3.7292</td>
</tr>
<tr>
<td>a1</td>
<td>-0.3002</td>
<td>-0.7559</td>
</tr>
</tbody>
</table>
According to Table 6-19, The $R^2$ and $R_a^2$ for the LIST model are 0.5909 and 0.5538 respectively. These values are over 50%. The $R^2$ and $R_a^2$ for the QIST model are 0.6910 and 0.6383 respectively. These values are also over 50% and higher than those of the LIST model. In terms of Table 6-19, the t-values of the biomass price and total disposable personal income in the LIST model and QIST model show that the biomass price and the total disposable personal income are significant at the significance level of 1% and have significant relation with biomass consumption, which is consistent with the classical economic theory. The t-values of the LIST model and QIST model also show that the pure $y$ coordinates are significant at the significance level of 1%.

In order to select the best model to identify the spatial influencing factors of biomass consumption in U.S., we should first test the spatial autocorrelation among residuals for all models. We adopt the Moran’s I test (because in this experiment we use spatial statistical model selection procedure to select model) and take the first-order spatial contiguity weigh matrix $W_1$ and the second-order spatial contiguity weight matrix $W_2$ respectively as the spatial weight of Moran’s I test. Table 6-20 reports the Moran’s I test values, the mean of Moran’s I test, the variance of Moran’s I test and the p-value of Moran’s I test. According to Table 6-20, at the significance level of 1%, the Moran’s I tests with spatial weight $W_1$ show that except that the residuals resulting from the linear spatial trend model are spatially autocorrelated, the residuals stemming from quadratic spatial trend model, LIST model and QIST model have no spatial autocorrelation. At the same significance level, the Moran’s I tests with spatial weight $W_2$ do not detect any spatial autocorrelation among residuals for these four models.

<table>
<thead>
<tr>
<th>$a_2$</th>
<th>-3.4788</th>
<th>-3.0219</th>
<th>-61.8014</th>
<th>-3.6019</th>
</tr>
</thead>
<tbody>
<tr>
<td>$a_1^2$</td>
<td>0.0352</td>
<td>1.439</td>
<td>0.9580</td>
<td>2.8705</td>
</tr>
<tr>
<td>$a_2^2$</td>
<td>0.0859</td>
<td>-0.9580</td>
<td>0.6423</td>
<td>3.6019</td>
</tr>
</tbody>
</table>

LIST model= linear IST model, QIST model= quadratic IST model.

$R_a^2$ ≡ adjusted R square.
<table>
<thead>
<tr>
<th>Model</th>
<th>I</th>
<th>E(I)</th>
<th>σ(I)</th>
<th>P-Value</th>
<th>I</th>
<th>E(I)</th>
<th>σ(I)</th>
<th>P-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Linear model</td>
<td>0.1831</td>
<td>-0.0588</td>
<td>0.0911</td>
<td>0.0079</td>
<td>-0.0268</td>
<td>0.0472</td>
<td>0.0645</td>
<td>0.7520</td>
</tr>
<tr>
<td>Quadratic model</td>
<td>-0.0361</td>
<td>-0.1050</td>
<td>0.0843</td>
<td>0.4139</td>
<td>-0.0367</td>
<td>-0.0640</td>
<td>0.0622</td>
<td>0.6604</td>
</tr>
<tr>
<td>LIST model</td>
<td>-0.0007</td>
<td>-0.0622</td>
<td>-0.0896</td>
<td>0.4928</td>
<td>0.0242</td>
<td>-0.0511</td>
<td>0.0643</td>
<td>0.2422</td>
</tr>
<tr>
<td>QIST model</td>
<td>-0.1835</td>
<td>-0.1056</td>
<td>-0.0830</td>
<td>0.3474</td>
<td>0.0173</td>
<td>-0.0675</td>
<td>0.0621</td>
<td>0.1719</td>
</tr>
</tbody>
</table>

LIST = linear IST model, QIST = quadratic IST model.
7 Discussion

7.1 Discussion of OFLR model

The advantages of the OFLR model design and OFLR model description have been discussed respectively in (Chapter 3 Design of OFLR Optimization Model for Supply) and (Chapter 4 Description of OFLR model). In this section, we further discuss some advantages of OFLR model not mentioned before and discuss some weakness and possible solutions.

The main purpose of the OFLR model is to determine the optimal locations and optimal capacity sizes of solar and wind plants by integrating remote sensing data. Remote sensing data which is input into the OFLR model aims to improve the final results and therefore contribute to better present the country level or larger regional level policy suggestions which play a key role in making macro solar and/or wind energy planning. The main question then gives rise to that whether remote sensing data can fulfill the aims. The answer is positive. Remote sensing data can improve the assessment accuracy of solar and wind energy potentials, impose remote sensing-based environmental constraints on OFLR model and extend the OFLR model to country level or larger regional level. Another issue is whether remotes sensing data is the necessary input parameter. Remote sensing data is one of the input data. As a matter of fact, without remote sensing data, the final results can also be obtained and the associated policy suggestions can be presented as well. However, these final results are not as good as the final results with remote sensing data.

In the OFLR model, the by-products like heat are also considered. This consideration is of importance. According to Dundett et al. (2008), when by-products were considered, the optimal locations and optimal capacity sizes of renewable energy plants were indeed influenced.

In the OFLR model, not only the continuation capacity size but also the pre-determined capacity sizes for solar and wind energy facilities are considered. In renewable energy planning, we normally have no any requirements for renewable energy plants’ capacity sizes. The photovoltaics with capacity of 0.152GW is no matter. However, in some cases
some requirements will be imposed to renewable energy plants’ capacity sizes. For example, the capacity size of photovoltaics has to conform to an implemental minimum capacity size. It would make no sense to set up 10kW photovoltaics. In these cases, the consideration for the pre-determined capacity sizes of energy plants in the OFLR model is pretty appreciated.

Model is a simplification of the complex reality and only extracts the main elements which are necessary to reveal the underlying complex reality. So in order to develop model, the simplification is necessary and important. The simplification to the supply chain system considered by the OFLR model removes battery because it deems that battery is not the essence of this supply chain system. Moreover, the simplification only considers HVDC and higher efficiency transportation plants because it regards that hydrogen and HVAC are out-of-date techniques. Such simplifications reduce the difficulty of formulating the OFLR model. However, a question will appear that whether the simplification exactly capture the main elements of underlying reality. The answer to question has close relation with the concept of main elements. The definition of main elements varies from one person to another person and relies primarily upon their understanding of the underlying phenomenon. However, some common accepted main elements are the same for one phenomenon. In this sense, the answer to the question is positive.

In the OFLR model, the studied area is transformed into grid. The grid structure of studied area can naturally integrate remote sensing data into the OFLR model and thus implement the main purpose of the OFLR model. Further, the grid structure of studied area enhances the coupling of the OFLR model with remote sensing model and GIS model. However, a problem related to this grid structure is presented: which spatial resolution is the appropriate spatial resolution of grid to cover the studied area. The solution to this problem is still under the way.

The economic theory employed by the OFLR model is the equilibrium theory with fixed demand. The theory is easily understood and commonly accepted. Therefore, the policy suggestions presented by the OFLR model are easily accepted. Moreover, because the equilibrium theory is one of the basic macro theories, in order to make the OFLR model provide more policy suggestions, the variables of the OFLR model are elaborated as the
basic variable of macro theories. Consequently, when putting two or more variables of OFLR model together, we can get more valuable economic meaning. Hence, we can present more policy suggestions. However, this theory assumes that the demand is fixed. In reality, we can also encounter the dynamic demands. The dynamic demands are not considered in OFLR model because the formulation for model with the dynamic demands is extremely complicated. It will involve more economic theories and need more mathematic knowledge.

In the OFLR model, the unit conversion is not considered. In renewable energy system, different units will be used. These units involve facilities, commodities and currency etc. the conversions among these units are often complicated. In reality, different applications will need different sets of units. It makes no sense to make a set of standard units for models. Moreover, the determination of unit for facilities is complicated. For example for renewable energy plant with single product, the units of capacity size and production are easy to be defined. The conversion factor of units is also easily determined. However, for renewable energy plant with two or more products, the situation becomes complex. We should choose the commodity group to determine the units of capacity size and production. Generally, this determination of units for capacity size and production varies from one research to other research and from one person to other person. In order to avoid involving complicated processing of units, the OFLR model does not take into account the conversion factors of units. Such processing is also found in TIMES model (Lou lou et al. 2005). This consideration makes modeler concentrate on development of model rather than the detail unit conversion.

Instead of focusing on all renewable energy, the OFLR model restricts the attention to solar and wind energy. This restriction can contribute to the formulation of the OFLR model. As a matter of fact, the recently developed renewable energy models focus on one or more kinds of renewable energy but not all (NREL 2000; IMAGEteam 2001). However, this restriction will make the OFLR model only determine the optimal locations and optimal capacity sizes of solar energy facilities and/or wind energy facilities and create the spatially explicit location planning strategies for solar and/or wind energy.

In the OFLR model, only electricity supply chain is considered. Without big change, other renewable products supply chains can, however, be added to the OFLR model.
The OFLR model distinguishes the supply regions and the demand regions. Solar and wind energy are supplied from supply regions to demand regions. However, a problem will appear whether one region is only the supply region or the demand region. The answer is negative. Each region has a chance of simultaneously the supply region and the demand region. They are not only self-providers but also supplier/demander. Such case can be coped with by the OFLR model. The OFLR model deals with them as not only the supply regions but also simultaneously the demand regions. They can supply solar and wind energy to any other regions and also can get solar and wind energy from any other regions. The most importance is that they can supply solar and wind energy to themselves through transportation plants. Therefore they can be modeled by the OFLR model.

The OFLR model does not take into account the past setup solar and wind energy facilities and transportation facilities. This problem can be solved by assigning the start nonzero values to the corresponding solar and wind energy facilities and transportation facilities if they have past setup. The start nonzero values are exactly the past setup.

In the OFLR model, two kinds of remote sensing-based constraints with remote sensing data are considered because they are widely encountered in reality. This consideration enhances the flexibility of OFLR model.

### 7.2 Discussion of IST model

The main purpose of IST model is to solve the ill-conditioned least square problem which often happens in spatial trend model while the residuals are not spatially autocorrelated and the estimation for its parameters is simple. When spatial trend model has no ill-conditioned least square problem but low goodness of fit, the IST model should still be developed. Even if the underlying phenomenon has no any relationship with locations, the IST model can still be used and reduce to identify the non-location spatial factors of the underlying phenomenon. When the underlying phenomenon can be explained completely by locations, the IST model will, in this case, degenerate to spatial trend model.

The definition of location plays a key role in IST model. It is no doubt that bad definition of location will make IST model invalid and therefore the IST model can not identify the spatial influencing factors for renewable energy consumption. However, the definition of
location varies from person to person, from data sets to data sets and from problem to problem. For point data, the definition of location is unambiguous. However, for area data, the exact definition of location is extremely hard to make. Cressie (1991) defined the capital location as the location of city in order that the location lies within the considered city. However, not all research areas are the cities or have capital. We recommend taking the centroid of MBR of studied area as the location of this area. However, the centroid of MBR of studied area still has chance to lie out the area. The IST model does not provide the fix definition of location. Various definitions of location can be realized and adopted in the IST model through modifying $A$ of IST model. Then, a question will give rise to that the adopted definition of location for IST model is reasonable. This can be assessed by final regressed results of IST model.

The IST model does not provide the detail appearance of non-location spatial influencing factors of renewable energy consumption. The determination of non-location spatial influencing factors of renewable energy consumption varies from author to author. The influencing factors of biomass consumption considered by Lambert D’Apote (1998) and Victor and Victor (2002) are population, income and energy price. In LEAP model (LEAP 2002) these three factors are also identified to trace renewable energy consumption. However, Nakićenović et al. (1998), IEA (2002) and World Bank (2002) argued that GDP was an influencing factor of biomass consumption. As a matter of fact, the spatial influencing factors of renewable energy consumption are far more than these four factors. For example, supply is also one of factors (Basile 1980). Moreover, in reality the situation we will encounter is we do not always acquire the proper data set for IST model. In this case, if the detail appearance of non-location spatial factors of the IST model is fixed, the model can not be used. In order to allow more spatial influencing factors to be considered, the detail appearance of non-location spatial factors of the IST model is not fixed.

Since the definition of location and detail appearance of non-location spatial factors of the IST model are not fixed, an issue will arise that how to assure the adopted factors are sufficient for renewable energy consumption. The answer needs to get a hand from the spatial statistical model selection procedure proposed. The main purpose of the spatial statistical model selection procedure is to assess which IST model does best identify the spatial influencing factor of renewable energy consumption. In this procedure, only Moran’s I is adopted. However, the Wald test, LR test and LM test can be also considered.
Just like her name, the spatial statistical model selection procedure is solely applied to select spatial statistical model.

In essence, the IST model is a model frame which focuses on especially the spatial location factor rather than a detail model expression.

### 7.3 Discussion of case study of OFLR model

The first problem we encountered in this experiment is the determination of the appropriate spatial resolution for grid to cover studied area. Different spatial resolutions of grid will generate different amount of grids such that the final results will be different. Determining an appropriate spatial resolution for grid is a greatly difficult thing. So many factors, such as the understanding of the problems, the aims of the experiment and the availability of the data sets, etc., should be taken into account. In this experiment, we only take into account the availability of data sets, the computational time of the OFLR model and the data processing to determine the spatial resolution of grid of considered region. We obtain the solar direct normal irradiation data with the spatial resolution of 1×1 degrees and obtain the land slope data with the spatial resolution of 0.5×0.5 degrees as well as the landuse data with the spatial resolution of 300×300 m². Then we take the spatial resolution of 0.5×0.5 degrees as the spatial resolution of grid in this experiment. If we take the spatial resolution of 0.05×0.05 degrees, the amount of the potential locations of photovoltaics are up to 5200 rather than 52. The amount of the potential HVDC will be up to 26000 rather than 260. So to calculate the suitable area for each grid and run the model, it will take a longer time.

The potential of solar energy plays a major role in the supply chain system of solar energy (Wang et al. 2009; Wang and Koch 2010). The potential of solar energy has to be enough to satisfy the demand for solar energy. This experiment takes into account the theoretical potential of solar energy and the geographical potential of solar energy with the help of remote sensing data. Table 7-1 documents the theoretical potential of solar energy, the geographical potential of solar energy and the improvement for randomly selected grids. The improvement refers to the ratio of the difference of the theoretical solar potential extracting the geographical solar potential by the theoretical solar potential. On the basis of Table 7-1, we can evidently know that the improvement is very high, up to 0.9889. The
improvement is due to the consideration of the available area of land category which can be used to photovoltaics, and the introduction of the suitable factors to landuse remote sensing data.

<table>
<thead>
<tr>
<th>Location name</th>
<th>Theoretical solar potential (GWh/a)</th>
<th>Geographical solar potential (GWh/a)</th>
<th>Improved</th>
</tr>
</thead>
<tbody>
<tr>
<td>S36</td>
<td>2246788.027</td>
<td>24907.961</td>
<td>0.9889</td>
</tr>
<tr>
<td>S37</td>
<td>2238128.836</td>
<td>29284.484</td>
<td>0.9869</td>
</tr>
<tr>
<td>S38</td>
<td>2181709.048</td>
<td>29460.369</td>
<td>0.9865</td>
</tr>
<tr>
<td>S39</td>
<td>2169303.376</td>
<td>25487.029</td>
<td>0.9883</td>
</tr>
<tr>
<td>S40</td>
<td>2169303.376</td>
<td>24911.996</td>
<td>0.9885</td>
</tr>
<tr>
<td>S41</td>
<td>2082833.202</td>
<td>25810.768</td>
<td>0.9876</td>
</tr>
<tr>
<td>S42</td>
<td>2091848.359</td>
<td>24909.217</td>
<td>0.9881</td>
</tr>
<tr>
<td>S43</td>
<td>2076932.231</td>
<td>20827.027</td>
<td>0.9900</td>
</tr>
<tr>
<td>S44</td>
<td>2015182.308</td>
<td>25710.947</td>
<td>0.9872</td>
</tr>
<tr>
<td>S45</td>
<td>2002776.935</td>
<td>25836.609</td>
<td>0.9871</td>
</tr>
</tbody>
</table>

It is shown in this experiment that the results of scenario I and scenario II are the same, indicating that the geographical potential of solar energy in this experiment has no more influence than the theoretical potential of solar energy on the optimal locations and optimal capacity sizes of photovoltaics as well as the electricity prices. However, if we further analyze the results and data, we can find that the optimal capacity size of photovoltaics in grid is still smaller than the geographical potential of solar energy in the grid for four scenarios. According to the results of scenarios II and III, if we regard the location where the optimal photovoltaics is shut down as the location with zero geographical potential of solar energy while the theoretical potential of solar energy for this location has no any change and contribute the reason that the optimal photovoltaics is shut down to the fact of the location having no enough geographical solar potential to set up photovoltaics, we can clearly know that the geographical potential of solar energy has great influence on the optimal locations and optimal capacity sizes of photovoltaics, especially when the theoretical potential of solar energy is larger than the optimal capacity size of photovoltaics while the geographical potential is smaller than the optimal capacity size of photovoltaics. Hence, we should pay more attention to the geographical potential of solar energy than to the theoretical potential of solar energy because the former takes account of the impact of landuse.

In this experiment, the potential geographical locations of photovoltaics are explicitly
determined with the help of land slope remote sensing data. This consideration is necessary and important for creating the location planning strategy for photovoltaics (Leduc et al. 2009; Wang and Koch 2010). The explicitly geographical locations of photovoltaics are proved to influence the optimal locations and optimal capacity sizes of photovoltaics (see results of scenarios II and III). Therefore, the explicitly geographical locations of photovoltaics should be cautiously dealt with.

The spatial distribution of HVDC is also proved to influence the optimal locations and optimal capacity sizes of photovoltaics (see results of scenarios II and IV). So the spatial distribution of HVDC should also be cautiously dealt with.

It is noticed in this experiment that the optimal locations of photovoltaics all are situated in Germany in this experiment. Among the five studies countries, the photovoltaics price of Germany is the lowest, which is 5.50$/W (Table 6-4). Because the photovoltaics price of each potential location is the same with that of country to which the potential location belongs, the photovoltaics prices of the potential locations situated in Germany are also the lowest. It indicates that the photovoltaics price directly influences the optimal location of photovoltaics (Wang et al. 2009; Wang and Koch 2010). To verify this statement, we randomly choose one of supply regions which do not belong to Germany and decrease the photovoltaics price of this supply region to the cheapest. The chosen supply region is S18, which belongs to France. Then provided that only the photovoltaics price of S18 is changed, we perform the four scenarios again. The results confirm that S18 does become the one of optimal locations to set up photovoltaics. Consequently, in order to set up photovoltaics in one location, the feasible consideration is to decrease the photovoltaics price for this location. Therefore, we should pay more attention to the research and development of photovoltaics than before.

The experiment found that the optimal locations of photovoltaics had intimate relation with the explicitly geographical locations of demand regions. They were close to demand regions as much as possible, suggesting that the geographical location of demand region will influence the optimal location of photovoltaics. This was because in this experiment we took into account the cost of transportation facilities (i.e. HVDC). Therefore, when creating location planning strategy for renewable energy facilities, we should take into account the explicitly geographical locations of demand regions.
It is found in this experiment that the electricity price has close relation with the line distance of optimal HVDC. When the line distance of optimal HVDC increases, the electricity price will increase. For example, the electricity price of demand region D1 supplied by supply region S36 is 5.9776$/kWh in scenario II (Table 6-10) whereas the electricity price of the same demand region D1 supplied by supply region S37 is 5.9813$/kWh in scenario III (Table 6-13). The line distance of HVDC for the former is shorter than that of HVDC for the latter. This is consistent with the fact.

The quality of data has an intimate relation with the results. Data quality is a classical issue in applications. Good data can produce the good and expected results, and vice versa. In this experiment, data all come from internet. Different data processing with different data criteria will generate different accumulated error for data. Hence in order to eliminate the accumulated error, knowing about the data processing and data criteria is very important for quality control. However, such information is so difficult to be obtained that we have no way to control the accumulated error in data. The seriously contaminated data will make the final results invalid. In this case, the potential contaminated data are the solar photovoltaics module price data, landuse data and solar irradiation data, etc.

7.4 Discussion of case study of IST model

The main purpose of spatial trend model and IST model intends to capture the spatial variation of the underlying feature against locations. In order to fulfill the purpose, the definitions of locations for the studied areas play a major role. In this experiment, in order to avoid the locations falling out the 49 U.S. states, we take the centroids of the MBR of these 49 states rather than the geometric centroids of these 49 states as the locations of them. The centroids of the MBR of these 49 states all fall within these states. Other methods, such as taking the capitals of the states as the locations of states, are not considered in this experiment.

The linear spatial trend model has rather low $R^2$, only 0.1604. Increasing the order of the spatial trend model can improve $R^2$. The $R^2$ of the quadratic spatial trend model is 0.4142 and the $R^2$ of the cubic spatial trend model can even reach 0.6547. However, the ill-conditioned least squares problem occurs for the cubic spatial trend model. Although for quadratic spatial trend model, the ill-conditioned least squares problem does not
happen, its $R^2$ is still low (lower than 50%). In this case, we should consider other factors. The presences of the biomass price and total disposable personal income into the spatial trend models improve the $R^2$, increasing from 0.1604 up to 0.6910. The highest $R^2$, which is 0.6910, belongs to QIST model and the value is satisfied.

The Moran’s I tests show that the residuals resulting from quadratic spatial trend model, LIST model and QIST models are not spatially autocorrelated (Table 6-20). It infers that biomass consumptions of the 49 U.S. states will potentially demonstrate the quadratic spatial trend and have close relation with the biomass price and total disposable personal income, which is consistent with Lambert D’Apote (1998) and Victor and Victor (2002). With respect to the $R^2$, $R_a^2$, AIC and SIC of all models, we definitely deem the QIST model as the best model identifying the spatial influencing factors of biomass consumptions in the 49 U.S. states because the $R^2$ and $R_a^2$ of this model are the highest and the AIC and SIC of this model are the smallest. In other words, biomass consumptions of the 49 U.S. states are related with not only the spatial locations but also the biomass price and total disposable personal income.

If we, however, further analyze these results, we can interestingly find that the spatial variations of biomass consumptions in the 49 U.S. states are strongly towards north-south and are weakly towards east-west (see Table 6-18 and Table 6-19). In order to better reveal the spatial influencing factors of biomass consumptions in the 49 U.S. states, we remove the terms related to $x$ coordinate from the spatial trend models and the IST models. Accordingly, we call the resulted models as the single spatial trend models and the single IST models. Because the spatial trend models with the order higher than six have no good physical explanation, we only consider the models with the order smaller than or equaling to six. Table 7-2 records the estimated values of the parameters of single spatial trend models with the ordinary least square estimated method and Table 7-3 reports the estimated values of the parameters of single IST models with the same estimated method. Table 7-4 records the results of Moran’s I test for the residuals of all these models.

| Table 7-2. Estimated values of parameters of single spatial trend models |
|---------------------------------|-----------------|-----------------|-----------------|-----------------|
| Variable | Linear single model | Quadratic single model | Cubic single model |
|         | coefficient | t-value | coefficient | t-value | coefficient | t-value |
| constant | 225.2701 | 3.6839 | 1617.4443 | 4.0599 | 3678.6697 | 1.2813 |
| $a_2$   | -4.3501 | -2.8293 | -76.9464 | -3.7310 | -240.5755 | -1.0615 |
| $a_2^2$ | 0.9330 | 3.5280 | 5.2133 | 0.8822 |
| $a_2^3$ | -0.0369 | -0.7250 |
Table 7-3. Estimated value of parameters of single IST models

<table>
<thead>
<tr>
<th>Variable</th>
<th>LSIST model</th>
<th>QSIST model</th>
<th>CSIST model</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>coefficient</td>
<td>t-value</td>
<td>coefficient</td>
</tr>
<tr>
<td>constant</td>
<td>212.8520</td>
<td>4.4025</td>
<td>1,176.5392</td>
</tr>
<tr>
<td>TDINCOME</td>
<td>0.0996</td>
<td>3.9736</td>
<td>0.1012</td>
</tr>
<tr>
<td>a²</td>
<td>-3.3438</td>
<td>-2.9543</td>
<td>-54.1338</td>
</tr>
<tr>
<td>a³</td>
<td>0.6533</td>
<td>3.2727</td>
<td>-1.1428</td>
</tr>
<tr>
<td>a²</td>
<td>0.0155</td>
<td>0.4105</td>
<td></td>
</tr>
<tr>
<td>R²</td>
<td>0.5856</td>
<td>0.6668</td>
<td>0.6681</td>
</tr>
<tr>
<td>R₂</td>
<td>0.5580</td>
<td>0.6365</td>
<td>0.6295</td>
</tr>
<tr>
<td>AIC</td>
<td>10.0338</td>
<td>9.8567</td>
<td>9.8936</td>
</tr>
<tr>
<td>SIC</td>
<td>10.1882</td>
<td>10.0498</td>
<td>10.1253</td>
</tr>
</tbody>
</table>

LSIST model = linear single IST model, QSIST model = quadratic single IST model, CSIST model = cubic single IST model.

On the basis of Table 7-2 and Table 7-3, we evidently know that the spatial variations of biomass consumptions in the 49 U.S. states are significantly towards North-South direction because the t-values of terms related to y coordinate in all single IST models and single spatial trend models excluding cubic single spatial trend model are strongly significant at the significance level of 1%. According to Table 7-4, the Moran’s I tests with spatial weight $W_1$ show that at the significance level of 1%, except that the linear single spatial trend model has spatially autocorrelated residuals, all other five models have...
Discussion

no spatially autocorrelated residuals. At the same significance level, the Moran’s I tests with spatial weight $W_2$ do not detect any spatial autocorrelation among residuals for these six models. In terms of Table 7-2, Table 7-3 and Table 7-4, we evidently know that among these six models, the QSIST model is the best model identifying the spatial influencing factors of biomass consumption in U.S. with respect to the $R^2$, $R_a^2$, AIC and SIC, suggesting that the spatial variations of biomass consumptions in the 49 U.S. states are quadratic and are towards the North-South direction, and biomass consumptions of these states have a positive relation with the total disposable personal income and have a negative relation with the biomass price.

Compared with the method proposed by Haining (2003), the IST model developed in this dissertation can not only solve the ill-conditioned least squares problem and eliminate the spatial autocorrelation among residuals but also take into account the other factors and be easily combined into other models. In addition, the estimation for its parameters is simple. The QSIST model is a good example.

The determination of the spatial weight matrices plays an important role in testing the spatial autocorrelation among residuals. The spatial weight matrices should be determined in advance and then can be contained into the spatial autocorrelation test. Certainly, the appropriate spatial weight matrices will indicate the real spatial autocorrelation among residuals. The wrong spatial weight matrices will, however, contaminate the spatial autocorrelation among residuals. In general, the spatial weight matrices can be determined through the spatial non-nested test. In this experiment, we only consider two kinds of widely applied spatial weight matrices. Further work will consider more spatial weight matrices.

Model selection is extremely important for understanding the underlying phenomenon. Selecting an appropriate model is a complicated thing. The spatial statistical model selection procedure proposed in this dissertation mainly considers the character of spatial datasets and the model assumption which is the necessary step to modeling the underlying phenomenon. As a result, the spatial statistical model selected by this selection procedure is preferred. In terms of this selection procedure, the best model identifying the spatial influencing factors of biomass consumptions in the 49 U.S. states is the QSIST model. In other words, the spatial trend of biomass consumptions in the 49 U.S. states is quadratic.
In this experiment, we first investigate the spatial location factor of biomass consumption. Actually, this investigation should be first carried out because it will provide the preliminary knowledge for further modeling biomass consumption and will contribute to develop a stationary spatial statistical model for biomass consumption. However, if we first investigate the non-location spatial influencing factors of biomass consumption, the results will be different.
8 Conclusions and Future Work

8.1 Conclusions

- The OFLR model developed in this dissertation aims to mainly determine the optimal locations and capacity sizes of a set of solar energy facilities and wind energy facilities in electricity supply chain system of solar and wind energy through integrating remote sensing data. In addition to determining the most appropriate locations and capacity sizes of solar and wind energy facilities, the model can be able to
  1) integrate remote sensing and GIS,
  2) perform in large geographical areas, such as at country level or larger regions,
  3) explicitly take into account the spatial information about supply regions, demand regions and facilities, etc.,
  4) take into account the spatial variation of demand,
  5) take into account the spatial variation of supply,
  6) include policy tools, such as setting carbon emission tax on emissions from production and transportation,
  7) take into account the environmental constraints,
  8) take into account new advanced solar and wind energy facilities,
  9) take into account not only the continual capacity sizes of solar and wind energy facilities but also the pre-determined discrete capacity sizes of solar and wind energy facilities,
  10) take into account electricity import and export,
  11) determine the prices of electricity for demand regions,
  12) take into account the by-products,
  13) be applied to the supply chain system with solar plants and/or wind plants.

The OFLR model is a valuable tool for decision makers in order to make the most suitable solar and/or wind energy planning strategy regarding the locations and capacity sizes of solar and/or wind plants with electricity product.

- The Improved Spatial Trend model (IST model) developed in this dissertation can identify the spatial influencing factors for renewable energy consumptions,
especially the location factor. It can solve the ill-conditioned least squares problem which often occurs in spatial trend model while the residuals are not spatially autocorrelated. Moreover, the estimation for its parameters is simple.

- The spatial statistical model selection procedure presented in this dissertation considers the spatial autocorrelation among residuals and is thus preferred in selecting the best IST model or other spatial statistical model for renewable energy consumption. It is simple and easily implemented.

- The consideration for the explicitly geographical locations of photovoltaics (one kind of solar energy facility) is necessary and important in creating the location planning strategy for photovoltaics in five European countries. These explicitly geographical locations will influence the optimal locations and capacity sizes of photovoltaics in five European countries. As a result, we should consider the spatial distribution of supply when creating location planning strategy for photovoltaics in five European countries.

- The optimal locations and capacity sizes of photovoltaics in five European countries will be influenced by solar potential. The geographical potential of solar energy is better than the theoretical potential of solar energy (improvement can even reach 0.9889) because it takes into account the impact of landuse which can be acquired from remote sensing images.

- The optimal locations and capacity sizes of photovoltaics in five European countries have strong relation with the photovoltaics price. The optimal locations all are situated in Germany, where the photovoltaics price is the lowest. Therefore, the research and development of photovoltaics should receive more attention than it has previously.

- The optimal locations and capacity sizes of photovoltaics in five European countries will be influenced by the explicitly geographical locations of demand regions. The optimal locations of photovoltaics are close to demand regions as much as possible. As a consequence, when creating location planning strategy for photovoltaics in five European countries, we should take into account the explicitly geographical locations of demand regions. That is, we should consider the spatial distribution of demand.

- Biomass consumptions of the 49 U.S. states have relation with not only the

---

3 The five European countries are respectively Austria, France, Germany, Italy, and Switzerland.

4 Actually, it is 48 states and one federal district.
locations of these states but also the biomass price and the total disposable personal income. The spatial trend of biomass consumptions of these states is towards the North-South direction and is quadratic. Biomass consumptions of these states have a positive relation with the total disposable personal income and have a negative relation with the biomass price. As a result, when creating scientific biomass consumption strategy for the U.S., decision makers should not only take into account the location effect but should also consider the biomass price and total disposable personal income.

### 8.2 Future work

In this dissertation, two models are proposed. Strictly speaking, these two models are two modules of single renewable energy model. In theory, these two models have been integrated into a single renewable energy model. The IST model can generate the spatial consumption data for the OFLR model. In renewable energy planning, supply planning and consumption planning are dependent on and complement one another. They are two components of renewable energy planning. When making long-term renewable energy supply planning, consumption planning should be simultaneously considered and when making long-term renewable energy consumption planning, supply planning should be simultaneously considered. It makes no sense that only supply planning or consumption planning be considered in creating renewable energy planning. For example, if only supply planning is considered, then we do not have any information about consumption and therefore we can not create scientifically optimal supply planning, and vice versa. The OFLR model can provide supply planning strategy whereas the IST model can provide consumption planning strategy. When these two models are integrated, a comprehensive renewable energy planning strategy can then be created (Belyaev et al. 1976) and is preferred. Because of time constrains, the linkage of these two models has not been conducted in this dissertation but should be integrated in a follow up approach.
References


Munasinghe M. (1988). Integrated National Energy Planning and Management:


World Bank. (2002). World Development Indicators, CD-ROM.
Appendix I: Codes of 49 U.S. states and related name

<table>
<thead>
<tr>
<th>STATE_CODE</th>
<th>STATE_NAME</th>
<th>STATE_CODE</th>
<th>STATE_NAME</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Alabama</td>
<td>26</td>
<td>Nebraska</td>
</tr>
<tr>
<td>2</td>
<td>Arizona</td>
<td>27</td>
<td>Nevada</td>
</tr>
<tr>
<td>3</td>
<td>Arkansas</td>
<td>28</td>
<td>New Hampshire</td>
</tr>
<tr>
<td>4</td>
<td>California</td>
<td>29</td>
<td>New Jersey</td>
</tr>
<tr>
<td>5</td>
<td>Colorado</td>
<td>30</td>
<td>New Mexico</td>
</tr>
<tr>
<td>6</td>
<td>Connecticut</td>
<td>31</td>
<td>New York</td>
</tr>
<tr>
<td>7</td>
<td>Delaware</td>
<td>32</td>
<td>North Carolina</td>
</tr>
<tr>
<td>8</td>
<td>District of Columbia</td>
<td>33</td>
<td>North Dakota</td>
</tr>
<tr>
<td>9</td>
<td>Florida</td>
<td>34</td>
<td>Ohio</td>
</tr>
<tr>
<td>10</td>
<td>Georgia</td>
<td>35</td>
<td>Oklahoma</td>
</tr>
<tr>
<td>11</td>
<td>Idaho</td>
<td>36</td>
<td>Oregon</td>
</tr>
<tr>
<td>12</td>
<td>Illinois</td>
<td>37</td>
<td>Pennsylvania</td>
</tr>
<tr>
<td>13</td>
<td>Indiana</td>
<td>38</td>
<td>Rhode Island</td>
</tr>
<tr>
<td>14</td>
<td>Iowa</td>
<td>39</td>
<td>South Carolina</td>
</tr>
<tr>
<td>15</td>
<td>Kansas</td>
<td>40</td>
<td>South Dakota</td>
</tr>
<tr>
<td>16</td>
<td>Kentucky</td>
<td>41</td>
<td>Tennessee</td>
</tr>
<tr>
<td>17</td>
<td>Louisiana</td>
<td>42</td>
<td>Texas</td>
</tr>
<tr>
<td>18</td>
<td>Maine</td>
<td>43</td>
<td>Utah</td>
</tr>
<tr>
<td>19</td>
<td>Maryland</td>
<td>44</td>
<td>Vermont</td>
</tr>
<tr>
<td>20</td>
<td>Massachusetts</td>
<td>45</td>
<td>Virginia</td>
</tr>
<tr>
<td>21</td>
<td>Michigan</td>
<td>46</td>
<td>Washington</td>
</tr>
<tr>
<td>22</td>
<td>Minnesota</td>
<td>47</td>
<td>West Virginia</td>
</tr>
<tr>
<td>23</td>
<td>Mississippi</td>
<td>48</td>
<td>Wisconsin</td>
</tr>
<tr>
<td>24</td>
<td>Missouri</td>
<td>49</td>
<td>Wyoming</td>
</tr>
<tr>
<td>25</td>
<td>Montana</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>