EFFECTS OF SPATIAL RESOLUTION AND LANDSCAPE STRUCTURE ON LAND COVER CHARACTERIZATION

Wenli Yang

University of Nebraska-Lincoln

Follow this and additional works at: http://digitalcommons.unl.edu/geographythesis

Part of the Environmental Monitoring Commons, and the Other Environmental Sciences Commons

Yang, Wenli, "EFFECTS OF SPATIAL RESOLUTION AND LANDSCAPE STRUCTURE ON LAND COVER CHARACTERIZATION" (1997). Theses and Dissertations in Geography. 24.

http://digitalcommons.unl.edu/geographythesis/24

This Article is brought to you for free and open access by the Geography Program (SNR) at DigitalCommons@University of Nebraska - Lincoln. It has been accepted for inclusion in Theses and Dissertations in Geography by an authorized administrator of DigitalCommons@University of Nebraska - Lincoln.
DISSERTATION TITLE

Effects of Spatial Resolution and Landscape Structure on Land Cover Characterization

BY

Wenli Yang

SUPERVISORY COMMITTEE:

APPROVED

Signature
James W. Merchant
Typed Name
Donald C. Rundquist
Typed Name
Clinton M. Rowe
Typed Name
Elizabeth A. Walter-Shea
Typed Name

DATE
July 17, 1997
July 17, 1997
17 July 1997
July 17, 1997

GRADUATE COLLEGE
UNIVERSITY OF NEBRASKA
EFFECTS OF SPATIAL RESOLUTION AND LANDSCAPE STRUCTURE ON LAND COVER CHARACTERIZATION

by

Wenli Yang

A DISSERTATION

Presented to the Faculty of
The graduate College at the University of Nebraska
In Partial Fulfillment of Requirements
For the Degree of Doctor of Philosophy

Major: Geography

Under the Supervision of Professor James W. Merchant

Lincoln, Nebraska

July, 1997
This dissertation addressed problems in scaling, problems that are among the main challenges in remote sensing. The principal objective of the research was to investigate the effects of changing spatial scale on the representation of land cover. A second objective was to determine the relationship between such effects, characteristics of landscape structure and scaling procedures. Four research issues related to spatial scaling were examined. They included: 1) the upscaling of Normalized Difference Vegetation Index (NDVI); 2) the effects of spatial scale on indices of landscape structure; 3) the representation of land cover databases at different spatial scales; and 4) the relationships between landscape indices and land cover area estimations.

The overall bias resulting from non-linearity of NDVI in relation to spatial resolution is generally insignificant as compared to other factors such as influences of aerosols and water vapor. The bias is, however, related to land surface characteristics. Significant errors may be introduced in heterogeneous areas where different land cover types exhibit strong spectral contrast. Spatially upscaled SPOT and TM NDVIs have information content comparable with the AVHRR-derived NDVI. Indices of landscape structure and spatial resolution are generally related, but the exact forms of the
relationships are subject to changes in other factors including the basic patch unit constituting a landscape and the proportional area of foreground land cover under consideration. The extent of agreement between spatially aggregated coarse resolution land cover datasets and full resolution datasets changes with the properties of the original datasets, including the pixel size and class definition. There are close relationships between landscape structure and class areas estimated from spatially aggregated land cover databases. The relationships, however, do not permit extension from one area to another. Inversion calibration across different geographic/ ecological areas is, therefore, not feasible. Different rules govern the land cover area changes across resolutions when different upscaling methods are used. Special attention should be given to comparison between land cover maps derived using different methods.
ACKNOWLEDGEMENTS

Throughout the course of my graduate program, I have had the support of special people without whose help the accomplishment of my doctorate would not have been possible.

My deepest and most sincere thanks must first go to my advisor, Dr. James Merchant, whose guidance and encouragement were instrumental to my study. He spent countless hours listening and providing direction while always allowing me the freedom to explore my own way. He also invested a great deal of time reading and commenting my various manuscripts. He financially supported my entire research through the grants awarded to him by the National Aeronautics and Space Administration (grant NAGW-3940) and the USGS/EROS Data Center. I will always regard him as my best mentor.

I own a very special debt of gratitude to Dr. Donald Rundquist, not only because he similarly provided invaluable guidance throughout my entire study but also because he is the professor who brought me to Nebraska. Without him, I would not have had the chance to enter this particular research program.

I am also grateful to the other members in my Supervisory Committee, Dr. Clinton Rowe and Dr. Elizabeth Walter-Shea, who gave their time and expertise in guiding my study. Their constructive comments on both my research proposal and my thesis manuscript served to clarify my thoughts and to improve significantly the dissertation.
Thanks are extended to Dr. Bruce Wylie of the USGS/EROS Data Center who provided the SPOT image classification and worked with me in the Niobrara fields.

I would also like to express my gratitude to the staff and fellow graduate students in CALMIT, the Conservation and Survey Division, and the Department of Geography, although I could not mention each of their names because the people I need to thank are too many.

My greatest debt of gratitude is owned to my parents. They have always selflessly loved me, supported me and given me constant comfort, no matter how far from home I roam. They always occupy a special place in my heart and my gratitude to them will never be enough.

The person whose constant source of support and encouragement I can never repay is my wife, Kemeng. It is her love that helped me survived this demanding period of my life. She endured more sacrifices than anyone else during my study, especially during our separations and in countless nights and weekends when I was doing research and writing. She gave me more love, patience and motivations than anyone else in helping me to materialize this academic dream. This work is dedicated to you, Kemeng!
TABLE OF CONTENT

Chapter 1 Introduction ................................................................................................................ 1
  1.1 Introduction .................................................................................................................... 1
  1.2 Background .................................................................................................................... 3
    1.2.1 Scale ....................................................................................................................... 3
    1.2.2 Scale and spatial resolution .................................................................................... 4
    1.2.3 Scaling of remotely sensed data .............................................................................. 5
  1.3 Definition of terms ......................................................................................................... 9
  1.4 Problem statement and research objectives .................................................................. 11
  1.5 Structure of the dissertation ......................................................................................... 16
  1.6 References ..................................................................................................................... 18

Chapter 2 Literature review ................................................................................................. 22
  2.1 Introduction .................................................................................................................... 22
  2.2 Classification accuracy and spatial resolution .............................................................. 23
  2.3 Spatial scaling of remote sensing data and classified land cover maps ...................... 27
  2.4 Scaling studies in landscape ecology ............................................................................ 38
  2.5 Summary and comments .............................................................................................. 41
  2.6 References ..................................................................................................................... 44
Chapter 3 Development of multiscale databases ............................................................... 50

3.1 Introduction ................................................................................................................ 50

3.2 Study areas ................................................................................................................ 50

3.3 Remotely sensed data ................................................................................................ 51

  3.3.1 TM data ............................................................................................................. 51

  3.3.2 SPOT data ......................................................................................................... 55

  3.3.3 AVHRR data ..................................................................................................... 58

3.4 Ancillary data ............................................................................................................. 61

  3.4.1 Ecoregion map ............................................................................................... 61

  3.4.2 Other ancillary data ....................................................................................... 62

3.5 Land cover classification ........................................................................................... 63

  3.5.1 TM land cover classification and labeling ................................................... 63

  3.5.2 SPOT land cover classification ..................................................................... 64

  3.5.3 AVHRR land cover classification ................................................................. 65

  3.5.4 Accuracy assessment of the land cover classification ................................. 67

3.6 Creation of multiscale databases ............................................................................... 68

  3.6.1 The "averaging" upscaling procedure ........................................................... 68

  3.6.2 Upscaling of the TM data ............................................................................. 70

  3.6.3 Upscaling of SPOT data ................................................................................. 71

  3.6.4 Upscaling of AVHRR data ............................................................................. 71

  3.6.5 Upscaling of land cover data ........................................................................ 72

3.7 Summary .................................................................................................................... 74
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.8 References</td>
<td>76</td>
</tr>
<tr>
<td>Chapter 4 Spatial upscaling of the Normalized Difference Vegetation Index</td>
<td>102</td>
</tr>
<tr>
<td>4.1 Introduction</td>
<td>102</td>
</tr>
<tr>
<td>4.2 Data and method</td>
<td>105</td>
</tr>
<tr>
<td>4.3 Results and analyses</td>
<td>110</td>
</tr>
<tr>
<td>4.3.1 Difference in NDVI values upscaled using the two methods</td>
<td>110</td>
</tr>
<tr>
<td>4.3.2 Difference in NDVI values upscaled using the two methods</td>
<td>111</td>
</tr>
<tr>
<td>4.3.3 Relationship between NDVI difference and land cover</td>
<td>113</td>
</tr>
<tr>
<td>4.3.4 Change in image information content</td>
<td>115</td>
</tr>
<tr>
<td>4.4 Summary and conclusions</td>
<td>119</td>
</tr>
<tr>
<td>4.5 References</td>
<td>121</td>
</tr>
<tr>
<td>Chapter 5 Characterizing the effects of spatial resolution on landscape indices:</td>
<td>138</td>
</tr>
<tr>
<td>a simulation study</td>
<td>138</td>
</tr>
<tr>
<td>5.1 Introduction</td>
<td>138</td>
</tr>
<tr>
<td>5.2 Methods</td>
<td>141</td>
</tr>
<tr>
<td>5.3 Results and analyses</td>
<td>144</td>
</tr>
<tr>
<td>5.3.1 Effects of proportional foreground on landscape metrics</td>
<td>144</td>
</tr>
</tbody>
</table>
5.3.2 Integrated effects of spatial resolution and proportional area on landscape metrics .............................................................. 149

5.4 Summary and conclusions .......................................................................................................................................................... 155

5.5 References ................................................................................................................................................................................. 160

Chapter 6 Spatial upscaling of land cover databases .................................................................................................................. 191

6.1 Introduction ................................................................................................................................................................................ 191

6.2 Data and methods ...................................................................................................................................................................... 192

6.3 Results and analyses ................................................................................................................................................................. 196

6.3.1 SPOT-derived land cover in the Niobrara Preserve and vicinity ......................................................................................... 196

6.3.2 TM-derived land cover in the Lincoln/Omaha study area ................................................................................................. 198

6.3.3 AVHRR-derived 159-class land cover of the conterminous U.S. ..................................................................................... 200

6.3.4 AVHRR-derived 25-class land cover in the conterminous U.S. ......................................................................................... 203

6.3.5 Stratification by ecoregion ................................................................................................................................. 205

6.3.6 Comparison between the Kappa values and image information content ........................................................................ 208

6.4 Summary and conclusions ...................................................................................................................................................... 211

6.5 References ................................................................................................................................................................................. 216

Chapter 7 Spatial upscaling and land cover area estimation: linking land cover to landscape structure .............................................. 245

7.1 Introduction ................................................................................................................................................................................ 245
7.2 Data and methods ................................................................. 247

7.3 Results and analyses ............................................................. 253

7.3.1 Change in class area across resolutions: the Niobrara dataset ........ 253

7.3.2 Statistical analyses for the Niobrara 17 land cover classes .......... 257

7.3.3 Area change for the aggregated 25-class U.S. land cover data .... 260

7.3.4 Statistical analyses for the 25-class U.S. land cover data ......... 265

7.3.5 Statistical analysis for the 159-class U.S. land cover data ....... 268

7.3.6 Effects of upscaling techniques on land cover representation and on the relationship between land cover area and landscape structure indices .................................................. 271

7.3.7 Change of landscape level indices across resolution ............ 278

7.4 Summary and conclusion ...................................................... 282

7.5 References ........................................................................ 286

Chapter 8 Summary and recommendations for future research ........... 314

8.1 Uniqueness of the study ....................................................... 314

8.2 Spatial upscaling of NDVI ................................................... 315

8.3 Effects of spatial resolution on landscape structure indices ........ 317

8.4 Upscaling of land cover databases ....................................... 319

8.5 Relationships between land cover upscaling and landscape structure .... 321

8.6 Conclusions and recommendations for future research ............ 324

8.7 References ........................................................................ 329
LIST OF FIGURES

Figure 3.1 Locations of the Nebraska study areas ........................................................... 79
Figure 3.2 Pseudo color TM composite of the Lincoln/Omaha study area ...................... 80
Figure 3.3 TM NDVI image in the Lincoln/Omaha study area ....................................... 81
Figure 3.4 Pseudo color SPOT composite of the Niobrara study area (6/7/1994) ............ 82
Figure 3.5 Pseudo color SPOT composite of the Niobrara study area (8/24/1994) ........... 83
Figure 3.6 SPOT NDVI images in the Niobrara study area ........................................... 84
Figure 3.7 AVHRR Maximum Value Composite NDVI image ....................................... 85
Figure 3.8 Level I ecoregions of the conterminous U.S. ............................................... 86
Figure 3.9 TM-derived land cover map in the Lincoln/Omaha study area ....................... 87
Figure 3.10 SPOT-derived land cover map in the Niobrara study area ............................ 88
Figure 3.11 The 25-class AVHRR-derived conterminous U.S. land cover map ............... 89
Figure 3.12 The 240m resolution land cover map of the Lincoln/Omaha study area, following the aggregation method ................................................................. 90
Figure 3.13 The 480m resolution land cover map of the Lincoln/Omaha study area, following the aggregation method ................................................................. 91
Figure 3.14 The 960m resolution land cover map of the Lincoln/Omaha study area, following the aggregation method ................................................................. 92
Figure 3.15 The 240m resolution land cover map of the Niobrara study area ............... 93
Figure 3.16 The 480m resolution land cover map of the Niobrara study area .......... 94
Figure 3.17 The 960m resolution land cover map of the Niobrara study area .......... 95
Figure 3.18 The 16km resolution 25-class land cover map of
the conterminous United States .................................................................. 96
Figure 3.19 The 32km resolution 25-class land cover map of
the conterminous United States .................................................................. 97
Figure 3.20 The 64km resolution 25-class land cover map of
the conterminous United States .................................................................. 98
Figure 3.21 The 240m resolution land cover map of the Lincoln/Omaha
study area, following the re-classification method ....................................... 99
Figure 3.22 The 480m resolution land cover map of the Lincoln/Omaha
study area, following the re-classification method ....................................... 100
Figure 3.23 The 960m resolution land cover map of the Lincoln/Omaha
study area, following the re-classification method ....................................... 101
Figure 4.1 Coarse resolution NDVI upscaled using two different methods,
results of simulation .................................................................................... 123
Figure 4.2 Differences between the two NDVI values across the resolutions,
results of the Lincoln/Omaha TM image .................................................... 124
Figure 4.3 Difference image between the two TM NDVI at the 480m resolution ...... 125
Figure 4.4 Differences image the two SPOT NDVI at the 480m resolution .......... 126
Figure 4.5 Percentage of pixels with varying degree of Diff(NDVI) ..................... 127
Figure 4.6 Differences between the two NDVI values across the resolutions,
results of the Niobrara SPOT image ................................................................. 128

Figure 4.7 Differences between the two NDVI values across the resolutions,
results of the AVHRR image ........................................................................ 129

Figure 4.8 Percentage of pixels having greater than 0.01 Diff(NDVI)
changes with the number of land cover classes ........................................ 130

Figure 4.9 Changes in NDVI standard deviation and entropy across the
resolutions, the Niobrara SPOT image .......................................................... 131

Figure 4.10 Changes in NDVI standard deviation and entropy across the
resolutions, the Lincoln/Omaha TM image .................................................... 132

Figure 4.11 Changes in NDVI standard deviation and entropy across the
resolutions, the conterminous U.S. AVHRR image ......................................... 133

Figure 4.12 Comparison between the TM NDVI and the AVHRR MVC NDVI .... 134

Figure 4.13 Comparison between the SPOT NDVI and the AVHRR MVC NDVI .... 135

Figure 4.14 Regression between the SPOT NDVI and the
AVHRR MVC NDVI, June 1994 ...................................................................... 136

Figure 4.15 Regression between the SPOT NDVI and the
AVHRR MVC NDVI, August 1994 ................................................................. 136

Figure 4.16 Regression between the TM NDVI and the
AVHRR MVC NDVI, August 1991 ................................................................. 137

Figure 5.1 Simulated neutral landscape, circular unit patch (p=0.05, Res=30m) .... 163

Figure 5.2 Simulated neutral landscape, circular unit patch (p=0.20, Res=30m) .... 163

Figure 5.3 Simulated neutral landscape, circular unit patch (p=0.50, Res=30m) .... 164
Figure 5.4 Simulated neutral landscape, circular unit patch (p=0.05, Res=480m) 164
Figure 5.5 Simulated neutral landscape, circular unit patch (p=0.20, Res=480m) 165
Figure 5.6 Simulated neutral landscape, circular unit patch (p=0.50, Res=480m) 165
Figure 5.7 Simulated neutral landscape, square unit patch (p=0.05, Res=30m) 166
Figure 5.8 Simulated neutral landscape, square unit patch (p=0.20, Res=30m) 166
Figure 5.9 Simulated neutral landscape, square unit patch (p=0.50, Res=30m) 167
Figure 5.10 Simulated neutral landscape, square unit patch (p=0.05, Res=480m) 167
Figure 5.11 Simulated neutral landscape, square unit patch (p=0.20, Res=480m) 168
Figure 5.12 Simulated neutral landscape, square unit patch (p=0.50, Res=480m) 168
Figure 5.13 Simulated neutral landscape, dotted unit patch (p=0.05, Res=30m) 169
Figure 5.14 Simulated neutral landscape, dotted unit patch (p=0.20, Res=30m) 169
Figure 5.15 Simulated neutral landscape, dotted unit patch (p=0.50, Res=30m) 170
Figure 5.16 Simulated neutral landscape, dotted unit patch (p=0.05, Res=480m) 170
Figure 5.17 Simulated neutral landscape, dotted unit patch (p=0.20, Res=480m) 171
Figure 5.18 Simulated neutral landscape, dotted unit patch (p=0.50, Res=480m) 171
Figure 5.19 Change in landscape indices with proportional foreground area, CPLS 172
Figure 5.20 Change in landscape indices with proportional foreground area, SPLS 173
Figure 5.21 Change in landscape indices with proportional foreground area, DPLS 174
Figure 5.22 Change in largest patch index (LPI) across resolutions (CPLS) 175
Figure 5.23 Change in largest patch index (LPI) across resolutions (SPLS) 176
Figure 5.24 Change in largest patch index (LPI) across resolutions (DPLS) 177
Figure 5.25 Change in largest patch index (LPI) across P levels 178
Figure 5.26 Change in mean patch size (MPS) across resolutions (CPLS) .................... 179

Figure 5.27 Change in mean patch size (MPS) across resolutions (SPLS) .................... 180

Figure 5.28 Change in mean patch size (MPS) across resolutions (DPLS) .................... 181

Figure 5.29 Change in patch size standard deviation (PSSD) across resolutions (CPLS) ............................................................................. 182

Figure 5.30 Change in patch size standard deviation (PSSD) across resolutions (SPLS).......................................................................... 183

Figure 5.31 Change in patch size standard deviation (PSSD) across resolutions (DPLS) .......................................................................... 184

Figure 5.32 Change in landscape shape index (LSI) across resolutions (CPLS) ........... 185

Figure 5.33 Change in landscape shape index (LSI) across resolutions (SPLS) ........... 186

Figure 5.34 Change in landscape shape index (LSI) across resolutions (DPLS) .......... 187

Figure 5.35 Change in mean nearest neighbor distance (MNN) across resolutions (CPLS) ............................................................................. 188

Figure 5.36 Change in mean nearest neighbor distance (MNN) across resolutions (CPLS) ............................................................................. 189

Figure 5.37 Change in mean nearest neighbor distance (MNN) across resolutions (CPLS) ............................................................................. 190

Figure 6.1 The concepts of consistency rate, pixel heterogeneity, and degree of dominance ............................................................................. 218

Figure 6.2 Changes in consistency rate and Kappa statistics across resolutions, the SPOT-derived land cover data ............................................................................. 219
Figure 6.3  Heterogeneity of coarse resolution pixels, the SPOT-derived land cover data ......................................................... 220

Figure 6.4  Degree of dominance of the majority class, the SPOT-derived land cover data ....................................................... 221

Figure 6.5  Changes in consistency rate and Kappa statistics across resolutions, the TM-derived land cover data ..................... 222

Figure 6.6  Heterogeneity of coarse resolution pixels, the TM-derived land cover data .......................................................... 223

Figure 6.7  Degree of dominance of the majority class, the TM-derived land cover data .......................................................... 224

Figure 6.8  Changes in consistency rate and Kappa statistics across resolutions, the AVHRR-derived 159-class land cover .......... 225

Figure 6.9  Heterogeneity of coarse resolution pixels, the AVHRR-derived 159-class land cover ............................................... 226

Figure 6.10 Degree of dominance of the majority class, the AVHRR-derived 159-class land cover data ...................................... 227

Figure 6.11 Changes in consistency rate and Kappa statistics across resolutions, the AVHRR-derived 25-class land cover data ... 228

Figure 6.12 Heterogeneity of coarse resolution pixels, the AVHRR-derived 25-class land cover data ........................................ 229

Figure 6.13 Degree of dominance of the majority class, the AVHRR-derived 25-class land cover data ...................................... 230
Figure 6.14 Changes in consistency rate and Kappa statistics across resolutions stratified by ecoregions, the AVHRR-derived 159-class land cover ...... 231

Figure 6.15 Changes in consistency rate and Kappa statistics across resolutions stratified by ecoregions, the AVHRR-derived 25-class land cover ........ 232

Figure 6.16 The heterogeneity of coarse resolution pixels, the AVHRR-derived 159-class land cover in ecoregion #4 ........................................................ 233

Figure 6.17 Degree of dominance of the majority class, the AVHRR-derived 159-class land cover in ecoregion #4 ........................................................ 234

Figure 6.18 The heterogeneity of coarse resolution pixels, the AVHRR-derived 159-class land cover data in ecoregion #6 ................................................ 235

Figure 6.19 Degree of dominance of the majority class, the AVHRR-derived 159-class land cover data in ecoregion #6 ................................................ 236

Figure 6.20 The heterogeneity of coarse resolution pixels, the AVHRR-derived 25-class land cover in ecoregion #4 ................................................ 237

Figure 6.21 Degree of dominance of the majority class, the AVHRR-derived 25-class land cover data in ecoregion #4 ................................................ 238

Figure 6.22 The heterogeneity of coarse resolution pixels, the AVHRR-derived 25-class land cover data in ecoregion #6 ................................................ 239

Figure 6.23 Degree of dominance of the majority class, the AVHRR-derived 25-class land cover data in ecoregion #6 ................................................ 240

Figure 6.24 Change in AVHRR NDVI standard deviation with spatial resolutions in the four ecoregions ................................................ 241
Figure 6.25 Change in TM and SPOT NDVI standard deviation with spatial resolutions ................................................................................................... 242

Figure 6.26 Regression between Kappa statistics and NDVI standard deviation, AVHRR data .............................................................................................. 243

Figure 6.27 Regression between Kappa statistics and NDVI standard deviation, TM and SPOT data .................................................................................... 244

Figure 7.1 Changes in class area across resolutions, the SPOT-derived land cover data ........................................................................................... 288

Figure 7.2 Changes in class area across resolutions, the SPOT-derived land cover data ........................................................................................... 289

Figure 7.3 Changes in class area across resolutions, the SPOT-derived land cover data ........................................................................................... 290

Figure 7.4 Changes in class area across resolutions, the AVHRR land cover ........................................................................................... 291

Figure 7.5 Changes in class area across resolutions, the AVHRR land cover ........................................................................................... 292

Figure 7.6 Changes in class area across resolutions, the AVHRR land cover ........................................................................................... 293

Figure 7.7 Changes in class area across resolutions, the AVHRR land cover ........................................................................................... 294

Figure 7.8 Spatial distribution of the dryland cropland and woodland/cropland classes ........................................................................................... 295

Figure 7.9 Spatial distribution of the desert shrubland and shrubland/grassland classes ........................................................................................... 296

Figure 7.10 Consistency rates and Kappa statistics between fine and coarse resolution data and between data derived using different
upscaling techniques, results of the TM land cover data ......................... 297

Figure 7.11 Changes in class area across resolutions,
following the aggregation method ............................................................. 298

Figure 7.12 Changes in class area across resolutions,
following the aggregation method ............................................................. 299

Figure 7.13 Changes in class area across resolutions,
following the aggregation method ............................................................. 300

Figure 7.14 Changes in class area across resolutions,
following the aggregation method ............................................................. 301

Figure 7.15 Changes in class area across resolutions,
following the re-classification method ...................................................... 302

Figure 7.16 Changes in class area across resolutions,
following the re-classification method ...................................................... 303

Figure 7.17 Changes in class area across resolutions,
following the re-classification method ...................................................... 304

Figure 7.18 Changes in class area across resolutions,
following the re-classification method ...................................................... 305

Figure 7.19 Changes in LPI, MPS, PSSD, and LSI across resolutions, the
SPOT-derived land cover data ................................................................... 306

Figure 7.20 Changes in DLF, IJI, and CONTAG across resolutions, the
SPOT-derived land cover data ................................................................... 307
Figure 7.21 Changes in LPI, MPS, PSSD, and LSI across resolutions, the TM-derived land cover data ................................................................. 308

Figure 7.22 Changes in DLF, IJI, and CONTAG across resolutions, the TM-derived land cover data .............................................................................. 309

Figure 7.23 Changes in LPI, MPS, PSSD, and LSI across resolutions, the AVHRR-derived 159-class land cover data .............................................. 310

Figure 7.24 Changes in DLF, IJI, and CONTAG across resolutions, the AVHRR-derived 159-class land cover data .............................................. 311

Figure 7.25 Changes in LPI, MPS, PSSD, and LSI across resolutions, the AVHRR-derived 25-class land cover data ................................................ 312

Figure 7.26 Changes in DLF, IJI, and CONTAG across resolutions, the AVHRR-derived 25-class land cover data ................................................ 313
LIST OF TABLES

Table 1.1 Present and future earth observing satellites .................................................... 2
Table 1.2 Land cover characteristics input requirements and spatial scale for selected modeling applications and models .............................................................. 7
Table 3.1 Spectral and spatial resolutions of the TM sensor ............................................. 52
Table 3.2 Statistical characteristics of the TM subimage .................................................. 53
Table 3.3 Spectral and spatial resolutions of the SPOT sensor ........................................ 56
Table 3.4 Statistic characteristics of the SPOT image ...................................................... 57
Table 3.5 Spectral and spatial resolutions of the AVHRR sensor ................................... 59
Table 3.6 AVHRR NDVI data used in the study .............................................................. 61
Table 3.7 Level I ecoregions of the conterminous U.S. .................................................... 62
Table 3.8 Twenty-five grouped AVHRR-derived land cover class .................................. 66
Table 4.1 The range, mean, and standard deviation of the AVHRR MVC NDVI and upscaled SPOT and TM NDVI .............................................................. 117
Table 5.1 Regression between LSI and $P_i$ for the three landscapes ......................... 147
Table 5.2 Regression between MNN and the reciprocal of $P_i$ for the three landscapes ........................................................................................................... 148
Table 5.3 Regression between LSI and spatial resolution for the CPLS landscape at selected $P_i$ levels ...................................................................................... 153
Table 5.4 Regression between LSI and spatial resolution for the SPLS landscape
Table 5.5 Regression between LSI and spatial resolution for the DPLS landscape

Table 6.1 Range of Kappa statistics and degree of agreement

Table 6.2 Kappa values between full and coarse resolution land covers

Table 6.3 Difference in Kappa values between the 25-class and the 159-class land cover averaged from the four ecoregions in the conterminous United States

Table 6.4 Differences in Kappa values among the four ecoregions, averaged from pairwise comparisons

Table 6.5 Summary of linear regression between Kappa values and NDVI standard deviations for the 159-class AVHRR dataset

Table 6.6 Summary of linear regression between Kappa values and NDVI standard deviations for the 25-class AVHRR dataset

Table 6.7 Summary of linear regression between Kappa values and NDVI standard deviations for the TM dataset in Lincoln/Omaha area and for the SPOT dataset in the Niobrara study area

Table 7.1 Data organization for regressions of coarse resolution class areas and proportional errors against landscape indices, original area, and spatial resolution

Table 7.2 Neighboring relationship among different land cover types in the SPOT-derived land cover in the Niobrara study area

Table 7.3 Mixture matrix between the 20m and 960m SPOT-derived
Table 7.4 Mixture matrix between the 20m and 960m SPOT-derived land cover data ................................................................. 255

Table 7.5 Pairwise Pearson-product-moment correlation among class level landscape indices. SPOT-derived land cover data in the Niobrara study area ........................................... 256

Table 7.6 Regression summary for the forward multi-linear regression between the nine predictor variables and coarse resolution class areas. SPOT-derived land cover data in the Niobrara study area ........................................................................ 257

Table 7.7 Regression summary for the forward multi-linear regression between the nine predictor variables and proportional errors. SPOT-derived land cover data in the Niobrara study area ........................................................................ 258

Table 7.8 Neighboring relationship among different land cover types in the 25-class AVHRR-derived land cover in the conterminous U.S. ........................................... 259

Table 7.9 Mixture matrix between the 1km and 64m AVHRR-derived land cover data ........................................................................ 260

Table 7.10 Pairwise Pearson-product-moment correlation among class level landscape indices. AVHRR-derived 25-class land cover data ........................................................................ 261

Table 7.11 Regression summary for the forward multi-linear regression between the nine predictor variables and coarse resolution class areas, AVHRR-derived 25-class land cover data ........................................................................ 262

Table 7.12 Regression summary for the forward multi-linear regression between the nine predictor variables and proportional errors, AVHRR-derived land cover data ........................................................................ 263
Table 7.13 Pairwise Pearson-product-moment correlation among class level landscape indices. AVHRR-derived 159-class land cover data ........................................... 268

Table 7.14 Regression summary for the forward multi-linear regression between the nine predictor variables and coarse resolution class areas, AVHRR-derived 159-class land cover data ........................................................................ 269

Table 7.15 Regression summary for the forward multi-linear regression between the nine predictor variables and proportional errors, AVHRR-derived 159-class land cover data ........................................................................ 270

Table 7.16 Pairwise Pearson-product-moment correlation among class level landscape indices, TM-derived land cover data ........................................................................ 273

Table 7.17 Regression summary for the forward multi-linear regression between the nine predictor variables and coarse resolution class areas, TM-derived land cover data, following the aggregation method ........................................... 274

Table 7.18 Regression summary for the forward multi-linear regression between the nine predictor variables and proportional errors, TM-derived land cover data, following the aggregation method ........................................... 275

Table 7.19 Regression summary for the forward multi-linear regression between the nine predictor variables and coarse resolution class areas, TM-derived land cover data, following the re-classification method ........................................... 277

Table 7.20 Regression summary for the forward multi-linear regression between the nine predictor variables and proportional errors, TM-derived land cover data, following the aggregation method ........................................... 278
land cover data, following the re-classification method ........................................... 277

Table 7.21 Regression between LPI and spatial resolution ........................................... 281
Table 7.22 Regression between MPS and spatial resolution ........................................... 281
Table 7.23 Regression between PSSD and spatial resolution ....................................... 281
Table 7.24 Regression between LSI and spatial resolution .......................................... 282
Table 7.25 Regression between CONTAG and spatial resolution ................................. 282
CHAPTER 1

INTRODUCTION

1.1. Introduction

Information on land cover is critical for dealing with a broad spectrum of issues such as global warming, agricultural sustainability, land transformation and biodiversity. Ecological, hydrologic and climate models are usually dependent on accurate representation of land cover. Today most mapping, characterization and monitoring of the earth’s land cover, especially over large regions or the globe, is dependent on the use of remotely sensed data. At least ten satellites routinely collect imagery of the earth and its environs, and during the next fifteen years more than one dozen additional earth observing satellites are expected (Table 1.1).

There are, however, frequently considerable disparities between the levels of spatial detail collected by remote sensing systems and those required for various ecoclimatological modeling and land management applications. It is, therefore, often necessary to spatially transform satellite-derived data, either by generalizing finer measurements to coarser levels or by extracting detailed information from coarse-resolution satellite observations, so that land cover information can be used in models or to address other environmental issues. While it is clear that such spatial transformations can significantly affect the quality of information, the consequences of applying specific transformation approaches and methods in different landscapes are not well understood.
| Table 1.1. Present and Future Earth Observing Satellites*

**NASA Satellites**
- Earth Radiation Budget Satellite (ERBS)
- Upper Atmosphere Research Satellite (UARS)
- The Ocean Topography Experiment (TOPEX/Poseidon)
- Earth Probe Total Ozone Mapping Spectrometer (TOMS)
- The Landsat Program
- The Tropical Rainfall Measuring Mission (TRMM)
- Earth Observing System
  - (EOS-AM1, EOS-PM1, EOS-CHEM1, EOS-LASER ALT, EOS-AM2, EOS-PM2)

**Non-NASA US Satellites**
- NOAA's Polar Orbiting Satellite (NOAA 9, 10, 11, 12, 14)
- Geostationary Observational Environmental Satellite (GOES)
- Defense Meteorological Satellite (DMSP)
- Global Ocean Color Monitoring Mission’s SeaWiFS/SeaStar

**International Satellites**
- European Remote Sensing Satellite-1 (ERS-1) (European Space Agency)
- Meteosat (European Space Agency)
- Japanese Earth Resources Satellite-1 (JERS-1)
- GMS (Japan)
- SPOT (France)
- Meteor (Russia)
- Resors (Russia)
- Sich/Okean (Ukraine/Russia)
- MIR-Priroda (Russia)
- Envisat (European Space Agency)
- ERS-2 (European Space Agency)
- The Advanced Earth Observing Satellite (ADEOS) (Japan)
- Radarsat (Canada)

* From NASA’s Mission to Planet Earth Web Site (http://www.hq.nasa.gov/office/mtpe/spacecraft.html).
1.2. Background

1.2.1. Scale

The biosphere is extremely complex and is organized across a wide range of spatial and temporal dimensions. For example, studies of trace gas emission may extend over a few meters at local sites to tens or hundreds of kilometers for regional and global investigations, while such studies may encompass temporal scales ranging from minutes to months or years. Climatic processes can be examined in a single agricultural field of less than one square kilometer as well as at regional or global scales. Because of the great differences among the spatial and temporal scales, the scales of land cover data needed to model biospherical environments and processes vary significantly. Inconsistencies between the intrinsic scale of the biophysical process being studied and the spatial and temporal intervals of land cover observation can be sources of errors. Depending on the subject under investigation, data needed for studies of the biosphere may include measurements of pollen counts or soil particle attributes obtained through electron microscopes, or of radiation energy levels for different land cover types observed from satellites over thousands of square kilometers.

Scale and scaling (i.e., change of scale) issues are relevant to most scientific research activities, but have received special attention in geography (Harvey, 1969; Meentemeyer, 1989). Unfortunately, there remains substantial ambiguity in using the term scale (Woodcock and Strahler, 1987). Lam and Quattrochi (1992) gave three basic meanings of scale. The term "geographic (or observation) scale" is used to denote the
overall areal extent of a study. For example, land cover mapping for a specific urban area or wetland is usually referred to as "small scale", as compared to the mapping of vegetation for the conterminous United States or a whole continent (e.g., Goward et al., 1985; Loveland et al., 1991), which is a "large scale" study. "Cartographic scale" is defined as the ratio of the distance measured between two points on a map to the actual ground distance between the corresponding points. In this connotation, a "large scale" map covers a small area but with more detailed information, while a "small scale" map covers a large area with less detail. Therefore, a large (regional, continental) scale land cover mapping project usually results in a map having small cartographic scale, while a small scale (local) investigation of land cover results in a map having large cartographic scale. Here both usages of the term scale imply the level of observation needed for a study, since large scale (or, small cartographic scale) studies usually require relatively coarse observations (i.e., larger sampling grid size), while small scale (or, large cartographic scale) studies having more detail require finer observations (e.g., smaller sampling grid size). The third usage of scale refers to the spatial extent at which a particular biophysical process operates. For example, local climatic phenomena may occur in an area of a few square kilometers while upper air circulation occurs around the whole globe.

1.2.2. Scale and spatial resolution

In remote sensing spatial scale is usually equated with spatial resolution. Strictly
speaking, spatial resolution refers to a sensor's effective instantaneous field of view (IFOV). In most cases, however, spatial resolution is described by image pixel size which is the size of an object on the ground, determined by the sensor's IFOV (Forshaw et al. 1983; Woodcock et al. 1988a and b; Belward, 1992). The higher the resolution (i.e., smaller the pixel size) of an image, the more ground detail it provides.

Spatial resolution is also generally related to the overall areal coverage of an image. A coarse resolution image usually covers a larger area than a fine resolution scene. For example, an image obtained by the Advanced Very High Resolution Radiometer (AVHRR) at 1.1 km resolution covers about 2400 by 2400 square kilometers. A Landsat Multispectral Scanner (MSS) image having 79m resolution covers an area of 180 by 180 square kilometers, while a SPOT image at 20m resolution covers an area of 60 by 60 square kilometers. Therefore, the relationship between spatial resolution and the geographic scale (Cao and Lam, 1997) is straightforward. A large scale study (e.g., continental-level vegetation monitoring) usually dictates the use of coarse resolution images while a small scale study (e.g., city land use mapping) usually demands fine resolution images.

1.2.3. Scaling of remotely sensed data

In most scientific research, the investigators select the scale of measurement or observation that corresponds to that of the phenomena being studied. However, remotely sensed data are only obtained at a few pre-selected spatial resolutions and these are
unable to satisfy users' diverse scale requirements from local to regional and global studies. Thus, remotely sensed data are often required with spatial resolutions different from their original ones and the creation of multi-scale data sets is necessary (Justice et al., 1989). Loveland et al. (1995) discussed the scale requirements for a number of climatological, hydrological and ecosystems models (Table 1.2). Substantial spatial transformation is required for remotely sensed data to be used in such models.

The process of scale change is extremely complex. Gardner et al. (1989) classified the response of ecological process to changes in scale into four conditions: 1) the process is scale invariant, in which case no transformations are required; 2) a process is similar in its effect across scales and scale transformations are simple; 3) the process varies little and remains dominant across scales, but additional factors and constraints increase uncertainty of the prediction of the process; and 4) the importance of the process or the process constraints changes with scale. In the last two cases, inappropriate integration of local processes may introduce significant errors in the estimation of regional processes.

One of the major factors that constrains our ability to translate information from one scale to another is spatial heterogeneity (Wessman, 1992). Scaling problems may not occur in spatially homogeneous landscapes because information can be integrated directly to coarser resolution (upsampling) or can be inferred at finer resolution (downsampling). In heterogeneous landscapes, however, measurements obtained at fine resolution often cannot be summed directly to produce regional estimates and those acquired at coarse resolution cannot be decomposed into finer resolution without additional information.
Table 1.2. Land-Cover Characteristics Input Requirements and Spatial Scale for Selected Modeling Applications and Models (From Loveland et al., 1995).

<table>
<thead>
<tr>
<th>Classification Scheme</th>
<th>Model</th>
<th>Spatial Scale</th>
<th>Associated Attributes</th>
</tr>
</thead>
<tbody>
<tr>
<td>General Circulation Models</td>
<td>NASA/GSFC SiB</td>
<td>4x5 degrees</td>
<td>SiB set and NDVI derivatives</td>
</tr>
<tr>
<td></td>
<td>University of Maryland-COLA Simplified SiB</td>
<td>4.5x7.8 degrees</td>
<td>SSiB set and derivatives</td>
</tr>
<tr>
<td></td>
<td>NCAR-CCM BATS</td>
<td>2x4 degrees</td>
<td>BATS set and NDVI derivatives</td>
</tr>
<tr>
<td>Mesoscale Meteorological Models</td>
<td>CSU-RAMS LEAF Nested Grids of 1, 10, 40 km</td>
<td>LEAF Set and NDVI derivatives</td>
<td></td>
</tr>
<tr>
<td></td>
<td>PSU-NCAR MM4 BATS</td>
<td>Nested Grids of 4, 12, 36 km</td>
<td>BATS Set and NDVI derivatives</td>
</tr>
<tr>
<td>Hydrologic Models</td>
<td>Watershed-Precipitation/Runoff Basic Classes</td>
<td>2.5, 5, 10 km</td>
<td>model specific</td>
</tr>
<tr>
<td>Agricultural Chemical/Runoff Anderson Level II</td>
<td>country level or 1 km</td>
<td>model specific</td>
<td></td>
</tr>
<tr>
<td>Ecosystem Models</td>
<td>RHESSys Basic Biomes</td>
<td>1-50 km</td>
<td>RHESSys Set and NDVI derivatives</td>
</tr>
<tr>
<td>CENTURY Anderson Level II</td>
<td>1-50 km</td>
<td>NDVI derivatives</td>
<td></td>
</tr>
<tr>
<td>Biogenic Emissions</td>
<td>Key species (oak, hickory, etc.)</td>
<td>20 km</td>
<td>NDVI derivatives</td>
</tr>
</tbody>
</table>

Table of Abbreviations
- BATS: Biosphere-Atmosphere-Transfer-Scheme
- COLA: Center for Ocean-Land-Atmosphere
- CSU-RAMS: Colorado State University-Regional Atmosphere Modeling System
- GSFC: Goddard Space Flight Center
- LEAF: Land-Ecosystem-Atmosphere-Feedback
- NCAR-CCM: National Center for Atmospheric Research, Climate Community Model
- NDVI: Normalized Difference Vegetation Index
- PSU-NCAR-MM4: Penn State University/National Center for Atmospheric Research-Mesoscale Meteorology
- RHESSys: Regional Hydrological Ecosystem Simulation System
- SiB: Simple Biosphere Model
- SSiB: Simplified Simple Biosphere Model
Considerable research has been directed at issues related to downscaling (e.g., Marsh et al., 1980; Jasinski and Eagleson, 1990; Cross et al., 1991; Quarmby et al., 1992; Foody and Cox, 1994; Schneider, 1994). Although an optimum solution to downscaling of a mixed-pixel is possible when certain knowledge (e.g., number of possible land cover classes) about the pixel can be obtained from other sources, there has been no systematic study with respect to spatial and spectral resolution (Raffy, 1994). Upscaling may seem easier because the task is simply to aggregate or summarize data from a finer to a coarser resolution. But this seemingly simple problem turns out to be not easy at all because most upscaling involves heterogeneous landscapes. Spatial aggregation of remote sensing information, either raw data such as Digital Number (DN) values and reflectances or derived measurements/products such as vegetation indices and land cover classes, are related to landscape heterogeneity in complex ways. Different coarse resolution products (images, maps, etc.) may result from the same fine resolution image, depending on the approaches used in aggregation and the spatial characteristics of the area under investigation. There are usually no simple relationships, not to mention linearity, between the source remotely sensed data and final products such as land cover maps. The problem of scale change in remote sensing is relevant to many specific applications, but it remains one of the main challenges of data interpretation (Malingreau and Belward, 1992; Raffy, 1993, 1994).
1.3. Definition of terms

Because terminology related to scaling issues is sometimes ill-defined, in this dissertation the following definitions are used:

1) **Scale**: the term 'scale' refers to the spatial resolution of remotely sensed imagery. Scale, spatial resolution, and pixel size are used synonymously in the context of this dissertation.

2) **Scaling, rescaling, upscaling, and downscaling**: scaling and rescaling are used to indicate processes for changing the spatial resolution of raw remotely sensed data (i.e., DN value) and their derivatives (e.g., vegetation index, land cover). Upscaling refers to change from a fine to coarse resolution. Downscaling is defined as the extraction of sub-pixel information from coarse resolution pixels.

3) **Fine (or high) resolution, coarse (or low) resolution, and full resolution**: fine (or high) resolution images have pixel sizes smaller than 100 meters while coarse (or low) resolution images are those having pixel sizes of 1km or larger. Full resolution images are images that have not been spatially rescaled. It should be pointed out that, in many instances, "original data" have been spatially resampled during geometric rectification. Here the "full resolution" data refers to those original data which have not undergone spatial upscaling. For example, the 1km resolution AVHRR data provided by the United States Geological Survey (USGS) EROS Data Center is considered as "full resolution" even though the data were actually resampled from raw observations. While the terms "fine" and "coarse" are used to indicate resolutions, the words "finer" and "coarser" are
frequently used to refer to relative orders of resolutions. For example, although 4km resolution is much coarser than 30m resolution, it is finer than 8km resolution.

4) **Grouping and aggregation**: 'grouping' means to coalesce a number of detailed land cover classes into a more general class, e.g., grouping "corn", "sorghum" and "soybean" classes into a single "crop" class. The term aggregation refers to change of a land cover map at a finer resolution into a coarser resolution. It is one approach of upscaling land cover data.

5) **Landscape**: a landscape is composed of land cover patches. A patch is defined as a cluster of neighboring pixels that are classified as a single land cover type. Therefore, a landscape refers to a mosaic of various types of land cover in a study area.

6) **Landscape structure**: landscape structure refers to the spatial distribution and configuration of the patches in a landscape. Landscape structure can be quantitatively described by "landscape structure indices" which quantify spatial relations among patches (e.g., patch size, patch shape, patch interspersion, and distance between patches).

7) **Land cover characterization**: Land cover characterization means to identify land cover according to its type, spatial pattern, physiognomy and seasonal character. In this dissertation, the term refers only to land cover type and its spatial pattern.

8) **Image information content**: the information content of an image refers to the values of individual pixels and their spatial arrangements. The pixel value can be DN value, reflectance, or a vegetation index. Image information content can be measured in several ways, including variance, entropy, contrast, etc.
1.4. Problem statement and research objectives

Many questions related to spatial scaling problems in land cover characterization remain unanswered. This research addresses the following questions:

1. What are the effects of spatial scaling on image information content? More specifically, are images obtained by sensors operating at a particular resolution (e.g., 1km) comparable, in terms of information content, to those spatially rescaled to that resolution?

2. How do indices of landscape structure change when spatial resolution changes and what are the implications of such change for land cover characterization?

3. How well is land cover information maintained when it is spatially upscaled to coarse resolutions and what are the influencing factors on spatial upscaling?

4. Are there significant relationships between land cover areas and landscape metrics across spatial resolutions? If such relationships exist, are they extendible through different datasets and across different areas?

5. What are the impacts of different scaling approaches on the representation of land cover and how do these relate to the spectral and spatial characteristics of the landscape under investigation?

The principal objective of this research is to characterize and explain the effects of changing spatial scale on the representation of land cover in relation to landscape structure and methods of upscaling. Specific objectives were developed corresponding to the aforementioned five research problems. They included:

1. Exploration of the relationship between two different scaling procedures applied
to the Normalized Difference Vegetation Index (NDVI) in relation to landscape heterogeneity.

Different methods can be used when fine resolution data is to be upscaled to a coarser dataset (see Justice et al., 1989). NDVI is used, almost without exception, in regional scale land cover characterization (e.g., Justice et al., 1991; Loveland et al., 1991). The basis for using NDVI to map land cover is the fact that healthy green vegetation reflects strongly in the near-infrared wavelengths of the spectrum due to internal mesophyll structure and low in the red wavelengths because of strong absorption by leaf chlorophyll and other pigments (Sellers, 1985, 1989; Sellers et al., 1992). The relationship between NDVI and vegetation biophysical parameters is observed, however, at leaf or canopy levels. To apply this relationship to coarse resolution remote sensing, scaling problems of two types need to be examined. One is the scaling of the relationship between NDVI and the red and near-infrared measurements obtained by a remote sensor. The other is the scaling of the relationship between NDVI and vegetation biophysical parameters. NDVI is apparently a non-linear transformation of reflectance and different scaling approaches will result in different NDVI values. A number of investigations have focussed on the non-linearity of NDVI (e.g., Jasinski, 1990; Aman et al., 1992, Friedl et al., 1995; De Cola, 1997). However, additional studies have been suggested (Aman et al., 1992; De Cola, 1997). In this study, I estimate the mean error caused by the non-linearity of NDVI, and link the error to spatial heterogeneity and spectral properties of the landscape. I also investigate the comparability, in terms of image information content, between NDVI upscaled from fine resolution sensors to that obtained directly from coarse
resolution sensors.

2. Examination of the effects of spatial resolution on landscape structure indices and the implications for land cover characterization.

Recent studies in land cover scaling issues have shown that errors in land cover area estimation at coarse resolution are related to the spatial pattern of the landscape under investigation (e.g., Woodcock and Strahler, 1987; Townshend and Justice, 1988; Moody and Woodcock, 1994). Attempts have been made to calibrate the areal errors in coarse resolution land cover maps, either upscaled from fine resolution maps or directly classified from coarse resolution data, using models based on landscape structure indices (e.g., Mayaux and Lambin, 1995; Moody and Woodcock, 1995). However, landscape patterns are often determined from maps classified using remotely sensed data (e.g., Simmons et al., 1992; Wickham and Ritters, 1995; Frohn, 1996), and thus the landscape indices themselves are influenced by the observation scale. For example, to estimate areal errors of an AVHRR-derived land cover map, one may use landscape indices derived from finer resolution images (e.g., SPOT, TM). Due to the scale dependency of the landscape indices, the estimations resulting from using indices measured from different images will be different. Therefore, as a first step, the relationship between landscape indices and spatial scale, which is not well understood (Meentemeyer and Box, 1987), must be examined.

3. Investigation of the relationships (i.e., agreements) between full resolution land cover dataset and those spatially upscaled from the full resolution dataset, and examination of the factors that affect such relationships.
Land cover information is needed in a wide variety of applications including climate and ecological modeling and resources assessment. Spatial resolutions (model grid cells) of such applications are often much coarser than those of remote sensing images (Loveland et al., 1995). Consequently, upscaling of land cover is needed in order for data to be compatible with global models (DeFries et al., 1997). Spatially upcaled land cover data, however, may not agree well with the original data and may introduce varied degrees of biases in modeling results (e.g., Oleson et al., 1996; DeFries et al., 1997). It is expected that the relationships between upscaled land cover data and the full resolution data are dependent on a number of factors including characteristics of the study area and classification schemes. No systematic investigation into such relationships, however, was found in the literature. In this research, I examine the degrees of agreement between full and upscaled land cover data using data derived from different satellites, in different ecological areas, and with different classification schemes. The relationships between changes in degree of agreement and in image information content (quantified by NDVI standard deviation) across resolutions are also explored.

4. Examination of the relationships between land cover areas estimated from upscaled land cover data and landscape structure metrics, and the extendibility of such relationships across different areas and datasets.

The presence and areal extent of a land cover type is related to spatial resolution as well as to the landscape properties. As resolution coarsens, land covers having small patch sizes tend to be lost. However, if the small patches are closely clumped, they may still be retained (e.g., Turner et al., 1989). Change of proportional areas of different land
cover types is closely related to the patch sizes and connectivity between patches (Moody and Woodcock, 1994). Recent studies have shown that regression models between proportional areal errors contained in the coarse resolution land cover and landscape structure metrics existed in specific study areas (Mayaux and Lambin, 1995; Moody and Woodcock, 1995). The operational use of the regression models, however, requires that they be invertible and generalizable across different areas. Additional research was suggested to explore the extendibility of such relationships (Moody and Woodcock, 1994, 1995). I investigate changes in land cover areas across a wide range of spatial resolutions and their associations with landscape structures in different areas. Statistical relationships between land cover area and landscape metrics are explored and the similarity/dissimilarity in the relationships derived from different datasets are examined.

5. Investigation of the impacts of different upscaling approaches on the representation of land cover across resolutions.

It has been shown that the nonlinear relationships between remote sensing measurements, such as red and near-infrared reflectances, and vegetation biophysical variables, such as fractional photosynthetically active radiation, will cause uncertainties in scaling (Friedl et al., 1995). In land cover characterization, there are no simple relationships between the source data (images) and the final product (land cover maps). Thus, different approaches used to rescale a land cover map from one resolution to another are likely to produce different results. Most previous studies of scale effects on land cover were based on land cover maps aggregated from fine resolution products (e.g., Moody and Woodcock, 1994). However, coarse resolution maps can also be obtained by
performing upscaling before classification, which is more similar to real applications
where coarse resolution land covers are often classified directly from images obtained by
coarse resolution sensors such as AVHRR. In this dissertation, I investigate the
relationships between the land cover maps obtained by different approaches and link such
relationships to the characteristics of the landscape under investigation.

The fulfillment of the objectives of the research will improve the understanding
of scaling effects on land cover representation and will contribute to the creation of
better land cover databases across a broad range of spatial scales.

1.5. Structure of the dissertation

Following a review of literature in chapter 2, the strategies and methodologies of
establishing multi-source, multi-resolution, and multi-product databases are discussed in
Chapter 3. The databases include original full resolution data as well as derived data
(e.g., NDVI and land cover) at resolutions ranging from 20m to 64km. Research
Objective 1, examination of image information content and NDVI scaling issues, is
developed in Chapter 4. In Chapter 5, I examine the relationships between landscape
structure and spatial resolution. Chapter 6 deals with the representation of land cover at
various resolutions. The last two research objectives, land cover area/landscape structure
relationships and the impact of upscaling approaches, which share many processing
procedures and results, are discussed in Chapter 7. Chapters 4 to 7 each include sections
on data and methodology, results and analyses, summary and conclusions, and literature
cited. Chapter 8, the concluding chapter, is a review of results and findings of this research and recommendations for future studies.
1.6. References


Gardner, R. H., O’neill, R. V., Turner, M. G., Dale, V. H., 1989, Quantifying scale-


Meentemeyer, V., 1989, Geographical perspectives of space, time, and scale, *Landscape


Simmons, M. A., Cullinan, V. I., and Thomas, J. M., 1992, Satellite imagery as a tool to


CHAPTER 2
LITERATURE REVIEW

2.1. Introduction

Land cover information is needed for a broad range of research and application activities such as climate and ecological modeling, agricultural and forest monitoring, and resources management. Land cover mapping and monitoring is one of the most important applications of remote sensing. Widespread use of remotely sensed data in land cover mapping stems in part from the fact that land cover information can be interpreted more or less directly from evidence visible on images and relatively little inference is required in many situations (Campbell, 1987). This, however, does not indicate that accurate land cover mapping is an easy task.

Spatial resolution is a major factor influencing the ability to map and characterize land cover. There has been considerable research on the assessment of spatial resolution of remotely sensed data required for mapping and monitoring the land surface (e.g., Sadowski et al., 1977; Forshaw et al., 1983; Cushnie, 1987; Weiler and Stow, 1991; Davis et al., 1992; Marceau et al., 1994a, 1994b; Moody and Woodcock, 1994; Mayaux and Lambin, 1995; Moody, 1996). Research in this field has been focussed in several areas. The first is direct investigation into the effects of spatial resolution on classification accuracy, usually in relatively small and well-known areas. The second focus has been on effects of spatial rescaling on land cover mapping. A third emphasis
has been on examination of relationships between spatial scale and landscape structure and the roles of such relationships in land cover mapping. Since this dissertation deals primary with effects of spatial scaling on land cover characterization, the review of literature is focused on the latter two areas.

2.2. Classification accuracy and spatial resolution

Most early studies of the influences of spatial resolution on land cover classification accuracy involved use of high resolution data (airborne cameras, scanners) to simulate datasets with various spatial resolutions (e.g., Simonett and Coiner, 1971; Kan et al., 1975; Clark and Bryant, 1977; Craig and Labovitz, 1980; Markham and Townshend, 1981; Wrigley et al., 1984). The resolutions under investigation were usually finer than that of the first generation Landsat Multispectral Scanner (MSS) image, whose pixel size is 79m. In those studies, accuracies of classification using data of various spatial resolutions for different land cover/land use types were taken as criteria to select suitable resolutions. Because the spatial resolutions of the airborne data used in those studies were usually much smaller than the intrinsic scale of the subject under investigation, an increase in resolution cell size tended to decrease within-class variances due to the averaging effects and thus resulted in a high degree of homogeneity for each class. The reduction in within-class variation usually compensate for the spectral mixture at the boundary of different classes. As a result, improvements in classification accuracy with coarsening resolution were usually observed (e.g., Simonett and Coiner, 1971;
Markham and Townshend, 1981). Further research focused on the selection of appropriate resolutions for land cover mapping showed that, due to the complex nature of land surface, there seems no unique spatial resolution appropriate for the mapping of all geographical entities (Justice et al., 1991; Marceau et al., 1994a). The optimal spatial resolution is closely linked to the structure of the landscape under investigation (Woodcock and Strahler, 1987; Marceau et al., 1994b).

In an empirical investigation of the effects of spatial resolution on forest classification, Sadowski et al. (1977) degraded airborne multispectral scanner data from 2x2 meters progressively to 64x64 meters, resulting in six datasets with successively coarser resolutions. Supervised classifications were performed on each of the datasets and the accuracies of the classifications were analyzed. The results showed that overall performance in classification improved as spatial resolution was degraded and that classification accuracy increased for hierarchies of more general (grouped) forest features. The improvement in performance at coarse resolution was believed to have resulted from reduction in scene variation. Latty and Hoffer (1981) conducted a study using Thematic Mapper Simulator (TMS) data. Classifications were performed on four datasets with different resolutions and ten land cover classes. Results showed that higher spatial resolution data produced lower overall classification accuracy and this phenomenon was particularly pronounced in forest cover types, which displayed relatively large levels of spectral variation. Similar relationships among spatial resolution, scene variance, and classification accuracy were found by other investigators (e.g., Cushnie, 1987).

While such studies demonstrated that classification results were associated with
scene variance (which decreased with coarsening resolution), Markham and Townshend (1981) proposed more complicated inter-relationships among various factors that may affect classification accuracy. Their study showed that scene variance varied significantly between land cover categories and between spectral bands for individual land cover types. The effects of scene variance on classification accuracy were dependent on, in a complex way, the relative locations of cover types within the feature space. The increasing convergence of a cover type in feature space contributed to better classification performance in coarse resolution datasets. However, mixed-pixels on the boundary in coarse resolution datasets would reduce the classification accuracy and thus offset the benefit of lower scene variance.

A different method was proposed by Woodcock and Strahler (1987) to examine the effects of spatial resolution and to select appropriate resolutions for specific applications. The authors studied the relationship between image information content and spatial resolution, instead of classification accuracy. Local variance was used to explore such relationships for three different environments: forest, urban and agricultural land. They degraded digitized aerial photographs from $0.75 \, \text{m}$ resolution to $50 \, \text{m}$ and TM images from $30 \, \text{m}$ to $1 \, \text{km}$, creating datasets with different spatial resolutions. Changes in image local variance along with spatial resolution were analyzed. Different relationships were found for each environment due to the difference in surface structure and size of ground objects. Local variance tended to increase as the resolution became coarser and reached a maximum where the resolution size was close to, but smaller than, the size of the objects in the scene. It decreased as resolution further coarsened to include
many ground objects. The pattern of change in local variance reflected the high and low resolution models in remote sensing proposed by Strahler et al. (1986). The study showed that image information content was directly associated with the relationship between the size of the objects in the scene and the spatial resolution of the sensor. The pattern of the change in local variance with resolution helped explain the results found by other investigators (e.g., Latty and Hoffer, 1981; Cushnie, 1987) in the study of the effect of spatial resolution on classification accuracy.

Marceau et al. (1994a) used high resolution airborne MEIS (Multi-detector Electro-optical Imaging Scanner)-II data to investigate the effects of spatial resolution on classification accuracy and image information content in a mid-latitude temperate forest environment. The authors derived four datasets having resolutions ranging from 5m to 30m from the original airborne data. Classifications of the datasets were performed for different levels of forest classes and accuracy assessments were conducted statistically. They found a close relationship between spatial scale and the internal variance of the spectral classes when the best classification accuracies were achieved. Logically, there exists an optimal spatial resolution for each entity of interest, corresponding to its intrinsic spatial and spectral characteristics. However, due to the complex mixture of ground features in natural landscapes, an optimal resolution usually does not exist. The authors further examined the variance at each level of spatial resolution and for different forest entities (Marceau et al., 1994b). They found that each valley in the variance curves in relation to spatial resolution could be interpreted in terms of particular aggregation level of the forest entities, such as individual tree crowns, and the proportion of crowns, soil,
and shadow characteristics of a forest stand. Thus, they suggested that it might be possible to establish relationships between particular mixtures of ground features and corresponding spatial resolutions if the results observed in the study could be extrapolated.

2.3. Spatial scaling of remotely sensed data and classified land cover maps

Studies of scaling issues, either of land cover maps or datasets used to derive land cover maps, are particularly important as efforts are made to develop improved global data for land applications and ecoclimatological modeling (Sklar and Costanza, 1990; Townshend, 1992; Goodchild et al., 1993; Loveland et al., 1995). A large number of studies in spatial scaling are focused on data acquired by the Advanced Very High Resolution Radiometer (AVHRR). AVHRR data have been used widely for regional and continental scale vegetation monitoring (e.g., Justice et al., 1991) due to the wide spatial and frequent temporal coverage of the image. However, the 1.1km-resolution AVHRR pixels are rarely homogeneous. Therefore, much research has been focused on comparison between data from AVHRR and sensors aboard the Landsat and SPOT satellites (e.g., Gervin et al., 1985). Those comparisons usually involved the upscaling of fine resolution data such as TM (30m resolution) to a resolution equivalent to that of AVHRR (e.g., Townshend and Justice, 1988; Belward and Lambin, 1990; Aman et al., 1992; Moody and Woodcock, 1994). Spatial upscaling into resolutions coarser than 1km were also common subjects of studies due to the requirements of continental and global level ecosystem models (e.g., Townshend and Justice, 1988; Justice et al, 1989;
Malingreau and Belward, 1992; Belward, 1992; Oleson et al., 1996; DeFries et al., 1997). Most investigations of AVHRR data involved the use of NDVI (e.g., Gutman, 1987, 1991; Goward et al., 1991; De Cola, 1997), because land cover classification using AVHRR is usually based on analysis of the temporal patterns of the NDVI (e.g., Tucker et al., 1985; Townshend et al., 1987; Loveland et al., 1991, 1995).

Relationships between image information content and spatial resolution at resolutions from 125m to 4km were examined by Townshend and Justice (1988). The red and near-infrared bands of Landsat MSS images over seven geographical areas, including agricultural land, rangeland, woodland, and tropical rain forests, were used in their study. Original data of 79m resolution were first degraded to 125m and then successively to a resolution of 4km at a reduction rate of 2, resulting in datasets with seven different spatial resolutions. NDVI was derived from these datasets and several measures of information content were calculated from the NDVI data (e.g., standard deviation, entropy, power, and mean value). It was found that temporal and spatial characteristics of the terrain have a complex interrelationship which must be linked to the spatial characteristics of the sensing system. Contrary to the conclusions drawn from studies using aircraft simulated high resolution data, the authors concluded that land transformations are better represented by higher spatial resolution data. However, improvements in representation do not occur as a simple function of increasing resolution but are complex functions of spatio-temporal characteristics of the terrain and the type of land cover under consideration. Because each scene was composed of a variety of land cover types with widely differing spatial characteristics, it was difficult to sort out the
optimum spatial resolution for a specific land cover type. The study was extended to a resolution of 64km and it was again noticed that the overall variance of the scenes was contributed by a variety of landscapes with differing intrinsic spatial characteristics, which made the selection of appropriate resolution difficult (Townshend and Justice, 1990). These observations were similar to those obtained by Marceau et al. (1994a, b) in a predominantly forested region using fine resolution airborne data.

A study on the assessment of different resampling procedures was conducted by Justice et al. (1989). To reduce the data volume to a manageable extent, an on-board sampling procedure is adopted to produce AVHRR 4 km Global Area Coverage (GAC) data from 1 km Local Area Coverage (LAC) data. The sampling is an average/skip process. For a given scan line, the first four pixels are averaged and the next pixel is skipped. This sequence is continued along the line. The next two lines of data are skipped entirely and the "averaging" procedure is resumed for the following line. The authors compared this sampling procedure with other methods: 1) averaging all LAC pixels covered by one GAC pixel; 2) averaging four dispersed pixels; 3) selecting the single value from the center pixel; and 4) using the median value of all pixels in a window. Correlation analysis was performed between the original LAC data and each of the GAC datasets created by different procedures. The results showed that different resampling procedures provided different representations of the original data. The average/skip procedure adopted by NOAA is among the poorest while the average of all LAC pixels in a window is the best. Furthermore, the relationship between LAC and GAC depends on the spatial-spectral heterogeneity of the terrain in the scene. A
spectrally heterogeneous terrain is likely to result in larger differences between LAC and GAC. The study also showed differences in NDVI derived before and after the resampling of red and near-infrared bands. Calculating NDVI directly from GAC data will provide lower values than would be obtained by calculating them prior to averaging.

Belward and Lambin (1990) degraded a full Landsat MSS scene and an equivalent AVHRR HRPT (High Resolution Picture Transmission) near-nadir sub-image of West Africa successively to the scale of AVHRR GAC resolution and above (8 km). Curves of the mean local variances of both digital number and NDVI against cell size were plotted. Local variances of the two images showed only small changes. The authors attributed the results to the fact that the landscape in the study area was heterogeneous at a wide variety of scales. Thus, there were no distinctive scales that could result in sharp variations in local variance. However, results showed large differences in the local variances between MSS and AVHRR data. Because local variance measurement is image specific, the authors believed that the absolute values of local variance between MSS and AVHRR images were not comparable and slopes of the curves, instead, should be examined. They concluded from the opposite slopes between MSS and AVHRR that it was not possible to simulate AVHRR HRPT images from Landsat MSS images by simple aggregation of pixels, and therefore, sensor-specific factors must be considered. To examine the effect of spatial autocorrelation in AVHRR data, the MSS simulated 1 km data was smoothed by a 3x3-pixel filter and then collapsed to 8 km resolution. Local variance was again plotted against resolution. The curve of this smoothed MSS data more closely matched that of AVHRR data. The authors believed that the results demonstrated...
the autocorrelation between neighboring pixels of the AVHRR data. Furthermore, the authors concluded that the high spatial autocorrelation in AVHRR data were of little consequence when the sizes of ground objects were far greater than AVHRR pixel size. But, autocorrelation must be taken into account when AVHRR data are used in small study area. The spatial limitations of the AVHRR data make the accurate assignment of thematic content to specific groups of pixels difficult.

Because of the extensive use of NDVI in various land cover applications, a large portion of scale-related research has been directed towards examination of the relationships between NDVI and spatial resolution (see, for example, Townshend and Justice, 1988; Justice et al., 1989). Jasinski (1990) examined the NDVI in terms of the variability in subpixel landscape components, and with respect to variations in pixel sizes, within the context of the stochastic-geometric canopy reflectance model. Results indicated that, for Poisson distributed plants and for one deterministic distribution, NDVI increased with increasing subpixel fractional canopy amount, decreasing soil background reflectance, and increasing shadows. The simulation and analytical model also demonstrated that the variance caused by both soil reflectance variability and shadow variability was reduced as the pixel size increased. The variance due to shadow variability approached zero for spatially homogeneous plant distributions at large sampling scale ratios.

An investigation of the spatial characteristics of NDVI up to the scales commonly used in global modelling was conducted by Justice et al. (1991) for the whole continent of Africa. NDVI generated from AVHRR GAC data was spatially degraded from 8 km
resolution to 512 km, and scale variance analysis was performed for the resultant datasets. Although there was a general increase in variance at the coarser spatial resolution, considerable variation existed among the scale variances obtained from different areas. It appeared that the spatial variability of NDVI at the coarsest scale resulted from response of the zonal vegetation to varying climate conditions, while in the finest resolution (8 km), variability in NDVI was more associated with local difference in vegetation characteristics controlled by differences in topography, lithology and soil moisture. Interpretation of changes in the scale variance across resolutions was confounded in some cases due to the complexity of interaction among pixel sizes and land cover components.

As indicated in the study of Justice et al. (1989), there are two mathematically unequal procedures used in deriving coarse resolution NDVI from fine resolution data: 1) degrading both visible and infrared observations of the finer resolution system to the scale of the coarse resolution sensor and then calculating the vegetation index; and 2) obtaining vegetation indices from finer resolution data first and then degrading the results. Hall et al. (1992) demonstrated theoretically that, for an area composed of n homogeneous patches, NDVI is linear in radiance only in cases where either the \( \text{corr}[(\text{nir-vis}), (\text{nir+vis})] = 0 \) or \( (1/n) \sum \Delta(\text{nir-vis})^2 \sum \Delta(\text{nir+vis})^2 \frac{1}{n} = 0 \). Aman et al. (1992) analyzed the correspondence between NDVIs obtained from these two procedures by simulating AVHRR data from high spatial resolution SPOT and TM data over two geographical locations. Results of their investigation showed that the two NDVIs obtained, although algebraically different, were linearly correlated for both sites. The
authors concluded that NDVI, even though it is not a linear combination of radiances or reflectances, can be spatially integrated without significant loss of information. De Cola (1997) upcaled TM NDVI, using the two methods, from 30m to 16km in an area containing a large metropolis. It was found that the two NDVIs appeared to be generally commutative ($r=0.995$) although there were some differences. Similar results were found in predominantly grassland areas by Friedl et al. (1995). Further studies, however, were recommended to draw more generalized conclusions (Aman et al., 1992; De Cola, 1997).

Belward (1992) studied the behavior of local-variance/spatial-resolution for both HRPT and GAC data acquired from AVHRR for different ecological zones in West Africa. He found that the results were dependent on the ecological zones and on the part of the electromagnetic spectrum through which the data were recorded. Transitions across major ecological zones were detected across a range of resolutions. The study showed that in West Africa HRPT data could provide spatial information of land surface at full resolution, but the information was unreliable from GAC data. At resolutions coarser than $12 \text{ km}$, degraded GAC data were as good a source of land surface information as degraded HRPT data. These observations highlight the importance of a consideration of scale when using AVHRR data for vegetation monitoring, and emphasize the need for observations of different spatial and temporal scales.

Direct comparison among datasets obtained from different spatial levels is usually difficult because other factors such as viewing/illumination angles and atmospheric conditions will also influence the measurements. A controlled experiment was conducted by Demetriades-Shah et al. (1992) to compare datasets obtained from different platforms.
Concurrent measurements of ground-, helicopter-, and satellite-based reflectance were conducted at 13 tallgrass prairie sites. Ground measurements were taken looking vertically down, usually within 2 hours of noon. Landsat passed the site at around 1135 CDT obtaining data at an angle of about 5° and SPOT passed over at around 1235 CDT obtaining data at an angles of about 18° and 24°. The results showed the helicopter and satellite measurements were strongly correlated with hand-held radiometer NDVI values. The effects of atmosphere, sampling error, and differences of solar and viewing angles could be mostly overcome. Their findings suggested that the satellite measurements can be directly related to spectral measurements taken on or near the ground.

A number of recent studies focused on the estimation and calibration of errors introduced by spatial upscaling of land cover maps (e.g., Moody and Woodcock, 1994; Mayaux and Lambin, 1995; Moody, 1996) and on the effects of upscaling land cover data on global modeling (Oleson et al., 1996; DeFries et al., 1997). Scale-dependent errors in the estimation of land cover proportions were investigated by Moody and Woodcock (1994) by upscaling a TM-derived land cover map to a number of resolutions up to 1.02km. Results indicated that land cover classes with different patch sizes and connectivity between patches behave differently during the process of aggregation. Classes consisting of large, homogeneous patches grew larger as the sampling resolution was degraded, while classes characterized by highly clumped distributions, but small or intermediate sized patches, first grew and then decreased in size as the sampling resolution was progressively degraded beyond the typical patch size for that class. Classes composed of small fragmented units rapidly disappear as they were dominated
by more clumped cover types through the aggregation procedure.

As most pixels at 1.02km resolution are mixed-pixels, the authors proposed two possible ways of estimating the actual land cover proportions from coarse-resolution classification maps. First, they suggested using a high resolution maps to calibrate the mixtures of land cover components for various ecoregions. Second, they proposed developing regression relationships that predict land cover proportions at a coarse resolution as a function of land cover proportions at a fine resolution, the spatial properties of the fine resolution data, and the size of coarse resolution pixels. Both methods, however, involve significant effort, and consistencies across locations, classes, and times must be demonstrated before the methods can be actually adopted.

A regression-based calibration procedure was developed by Moody and Woodcock (1995a; Moody, 1996). Landscape indices derived from the original full resolution classification maps and the original class areas were used as predictor variables in standard forward regressions and tree-based regressions to estimate the proportional errors caused by spatial upscaling. The landscape indices selected included patch size, patch standard deviation, variance/mean ratio and patch diversity. Significant relationships between spatial characteristics, as reflected by landscape structure indices, and scale-dependent proportional errors were found. The authors suggested such relationships could lead to some generalizable understanding of scaling processes for a variety of landscape types.

Regression models that require the use of variables (e.g., landscape structure indices) measured from original full-resolution land cover maps, such as those developed
by Moody and Woodcock (1995a,b), are not generally applicable because the full resolution land cover maps needed to establish the regression models usually do not exist. Mayaux and Lambin (1995) proposed a two-step regression procedure. Using land cover maps developed independently from AVHRR and TM images, the authors regressed first the land cover areas obtained from the TM images against those obtained from the AVHRR images. In the second step regressions, the intercept and slope in the first regression were used as dependent variables and landscape indices were used as predictor variables. The landscape indices derived both from the TM images and from the AVHRR images were tested. The authors found that although values of the coefficient of determination were higher when TM-derived landscape indices were used than when AVHRR-derived indices were used, the improvement gained by using TM-derived indices was not large. The results suggested that landscape indices derived from the coarse resolution land cover maps might be used to replace those derived from fine resolution maps without significant increase in estimation errors. Using this method, the regression models could be made operational because all information could be obtained from coarse-resolution maps.

In an analysis of the sensitivity of land surface parameterization schemes to land cover datasets derived from remotely sensed imagery, Oleson et al. (1996) upscaled AVHRR-derived land cover having 1km original resolution and TM-derived land cover having 100m resolution to 0.2°, 0.4°, and 2.8° latitude/longitude grid cells. The upscaled land cover information was then used to parameterize the inputs of a global circulation model (GCM). It was found that there were substantial differences in areal proportions
of land cover types between the upscaled land cover datasets and a standard map-based
dataset commonly used in GCMs. The differences in land cover types resulted in
significant differences in model output in one 2.8° GCM cell, but differences in another
2.8° cell were smaller due to the similar biophysical parameterization of land cover types.
The authors believed that satellite-derived land cover datasets were needed to replace
outdated and inaccurate land cover information within current GCMs and that the choice
of land cover for model grid cells should be evaluated with respect to the land cover
composition and the response of each cover type to the range of atmosphere forcing.

In a similar study, DeFries et al. (1997) examined how the spatial upscaling of
land cover datasets would impact the output of global atmosphere-biosphere models. The
original land cover data used in the study had 8km resolution and the two model grid cell
sizes used were 1° and 4°, respectively. Two schemes were adopted to parameterize
model input: one was using land cover information upscaled to model cell sizes and the
other was parameterizing input according to the areal proportions of land cover within
each model cell. The latter method was unrealistic in operational use because of the large
computational load, but it could be used to test the difference between the two
parameterization schemes. Their results indicated that, with parameterization methods
currently used in the Simple Biosphere Model, only modest improvements in parameter
estimation would be obtained by describing each 1° grid cell as a mixture of vegetation
types derived from the 8km land cover dataset rather than as a single cover type. The
results, of course, were site specific and model specific. Further studies for other models,
in other regions, and for additional spatial resolutions of land cover information were
suggested (DeFries et al., 1997).

2.4. Scaling studies in landscape ecology

Scaling issues have been one of the important topics in ecology because patterns and processes are always related to scale. Processes important at one scale may not be important at another and patterns observable at one scale may not be perceivable at another. Researchers investigating the same problem may find their results do not match if their studies are conducted at different scales. There is an especially large body of literature regarding scaling problems in landscape ecology (Forman and Godron, 1986; Meentemeyer and Box, 1987; Wiens, 1992), a field that investigates the intrinsic relationship between spatial pattern and ecological processes. There are two different but interrelated aspects to the scale issues in landscape ecology. The first is identifying the scale of landscape patterns (e.g., Cullinan and Thomas, 1992; Hunsaker et al., 1994) and the second is understanding the effects of scale on the identification of patterns (e.g., Nellis and Briggs, 1989; Turner et al., 1989b; Qi and Wu, 1996). The latter is more relevant to remote sensing because, in many studies, landscape metrics are derived from land cover maps classified from remotely sensed data (e.g., Nellis and Briggs, 1989; Simmons et al., 1992; Wickham and Ritters, 1994; Benson and MacKenzie, 1995).

Turner et al. (1989b) presented a framework to investigate the effect of changing spatial scale on landscape pattern. Several landscape indices were used to examine the change of patterns across spatial scales. The authors attempted to derive general
relationships between scale and the landscape indices using simulated landscapes with only two types of patches and an existing land cover map. They found that a diversity index decreased linearly as observation cell size (grain) increased, while dominance and contagion indices were related to the change in the number of patch types. Their results indicated that the spatial scale at which landscape patterns were quantified influenced the results so that measurements made at different scales may not be comparable. The authors believed that although relationships among simple landscape indices measured at different scales might be established, they varied across landscapes so that extrapolation of such relationships to different regions may not be possible. They also observed that information was lost at coarse cell sizes and the nature of such loss was related to the original proportional area and contagion of land cover types. It was possible to predict the loss of information if the contagion of cover types was known.

Benson and MacKenzie (1995) compared the percent water area and number of water bodies observed from SPOT, TM and AVHRR images across spatial resolutions and examined the stability of several landscape indices derived from water bodies as resolution changes. The authors found that most landscape metrics were sensitive to scale change and the patterns of increase or decrease in the metrics were not monotonic. They believed that interpolation between the spatial resolutions of different satellite sensors was possible with an approach involving aggregation of pixels. However, reverse estimation from coarse resolution to finer resolution might be otherwise problematic unless an empirical relationship has been established. Like Turner et al. (1989b), they observed a change in land cover area as resolution changes. Non-dominant cover types decreased
as resolution coarsened. But the rate of decrease was determined by the aggregated arrangement of the cover type (water) in the landscape.

While most studies showed that landscape indices were sensitive to spatial resolution (e.g., Turner et al., 1989b; Benson and Mackenize, 1995), Wickham and Ritters (1995) found that landscape metrics should not be dramatically affected by the change in pixel size up to 80m. Seven landscape metrics were investigated at four different resolutions ranging from 4m to 80m. Although all but two metrics were significantly related to pixel size, the slope estimates were all small, indicating that the ranges of change in the metrics were limited. The reason that there were no significant influences of resolution on landscape indices in their study was the fact that the land cover type proportions remained essentially unchanged across the range of pixel sizes examined. Actually, the results of Wichham and Ritters’s study do not indicate that the landscape metrics they examined are not sensitive to pixel size. Rather, they indicate that landscape metrics are related to the spatial pattern as characterized by the land cover classification but not to the resolution itself. The effects of resolution are first reflected in the land cover map. If resolution coarsening does not exceed the patch sizes of the land cover types in an area, there will be no significant effects on land cover representation and, in turn, there will be little or no effect on landscape metrics derived from the land cover map.
2.5. Summary and comments

Considerable research has been directed towards spatial scaling issues in remote sensing and a variety of sensors, scale ranges and different regions have been covered. However, there is limited research in some areas. First, many studies have been conducted either on very broad and heterogeneous regions, such as whole continents (e.g., Justice et al., 1991), or on small areas with a limited number of cover types (e.g., Woodcock and Strahler, 1987). Scaling has been shown to be related not only to image resolution but also to land surface characteristics. In different regions, a different set of interactions may dominate the observed spatial variations at different scales. For example, Woodcock and Strahler (1987) found that, by using aerial photos and TM images, the local variance of an image decreased as its resolution exceeded the size of the objects in the scene, while Justice et al. (1991) observed a general increase in local variance from 8 km to 512 km resolution, although dimensions of most ground objects seemed to be less than the resolutions. The reason for this and other discrepancies is probably due to the great differences between study areas. Even in the same area, observed relationships vary from one cover type to another (e.g., Markham and Townshend, 1981). Using a study area with many complicated surface characteristics, such as a whole continent (e.g., Justice, et al., 1991), tends to make interpretation of results difficult. This suggests the need for "hierarchical" studies from relatively small, homogeneous areas to large, heterogeneous regions.

Second, most previous studies were conducted using only one type of remotely
sensed data. This is probably due to the few cases where investigators have simultaneous observations of different sensors over the same area. Because of dissimilarity of sensor characteristics, such as the discrepancies in band width and center wavelength, inconsistencies can be found in the results obtained from data of different sensors. For example, Aman et al. (1992) concluded that coarse resolution NDVI (AVHRR) could be simulated from fine resolution data (TM and SPOT). The authors examined only the simulated data but did not compare the results with actual AVHRR data. Belward and Lambin (1990) observed from simultaneous MSS and AVHRR images that it was not possible to simulate AVHRR data with MSS images by simple aggregation methods. It would be beneficial to conduct more studies on simulated and observed data from different sensors for better interpretation of contradictory results.

Third, most studies have been conducted for a single date and single season. There are few studies comparing time series of satellite data obtained with different pixel sizes over a given area (Malingreau and Belward, 1992). While multitemporal composite NDVI data have been shown to have a great advantage in characterizing "seasonally-distinct" land cover types (Loveland et al., 1991, 1995), it would be desirable to conduct investigations on the temporal patterns of NDVI/scale relationships.

Fourth, investigations of image information content of satellite remotely sensed data have often been conducted without looking at landscape heterogeneity. Early studies (using simulated satellite data), focused on the classification/resolution relationships, showed that the accuracy of land cover classification was closely related to pixel size of the image being used (e.g., Markham and Townshend, 1981). Using high resolution
airborne data, Woodcock and Strahler (1987) found a relationship between peak local variance and the size of scene objects. Because ground truth was difficult to obtain for coarse resolution images, results of many studies on information content were interpreted based on poorly-defined land cover classes.

Finally, there are few investigations into the effects of alternative scaling approaches on land cover representation. Land cover maps are usually upscaled using spatial aggregation methods (e.g., Turner et al., 1989b; Moody and Woodcock, 1994). Although it has been generally recognized that different scaling approaches will result in different representations of original data (Justice et al., 1989), additional research is warranted.
2.6. References


Mayaux, P., and Lambin, E. F., 1995, Estimation of tropical forest area from coarse spatial resolution data: a two-step correlation function for proportional errors due


Simonett, D. S., and Coiner, J. C., 1971, Susceptibility of environments to low resolution imaging for land-use mapping, Proceedings of the Seventh International


CHAPTER 3

DEVELOPMENT OF MULTISCALE DATABASES

3.1. Introduction

This chapter addresses the selection and description of study areas and preparation of databases. Study area selection was based on four criteria: 1) availability of remotely sensed and ancillary datasets; 2) accessibility for field investigation; 3) diversity in land surface characteristics, i.e., land cover types; and 4) diversity in landscape spatial structure. Major data sources used in the study included images acquired by Satellite Pour l’Observation de la Terra (SPOT), Landsat Thematic Mapper (TM), and Advanced Very High Resolution Radiometer (AVHRR). Ancillary datasets were utilized to assist land cover classification and validation, and to divide the large study area into homogeneous sub-regions. A number of datasets were derived from the original remotely sensed data, which included land cover classifications, at-satellite reflectance, and Normalized Difference Vegetation Index (NDVI). Both the original data and the derived products were upscaled to a variety of spatial resolutions to constitute multiscale databases.

3.2. Study areas

The conterminous United States was selected as a study area for AVHRR data
analysis. The study area included a variety of distinctive landscapes while the data volume was manageable. Two areas in Nebraska were selected for studies involving TM and SPOT data. The first Nebraska study area was located in the eastern part of the state (fig. 3.1). The area, about 10,000 km² in extent, includes Omaha and Lincoln, the two largest cities in Nebraska. The primary land use in this area, outside of the urbanized zone, was cropland. The major crop types grown include corn, sorghum, soybeans, wheat, oats, and alfalfa (Nebraska Agricultural Statistical Service, 1991). A second study area (fig. 3.1) was located in north central Nebraska along the Niobrara River. The area extended over approximately 4,000 km². The major land cover types in this area include native grasses, with small amounts of interspersed cropland. The two study areas represented very different landscape types.

3.3. Remotely sensed data

3.3.1. TM data

a) Characteristics of Thematic Mapper

The TM data used in this study were obtained from Landsat-5. The Landsat series of satellites were designed to meet the needs of resources managers and earth scientists by providing information about land surface characteristics and their temporal dynamics. Landsat-5 was launched in March 1, 1984 by the National Aeronautics and Space Administration (NASA). The satellite is in a sun-synchronous, near-polar orbit at an altitude of 705 km, having a 16-day revisit period.
The TM sensor aboard Landsat-5 is an optical-electronic system, which uses an oscillating mirror to scan perpendicular to the orbital path. The sensor simultaneously provides measurements over seven broad bands (Table 3.1). The six non-thermal bands have a spatial resolution of 30m. The levels of radiation energy from the earth's surface collected by the sensor are converted into digital signals. The digital data were quantified at an 8-bit level, resulting in 256 levels of digital numbers (DN) (i.e., 0 to 255).

Table 3.1. Resolutions of the TM sensor (from Jensen, 1986)

<table>
<thead>
<tr>
<th>Band</th>
<th>Spectral(µm)</th>
<th>Spatial(m)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.45-0.52</td>
<td>30x30</td>
</tr>
<tr>
<td>2</td>
<td>0.52-0.60</td>
<td>30x30</td>
</tr>
<tr>
<td>3</td>
<td>0.63-0.69</td>
<td>30x30</td>
</tr>
<tr>
<td>4</td>
<td>0.76-0.90</td>
<td>30x30</td>
</tr>
<tr>
<td>5</td>
<td>1.55-1.75</td>
<td>30x30</td>
</tr>
<tr>
<td>6</td>
<td>10.4-12.5</td>
<td>120x120</td>
</tr>
<tr>
<td>7</td>
<td>2.08-2.35</td>
<td>30x30</td>
</tr>
</tbody>
</table>

b) The TM image used in the study

A TM image covering the eastern Nebraska study area was acquired on August 26, 1991. The dataset was provided by the United States Geological Survey (USGS) EROS Data Center. The local time of the satellite overpass was about 1030. The image was geometrically corrected at the EROS Data Center and was geo-referenced to a UTM projection, with pixel size of 28.5x28.5 m². A nearest-neighbor resampling method was utilized in the geometric correction to minimize its impact on DN values. Only bands 1-5
and 7 were used in the study because the spatial resolution of band 6 was incompatible with those of the other bands.

The image was resampled, using a nearest-neighbor method, to 30x30 m² resolution prior to actual data processing. It comprises the "full resolution" TM dataset (30m resolution) used in this dissertation. The reason for this initial 28.5m to 30m resampling was that the 28.5m resolution could cause difficulties in comparing the results of this study area with those obtained from other study areas and by other investigators (most studies involving TM data upscaling have used 30m as the original resolution). The initial resampling might have effects on some coarse resolution data (e.g., 60m) upscaled from the 30m data but its effects became less obvious as resolution coarsens because the locational error became increasingly insignificant. A subimage of 3420x3420 pixels (102x102 km²) was subset from the full scene to cover the study area (fig. 3.2 and Table 3.2). There was no cloud contamination in the subimage.

| Table 3.2. Statistical Characteristics of the TM subimage |
|----------------|---|---|---|---|---|---|
| Band | Max | Min | Mean | Median | Mode | STD |
| 1    | 253 | 5   | 71.865 | 71  | 68  | 6.632 |
| 2    | 184 | 13  | 30.217 | 30  | 28  | 4.447 |
| 3    | 195 | 4   | 29.592 | 28  | 27  | 7.980 |
| 4    | 255 | 3   | 80.685 | 77  | 73  | 20.102 |
| 5    | 255 | 0   | 80.846 | 78  | 75  | 18.500 |
| 7    | 255 | 0   | 30.397 | 27  | 23  | 12.586 |
c) The TM NDVI image

The calculation of NDVI included three steps. First, the DN values of band 3 and 4 of the TM image were converted into radiances using appropriate calibration coefficients following Markham and Barker (1986):

\[ L_\lambda = L_{\text{min},\lambda} + (\frac{L_{\text{max},\lambda} - L_{\text{min},\lambda}}{QCAL_{\text{max},\lambda}}) \times QCAL_\lambda \]  

(3.1)

where

- \( L_\lambda \) is the radiance at band \( \lambda \) at sensor aperture in \( \text{mW cm}^{-2} \text{ sr}^{-1} \text{ \mu m}^{-1} \),
- \( QCAL_\lambda \) is the calibrated and quantified scaled radiance in units of DN,
- \( L_{\text{min},\lambda} \) is the spectral radiance at \( QCAL_\lambda = 0 \),
- \( L_{\text{max},\lambda} \) is the spectral radiance at \( QCAL_\lambda = QCAL_{\text{max},\lambda} \),
- \( QCAL_{\text{max},\lambda} \) is the range of rescaled radiance in DN.

Second, at-satellite (exoatmospheric) reflectances were computed from the radiances using Markham and Barker (1986):

\[ R_\lambda = \frac{\pi L_\lambda d^2}{E_{\text{sun},\lambda} \cos(\theta)} \]  

(3.2)

where:

- \( R_\lambda \) is the unitless effective at-satellite planetary reflectance of band \( \lambda \) (where \( \lambda \) is either near infrared or visible),
- \( L_\lambda \) is the spectral radiance at sensor aperture in \( \text{mW cm}^{-2} \text{ sr}^{-1} \text{ \mu m}^{-1} \),
- \( d \) is the earth-sun distance in astronomical units,
$E_{sun\lambda}$ is the mean solar exoatmospheric irradiance in $mW \ cm^{-2} \ \mu m^{-1}$ in band $\lambda$, and $\theta$ is the solar zenith angle in degrees.

Finally, the NDVI values were calculated from the red (band 3) and near-infrared (band 4) reflectances (fig. 3.3) using:

$$NDVI = \frac{R_{nir} - R_{red}}{R_{nir} + R_{red}}$$  \hspace{1cm} (3.3)

where $R_{nir}$ and $R_{red}$ are near infrared (TM4) and red (TM3) reflectances, respectively.

3.3.2. SPOT data

a) Characteristics of SPOT

SPOT is a European earth observing satellite system. The SPOT satellites operate in a sun-synchronous, near polar orbit at an altitude of 832 km. The satellites are inclined 98.7 degrees, cross the equator at 10:30 AM local time and have an orbital cycle of 26 days. They provide global coverage between 87 degrees north and south latitudes.

SPOT data are obtained from High Resolution Visible (HRV) sensors. Each satellite carries two HRVs and each HRV contains four CCD (Charge Coupled Device) subarrays. A 6000-element subarray is used in the panchromatic mode to record data at 10m resolution. Three 3000-element subarrays are used in the multispectral mode at 20m resolution. The multispectral mode captures data in two visible bands and one near-
infrared band. The panchromatic mode images data in the spectral range 0.51-0.73 micrometers (Table 3.3). The data are effectively encoded over an 8-bit 256 DN range.

Table 3.3. Spectral and Spatial Resolutions of the SPOT sensor

<table>
<thead>
<tr>
<th>Band</th>
<th>Spectral(μm)</th>
<th>Spatial(m)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.50-0.59</td>
<td>20x20</td>
</tr>
<tr>
<td>2</td>
<td>0.61-0.68</td>
<td>20x20</td>
</tr>
<tr>
<td>3</td>
<td>0.79-0.89</td>
<td>20x20</td>
</tr>
<tr>
<td>Panchromatic</td>
<td>0.51-0.73</td>
<td>10x10</td>
</tr>
</tbody>
</table>

*Summarized from Lillesand and Kiefer (1994).

**b) The SPOT images used in the study**

Two SPOT images over the Niobrara Valley Preserve and its vicinity were provided by the EROS Data Center. Multi-temporal images, acquired during the middle to latter part of the growing seasons in the study area, generally provided more information for vegetation classification than a single image. The images were acquired on June 7, 1994 and August 24, 1994 (fig. 3.4). The local times of the satellite overpasses were about 1140. The images were co-registered and geo-referenced to a UTM projection using nearest-neighbor resampling, with a pixel size of 20x20m². The sizes of both images were 3375x3500, covering an area of about 3,100km² (figs. 3.4 and 3.5, Table 3.4). There were a few clouds in both images. Cloud contaminated areas were not used in subsequent data analyses.
Table 3.4. Statistical Characteristics of the SPOT Images

<table>
<thead>
<tr>
<th>Band</th>
<th>Min</th>
<th>Max</th>
<th>Mean</th>
<th>Median</th>
<th>Mode</th>
<th>Std</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jun/1</td>
<td>36</td>
<td>219</td>
<td>65.537</td>
<td>66</td>
<td>67</td>
<td>7.082</td>
</tr>
<tr>
<td>Jun/2</td>
<td>20</td>
<td>194</td>
<td>49.593</td>
<td>51</td>
<td>51</td>
<td>8.763</td>
</tr>
<tr>
<td>Jun/3</td>
<td>18</td>
<td>209</td>
<td>95.985</td>
<td>95</td>
<td>93</td>
<td>12.923</td>
</tr>
<tr>
<td>Aug/1</td>
<td>30</td>
<td>188</td>
<td>54.547</td>
<td>55</td>
<td>55</td>
<td>6.745</td>
</tr>
<tr>
<td>Aug/2</td>
<td>13</td>
<td>181</td>
<td>42.875</td>
<td>44</td>
<td>46</td>
<td>9.016</td>
</tr>
<tr>
<td>Aug/3</td>
<td>14</td>
<td>178</td>
<td>76.329</td>
<td>74</td>
<td>71</td>
<td>12.088</td>
</tr>
<tr>
<td>Jun/NDVI</td>
<td>77</td>
<td>183</td>
<td>152.298</td>
<td>151</td>
<td>150</td>
<td>8.455</td>
</tr>
<tr>
<td>Aug/NDVI</td>
<td>75</td>
<td>181</td>
<td>149.169</td>
<td>146</td>
<td>143</td>
<td>10.317</td>
</tr>
</tbody>
</table>

c) The SPOT NDVI

The SPOT NDVI images (fig. 3.6) were derived using a procedure similar to that of the TM NDVI. However, the spectral radiances of the SPOT images were calculated using (Price, 1987):

\[ L(\lambda) = \alpha(\lambda) \times DN(\lambda) \]  

(3.4)

where

- \( L(\lambda) \) is the radiance in band \( \lambda \) in \( W \ m^{-2} \ sr^{-1} \ \mu m^{-1} \),
- \( \alpha(\lambda) \) is the calibration coefficient in band \( \lambda \) in \( W \ m^{-2} \ sr^{-1} \ \mu m^{-1} \), which was obtained from the header files of the image data, and
- \( DN(\lambda) \) is the DN of a pixel in band \( \lambda \).

At-satellite reflectance was calculated using equation 3.2, but the unit for spectral radiance, \( L(\lambda) \), and the mean solar exoatmospheric irradiance, \( E_{sun\lambda} \), is \( W \ m^{-2} \ sr^{-1} \ \mu m^{-1} \). \( E_{sun\lambda} \) was obtained from Price (1987).
NDVI was computed using equation (3.3). The SPOT near infrared and red bands utilized are bands 3 and 2, respectively.

3.3.3. AVHRR data

a) Characteristics of AVHRR

Since 1978, a series of Polar Orbiting Environmental Satellites (POES) has been launched by the National Oceanic and Atmospheric Administration (NOAA) to collect data over large areas of the Earth's surface. The NOAA satellites are in near-polar, sun-synchronous orbits with orbital periods of approximately 102 minutes and nominal altitudes of 870 km. Because the number of orbits per day (14.2) is not an integer, the exact sub-orbital tracks do not repeat on a daily basis. However, the local solar time of the satellite's passage is essentially unchanged for any latitude. For NOAA-11, which was launched in September, 1988, the approximate local solar time of the ascending node (northbound equator crossing) is about 1400 and that of the descending node (southbound equator crossing) is about 0200.

The Advanced Very High Resolution Radiometer (AVHRR) aboard the NOAA satellites is a broad band, four or five channel (depending on the model) scanner. The instantaneous field of view (IFOV) of the AVHRR results in a 1.1 km pixel size at nadir. The swath width of the sensor is approximately 2400 km. The NOAA-11 AVHRR captures Earth surface radiation in five bands, ranging from visible to thermal infrared wavelength. The radiometric resolution of the AVHRR data is 10-bit, i.e., 1024 DN
levels (Table 3.5).

Table 3.5. Resolutions of the AVHRR sensor

<table>
<thead>
<tr>
<th>Band</th>
<th>Spectral(μm)</th>
<th>Spatial(km)</th>
<th>Temporal(1/day)</th>
<th>Radiometric(bit)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.58-0.68</td>
<td>1.1x1.1</td>
<td>14.2</td>
<td>10</td>
</tr>
<tr>
<td>2</td>
<td>0.725-1.1</td>
<td>1.1x1.1</td>
<td>14.2</td>
<td>10</td>
</tr>
<tr>
<td>3</td>
<td>3.55-3.93</td>
<td>1.1x1.1</td>
<td>14.2</td>
<td>10</td>
</tr>
<tr>
<td>4</td>
<td>10.5-11.3</td>
<td>1.1x1.1</td>
<td>14.2</td>
<td>10</td>
</tr>
<tr>
<td>5</td>
<td>11.5-12.5</td>
<td>1.1x1.1</td>
<td>14.2</td>
<td>10</td>
</tr>
</tbody>
</table>

*summarized from NOAA (1988).

AVHRR data are provided in three formats. High Resolution Picture Transmission (HRPT) data are transmitted to ground station in real time and Local Area Coverage (LAC) data selectively recorded on board for subsequent playback. Both HRPT and LAC retain the original spatial resolution. Global Area Coverage (GAC) is resampled data with nadir spatial resolution of about 4 km. The data volume of GAC is sixteen times smaller than full resolution data.

b) Maximum value composite NDVI images

The measurements of the first two bands of AVHRR are used to calculate NDVI. NDVI derived from AVHRR data has been widely used in monitoring spatial patterns and temporal dynamics of terrestrial vegetation (e.g., Townshend and Justice, 1990). AVHRR-derived NDVI is, however, subject to contamination stemming from various
factors such as cloud cover, atmospheric aerosols and water vapor content, and viewing and illumination geometries (e.g., Justice et al., 1991). Therefore, AVHRR/NDVI data are often presented as images produced using a Maximum Value Composite (MVC) procedure. A MVC image is comprised of pixels having the maximum NDVI value observed over several (e.g., 10) successive days. The method usually tends to produce relatively cloud-free images and is believed to minimize effects of variable sun angle, water vapor, aerosols and directional surface reflectance (Holben, 1986).

c) The AVHRR images used in the study

The AVHRR data used in this study were extracted from the 1990, 1991 and 1994 Conterminous United States AVHRR digital datasets published on CD-ROMs by the EROS Data Center. Each CD-ROM set contains: 1) biweekly composited NDVI calculated from red and near infrared at-satellite reflectances; 2) calibrated AVHRR data (channels 1 through 5); 3) satellite viewing zenith angle; 4) solar zenith angle; 5) relative angle between sun and satellite azimuth; 6) scene identification numbers for each pixel used in the biweekly-compositing period; and other ancillary data. NDVI and channels 1 and 2 data of four composite periods were used in scaling analyses. These periods include August 17-30, 1990, August 16-29, 1991, May 27 - June 9 and August 19 - September 1, 1994 (Table 3.6 and fig. 3.7). Last three composite periods cover the acquisition date of the TM and SPOT data. The August 1990 period was selected because there is a corresponding 1990 land cover database and the season corresponds to the TM and one of the SPOT images. Other data corresponding to these composite
periods, including solar zenith angle and satellite viewing zenith angle, were used as reference data. AVHRR channels 1 and 2 data in the CD-ROMs were calibrated into at-satellite reflectance and linearly scaled to byte data using a procedure described by Eidenshink (1992). Thus, the reflectances were retrieved by reverse linear scaling and no conversion from DN value was needed.

Table 3.6. AVHRR NDVI Data Used in the Study

<table>
<thead>
<tr>
<th>Composite period</th>
<th>Data used for</th>
</tr>
</thead>
<tbody>
<tr>
<td>August 17-30, 1990</td>
<td>AVHRR NDVI upscaling</td>
</tr>
<tr>
<td>August 16-29, 1991</td>
<td>Comparison with TM NDVI (August 26, 1991)</td>
</tr>
<tr>
<td>May 27 - June 9, 1994</td>
<td>Comparison with SPOT NDVI (June 7, 1994)</td>
</tr>
</tbody>
</table>

3.4. Ancillary data

3.4.1. Ecoregion map

The map of Level I Ecological Regions of North America used in this study was obtained from the U.S. Environmental Protection Agency (EPA) through the Internet (ftp site: morpheus.cor.epa.gov). The map was last updated in January, 1995. The Level I ecoregion map depicts the coarsest level of the hierarchical framework that has been jointly prepared by the State of the Environment Directorate (Environment Canada), the National Institute of Ecology (Secretariat of Social Development, Mexico), and the EPA.
Environmental Research Laboratory in Corvallis, Oregon (cited from the readme file in the EPA ftp site). The ecoregion map was based on analysis of the patterns and the composition of biotic and abiotic phenomena that affect or reflect differences in ecosystem quality and integrity. These criteria include geology, physiography, vegetation, climate, soils, land use, wildlife, and hydrology (see the readme file at the EPA ftp site; Omernik, 1987; Omernik, in press).

The Level I ecoregion map was utilized because the regions identified in the Level I map were suitable for the analysis of upscaling the AVHRR-derived land cover (see below -- §3.5.3 for the AVHRR land cover). There are seven regions in the conterminous U.S. (nine in the whole of North America) (fig. 3.8 and Table 3.7). However, three of these regions (#3, #7, and #8) each contains less than ten percent of the area of the conterminous U.S. and were not used in the study.

Table 3.7. Level I Ecoregions of the Conterminous U.S.

<table>
<thead>
<tr>
<th>Region #3</th>
<th>Northern Glaciated Forests</th>
</tr>
</thead>
<tbody>
<tr>
<td>Region #4</td>
<td>Great Plains</td>
</tr>
<tr>
<td>Region #5</td>
<td>Western and Southern Deserts Basins and Ranges</td>
</tr>
<tr>
<td>Region #6</td>
<td>Northwestern Mostly Coniferous Forested Mountains</td>
</tr>
<tr>
<td>Region #7</td>
<td>Southern Cordillera</td>
</tr>
<tr>
<td>Region #8</td>
<td>Tropical Lowlands</td>
</tr>
<tr>
<td>Region #9</td>
<td>Eastern Temperate Forests</td>
</tr>
</tbody>
</table>

3.4.2. Other ancillary data

Other ancillary datasets utilized in this study include: crop acreage reports (USDA
Consolidated Farm Service Agency, unpublished reports), historical crop planting maps (Department of Agronomy, University of Nebraska, unpublished maps), crop statistics (Nebraska Agricultural Statistical Service, 1991), STATSGO soils (U.S. Department of Agriculture, 1994), Digital Line Graphs (U.S. Geological Survey, 1990), and aerial photographs. These datasets were used in assisting the land cover classification in the two Nebraska study areas.

3.5. Land cover classification

3.5.1. TM land cover classification and labeling

The six band TM data were used to develop a land cover map. A self-iterative, unsupervised clustering algorithm was used to create class seeds. A total of 120 initial clusters were generated. The spectral properties of these initial clusters and the distances between clusters were examined. Those clusters having spectral curves that were very similar and having distances very close to other clusters were merged to their closest clusters. One hundred clusters were retained as the seeds for second round classification. The selection of a cluster number of 100 was somewhat arbitrary, although it was based on two considerations: keeping as much spectral information as possible and eliminating unnecessary clusters.

In the next step of classification a supervised maximum likelihood classifier was used. The output of the classification was 100 new spectral clusters. Because the cluster seeds used in this maximum likelihood classification were not related to the land cover
classes in the study area, the classification was actually still unsupervised in terms of ground truth. Each of the new spectral classes was assessed by overlaying it on a pseudocolor composite image (Bands 4, 3, and 2) and assigning it to a land cover class. The land cover classes in urban areas were quite different from those in rural areas although some had similar spectral characteristics (e.g., alfalfa and golf course turf). To resolve this confusion, an urban mask layer was digitized. In the urban areas, labeling of land cover classes was based on visual interpretation, field investigation and aerial photos, while in rural areas, labeling was primarily based on historical crop maps, acreage reports and county crop statistics. A total of 19 land cover classes were identified in the study area. These included 10 rural classes, 8 urban classes and 1 shared class, open water (fig. 3.9).

3.5.2. SPOT land cover classification

A land cover classification derived from the two SPOT images was provided by Dr. Bruce Wylie of the EROS Data Center (Wylie, 1995, personal communication). The classification used the two three-band SPOT images as well as two NDVI images derived from the SPOT images. The images were combined into a single eight-band image. Unsupervised clustering was performed on the eight-band image. A total of 60 spectral clusters were generated from the classification. Each cluster was overlaid on the pseudo color composite (band 3, 2, and 1) of the original images for labeling. The clusters were examined according to black and white and near-infrared aerial photos, soil properties
(from STATSGO soils data), and field investigations. It was found that some of the original 60 clusters consisted of more than one land cover class. Class splitting was performed on those mixed clusters and a new 170-cluster map was produced. After labeling, a final 26-class land cover map was obtained. I further grouped the 26 classes into 17 more general classes (fig. 3.10).

3.5.3. AVHRR land cover classification

The conterminous United States 1 km AVHRR land cover database used in this study was a product of joint research conducted by the EROS Data Center and the Center of Advanced Land Management Information Technologies (CALMIT) of the University of Nebraska-Lincoln. To produce this database, the investigators first clustered maximum value composite (MVC) AVHRR/NDVI data acquired over the 1990 growing season into 70 seasonally-distinct land cover classes. Subsequently, the 70 classes were labeled, described, split and refined using a variety of ancillary datasets. The final product portrays 159 seasonally-distinctive land cover classes (Loveland et al., 1991, 1995). The concepts, methods and procedures used in producing this land cover map was described by Brown et al. (1993).

The original 159-class land cover database differs from conventional land cover maps in many respects as it incorporates seasonality and productivity as well as land cover (Loveland et al., 1995). Tables linking this database and other commonly used land cover classification systems were created to facilitate translations between different
systems. In this study, I grouped the 159 classes into 25 general classes primarily based on the "types" of land cover regardless of their seasonality and productivity (fig. 3.11 and Table 3.8). Although the grouping reduced overall information content, the 25-class database still distinguishes major ecological regions. This land cover database was used in examining interactive effects of both spatial aggregation and class grouping on the representation of land cover.

Table 3.8. Twenty-five Grouped Land Cover Classes

<table>
<thead>
<tr>
<th>Class</th>
<th>Class Label</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Dryland cropland</td>
</tr>
<tr>
<td>2</td>
<td>Irrigated cropland</td>
</tr>
<tr>
<td>3</td>
<td>Mixed dry/irrigated cropland</td>
</tr>
<tr>
<td>4</td>
<td>Grassland/cropland</td>
</tr>
<tr>
<td>5</td>
<td>Woodland/cropland</td>
</tr>
<tr>
<td>6</td>
<td>Grassland</td>
</tr>
<tr>
<td>7</td>
<td>Desert shrubland</td>
</tr>
<tr>
<td>8</td>
<td>Mixed shrubland/grassland</td>
</tr>
<tr>
<td>9</td>
<td>Chaparral</td>
</tr>
<tr>
<td>10</td>
<td>Savanna</td>
</tr>
<tr>
<td>11</td>
<td>Northern deciduous forest</td>
</tr>
<tr>
<td>12</td>
<td>Southeastern deciduous forest</td>
</tr>
<tr>
<td>13</td>
<td>Western deciduous forest</td>
</tr>
<tr>
<td>14</td>
<td>Southeastern coniferous forest</td>
</tr>
<tr>
<td>15</td>
<td>Western coniferous forest</td>
</tr>
<tr>
<td>16</td>
<td>Western woodland</td>
</tr>
<tr>
<td>17</td>
<td>Northern mixed forests</td>
</tr>
<tr>
<td>18</td>
<td>Southeastern mixed forests</td>
</tr>
<tr>
<td>19</td>
<td>Western mixed forests</td>
</tr>
<tr>
<td>20</td>
<td>Herbaceous wetlands</td>
</tr>
<tr>
<td>21</td>
<td>Forested wetlands</td>
</tr>
<tr>
<td>22</td>
<td>Barren</td>
</tr>
<tr>
<td>23</td>
<td>Subalpine forest</td>
</tr>
<tr>
<td>24</td>
<td>Alpine tundra</td>
</tr>
<tr>
<td>25</td>
<td>Water</td>
</tr>
</tbody>
</table>
3.5.4 Accuracy assessment of the land cover classifications

The TM land cover classification was compared with crop acreage reports obtained from the USDA Consolidated Farm Service Agency (CFSA) offices and with historical crop maps of agricultural research sites of the Department of Agronomy, University of Nebraska-Lincoln (UNL) (unpublished reports and maps of CFSA and UNL, 1991). One hundred and ten farm fields were digitized from the CFSA reports and UNL crop maps and then overlaid onto the classification map to assess the accuracy of the five most predominant classes (sorghum, soybean, corn, wheat/oat, and alfalfa). The overall accuracy of the five crop classes was 74%. The accuracy of the remaining classes was estimated through aerial photo interpretation and field investigation. About 128 pixels located in urban and suburban areas of Lincoln and Omaha were selected and subsequently used for verification. Of these sites, 80 were interpretable in aerial photos and another 28, located in easily accessible sites, were field-checked. Using such methods, the overall accuracy for barren, forest and water classes in rural areas and all classes in urban areas was determined to be over 85%, while accuracy for fallow, grass, and wet meadow ranged from 60% to 80%. The overall accuracy for the study area was estimated to be greater than 75%. It should be pointed out that the estimated accuracy may be somewhat biased because the selection of sample farm fields and pixels was based on availability of ground truth and was, therefore, not randomly distributed throughout the study area. Nevertheless, it is believed that the accuracy of the classification is sufficient for the purpose of this study of spatial scaling on land cover
representation.

The accuracy of the SPOT classification was assessed through field investigation, which was conducted in and near the Niobrara Preserve by Bruce Wylie (EROS Data Center) and me during the summer of 1995. It was estimated that the accuracy for forest, corn and other-crops is over 90% and the overall accuracy for the other classes is around 70% (Wylie, 1995, personal communication).

A solid accuracy assessment for the AVHRR land cover database is difficult as field validation may be prohibitively expensive while existing land cover/use maps may not be acceptable standards of reference. Nevertheless, preliminary results indicate that overall the classification performed well (Merchant et al., 1993). The database has been used as input for a number of ecological modeling studies (Steyaert et al. 1993).

3.6. Creation of multiscale databases

3.6.1. The "averaging" upscaling procedure

The creation of simulated multi-resolution images was based on an averaging algorithm conducted on the original full resolution image. A computer program was written to perform the averaging process. The program calculated the average value of a window, the size of which was defined by the ratio of user-prompted output resolution cell size to the original full resolution cell size. If the ratio was not an integer, the program used a closest integer value but would sometimes adjust the window size by one pixel to prevent error accumulation. For example, if the original cell size was 30m and
the output cell size is 1000m, the ratio of output to input was 33.33. In this case, the first
three output pixels will be computed from pixel 1 to 33, 34 to 66, and 67 to 100 of the
input image. That is, when roundup error exceeds half of the input cell size, an extra
input pixel will be included. This procedure ensures that locational errors are less than
one input cell size.

The coarse resolution images generated using the average procedure are not
directly comparable to those obtained from sensors with coarse resolutions because they
do not account for differing sensor characteristics, viewing/illumination geometries, and
atmospheric conditions. They are simulations ("coarse resolution TM, or SPOT images")
having parameters similar to those of actual TM or SPOT except for the resolution cell
size. The images were developed to allow investigation of scaling effects without the
influence of other factors such as varying viewing geometry and spectral resolution found
in different sensors.

However, the averaging procedure implicitly assumes that the response function
of the TM sensor is a square-wave function since all radiances averaged equally within
a coarse pixel and everything outside is excluded. A more accurate representation of an
image obtained with a coarse resolution sensor would involve modeling the sensor's
modulation transfer function between different resolutions to derive spatial filters for
simulating coarse resolution data (Justice et al., 1989; Markham and Townshend, 1981;
Sadowski and Sarno, 1976). In fact, the square-wave response function approximation
has been used in many other studies (e.g., Bian and Walsh, 1993; Collins et al., 1995;
Woodcock and Strahler, 1987). The assumption that the radiance in a coarse resolution
pixel is contributed evenly by proportional subpixel components is also the basis of linear unmixing models, which are used to estimate the proportional areas of component cover types in a mixed-pixel according to their spectral properties (e.g., Puyou-Lascassies et al., 1994; Quarmby et al., 1992; Shimabukuro and Smith, 1991). Therefore, it was judged that the averaging procedure was acceptable in this study.

3.6.2. Upscaling of the TM data

A total of eleven coarse resolution TM images were created using the aforementioned averaging procedure. These had resolutions of 60, 90, 120, 240, 360, 480, 600, 720, 840, 960 and 1,000 meters, respectively. Included in this series are images having resolutions approximately equivalent to currently operated sensor systems (e.g., AVHRR) and proposed sensors such as the Moderate Resolution Imaging Spectroradiometer (MODIS, 250m and 500m resolutions). The upper limit of 1km was determined because it is the resolution of the AVHRR data used in this study. The 1km resolution TM data were only used when comparisons to AVHRR were needed. When only TM data were involved (e.g., comparison between TM NDVIs upscaled using different methods), the coarsest resolution involved was 960m, close to 1km. The 1000m resolution was not used for TM upscaling (when no AVHRR data were involved) because there were locational errors (up to 10m, i.e., 0.33 times of original TM pixel size) in changing 30m resolution to 1000m resolution.

Three datasets were produced at each coarse resolution level. The first was the
coarse resolution DN data, which was directly upscaled from the full resolution (30m) TM image using the averaging procedure. The second dataset consisted of coarse resolution NDVI images calculated from the TM images at corresponding resolutions. And the third dataset was another series of NDVI images which were upscaled, using the same averaging procedure, from the full resolution NDVI (i.e., NDVI computed from the full resolution TM image). The difference between the second and the third datasets was that in the second series, NDVI was calculated after degradation of DN values while in the third series NDVI was computed from full resolution DN values and then upscaled.

3.6.3. Upscaling of SPOT data

Two series of coarse resolution SPOT NDVI data were derived. For the first series, the average upscaling procedure was conducted separately on SPOT HRV bands 2 and 3 of the August image. The coarse resolution data were computed at 40, 60, 80, 120, 240, 360, 480, 600, 720, 840, 960, and 1,000 meters. The two upscaled bands were used to calculate coarse resolution NDVI. The second series was upscaled from full resolution (20m) NDVI, which was derived from bands 2 and 3 of the original 20m resolution data.

3.6.4. Upscaling of AVHRR data

The AVHRR MVC NDVI image for the period of August 17-30, 1990 (fig. 3.7)
was used in the scaling analysis of AVHRR to regional scales. This specific composite period was selected because 1) there was an existing land cover database for 1990; 2) it was the peak of the growing season for most vegetation species in the conterminous U.S.; and 3) the seasonal composite period was consistent with the timing of the TM data, although there was one year difference. Two series of coarse resolution AVHRR NDVI images were produced by 1) averaging the original 1km MVC NDVI, and 2) averaging the red and near infrared reflectances first and then calculating coarse resolution NDVI. The datasets included two image series at 2, 4, 8, 16, 24, 32, 40, 48, 56 and 64 kilometer resolutions. The upper limit of 64 km was used because further coarsening would make the resultant image too small and because the 64km resolution was approximately equal to the 0.5x0.5 degree size used by some ecoclimatological models.

3.6.5. Upscaling of land cover data

a) Coarse resolution maps by aggregation

There are a number of rules commonly used to assign a representative class when a fine resolution land cover map is to be aggregated into a coarse map. For example, the random rule assigns a land cover class label to a larger (coarser) pixel by randomly selecting the label from a fine resolution pixel within the area of the coarse pixel. The priority rule first checks if a certain priority class exists within the area of the coarse pixel. The priority class will always be selected to represent the coarse pixel, regardless of the proportional area of the class with the coarse pixel.
The majority rule, which selects a majority class in the coarse resolution pixel, is the most commonly used method (e.g., Costanza and Maxwell, 1994; Oleson et al., 1996; Turner et al., 1989). It was used in this study in coarsening the full resolution land cover maps derived from SPOT, TM, and AVHRR data. A computer program, capable of window size adjustment when the coarse-to-fine ratio is not an integer, was developed to perform the aggregation procedure. Non-overlapping square windows, each representing one pixel in an output coarse-resolution image, were overlaid on the full resolution land cover map. Each pixel in the output image was assigned to the class that comprised the majority in a window. When two or more classes occurred with the same frequency, the window was randomly assigned to one of the candidate classes. This situation occurred most frequently when the difference between coarse and fine resolutions was small (e.g., 30m to 90m).

The resolution series included 60, 120, 240, 360, 480, 600, 720, 840, and 960 meters for the TM and SPOT land cover data (see, for example, figs. 3.12 - 3.17). The TM series also had 90m while the SPOT series also included 40m and 80m. The AVHRR series had 2, 4, 8, 16, 24, 32, 40, 48, 56, and 64 kilometers (see, for example, figs. 3.18, 3.19, 3.20). Both the original 159-class and the grouped 25-class AVHRR land cover datasets were spatially aggregated. The 159-class series is not illustrated here because the many classes were difficult to distinguish.

b) Coarse resolution maps by re-classification

The aggregation procedure described above is similar to map generalization, but
different from direct classification of coarse resolution images, where the integration of radiance occurs before classification. In the eastern Nebraska study area (TM image), re-classifications were performed on the coarse resolution TM images upscaled by the averaging method. The re-classifications created another series of coarse resolution land cover maps. Maximum likelihood classification algorithms were utilized in the re-classifications. To ensure that the decision rule in the re-classifications of coarse resolution images was consistent with that in the classification of full resolution TM image, the same statistics used to determine class number and class covariances in the 30m image classification were utilized. Labeling procedures identical to those described earlier were used to assign spectral classes to land cover classes. The coarse resolution land cover maps obtained by such "re-classification" (figs. 3.21, 3.22, 3.23) were compared with those generated by the aggregation method to investigate the effects of different upscaling approaches on the representation of land cover. This investigation was only conducted on the TM series maps because substantial ancillary data were used in labeling the SPOT and the AVHRR land cover maps and the generation of re-classification maps would be difficult.

3.7. Summary

This chapter discussed the selection of study areas, data used in the research and the methods of deriving coarse resolution data from the original full resolution datasets. The database established represents multi-sensor data over a broad range of spatial...
resolutions and in different ecological areas. The main components of the database include: 1) TM series data in eastern Nebraska’s predominantly agricultural area (data in this series included NDVI images and land cover maps, both upscaled using different methods, i.e., averaging NDVI directly vs. averaging DN values before calculating NDVI, and upscaling land cover using aggregation method vs. using a re-classification method); 2) SPOT series data in north-central Nebraska’s predominantly grassland area (data in this series included NDVI images upscaled using two different methods and land cover maps upscaled using the majority-rule aggregation method); and 3) AVHRR data covering the whole conterminous U.S. Included in this series were two differently upscaled NDVI datasets and the spatially aggregated land cover maps which included both the original 159-class and the grouped 25-class land cover datasets.

The aforementioned databases served as the main data sources for this study. A number of new datasets were derived (e.g., landscape heterogeneity maps and NDVI difference images) in later components of the study. These are discussed in the relevant chapters.
3.8. References


Markham, B. L., and Barker, J L., 1986, Landsat MSS and TM post-calibration dynamic ranges, exoatmospheric reflectances and at-satellite temperatures, EOSAT Landsat Technical Notes, No.1, pp.5-8.


Merchant, J. W., Yang, L., and Yang, W., 1993, Validation of continental-scale land cover data bases developed from AVHRR data, Proceedings of the Pecora XII Symposium, Sioux Falls, South Dakota, pp.63-72.


Figure 3.1 Locations of the Nebraska Study Areas
Figure 3.2 Pseudo color composite image of the Lincoln/Omaha study area (TM4 2 1, August 26, 1991)
Figure 3.3 TM NDVI image of the Lincoln/Omaha study area
Figure 3.4 Pseudo color composite image of the Niobrara Valley Preserve and vicinity (SPOT 3 2 1, June 7, 1994)
Figure 3.5 Pseudo color composite image of the Niobrara Valley Preserve and vicinity (SPOT 321, August 24, 1994)
Figure 3.6 SPOT NDVI images of the Niobrara study area
Figure 3.7 AVHRR Maximum Value Composite NDVI image of the conterminous United States (August 17-30, 1990)
Figure 3.8 The Level I ecoregion map of the conterminous United States
Figure 3.9 Land cover map of the Lincoln/Omaha study area  
(classified from the 30m resolution TM image)
Figure 3.10 Land cover map of the Niobrara study area

(Both the June and the August SPOT image were used in the land cover classification. The seventeen class map was grouped from a twenty-six class map derived by Bruce Wylie of the USGS EROS Data Center.)
Figure 3.11 The 25-class land cover map of conterminous United States
(derived from the 159-class database of Loveland et al, 1995)
Figure 3.12 Land cover map of the Lincoln/Omaha study area at 240m resolution, following the aggregation method
Figure 3.13 Land cover map of the Lincoln/Omaha study area at 480m resolution, following the aggregation method
Figure 3.14 Land cover map of the Lincoln/Omaha study area at 960m resolution, following the aggregation method.
Figure 3.15 Land cover map of the Niobrara study area at 240m resolution
Figure 3.16 Land cover map of the Niobrara study area at 480m resolution
Figure 3.17 Land cover map of the Niobrara study area at 960m resolution
Figure 3.18  Land cover map of conterminous United States
(derived from the 159-class database of Loveland et al, 1995, and aggregated to 16 km res.)
Figure 3.19 Land cover map of conterminous United States

(derived from the 159-class database of Loveland et al, 1995, and aggregated to 32 km res.)
Figure 3.20 Land cover map of conterminous United States
(derived from the 159-class database of Loveland et al., 1995, and aggregated to 64 km res.)
Figure 3.21 Land cover map of the Lincoln/Omaha study area at 240m resolution, following the re-classification method
Figure 3.22 Land cover map of the Lincoln/Omaha study area at 480m resolution, following the re-classification method
Figure 3.23 Land cover map of the Lincoln/Omaha study area at 960m resolution, following the re-classification method
CHAPTER 4

SPATIAL UPSCALING OF THE NORMALIZED DIFFERENCE VEGETATION INDEX

4.1. Introduction

Healthy green vegetation reflects strongly in the near-infrared (NIR) wavelengths of the electromagnetic spectrum, due to internal mesophyll structure, and reflects weekly in the red wavelengths because of a strong absorption by leaf chlorophyll and other pigments. This property of vegetation results in close relationships between plant biophysical variables, such as absorbed photosynthetically active radiation, and vegetation indices developed from NIR-red spectral reflectance ratios (Gausman, 1974; Sellers 1985, 1989; Tucker and Sellers, 1986; Sellers et al., 1992). Such relationships allow inference of vegetation condition from spectral radiances measured from remote sensing platforms.

The Normalized Difference Vegetation Index (NDVI), defined as the ratio of the difference between the NIR and red reflectance to the sum of the two (equation 3.3), has been used widely for research on the extent and condition of vegetation (e.g., Justice et al., 1985; Gallo and Flesch, 1989; Cihlar et al., 1991). NDVI has also been used as a primary variable in regional and continental scale land use/land cover mapping and monitoring (Norwine and Greegor, 1983; Tucker et al., 1985; Townshend et al., 1987; Loveland et al., 1991, 1995).

While the relationships between vegetation indices, including NDVI, and plant
biophysical parameters are often observed/developed from leaf or canopy level measurements/models (e.g., Sellers, 1985; Sellers et al., 1992; Verma et al., 1993), their applications often involve the use of data obtained from coarse resolution remote sensors, such as the Advanced Very High Resolution Radiometer (AVHRR). For the relationships observed at patch levels, usually on the order of meters or tens of meters, to be applied to coarse resolution satellite data (e.g., 1.1km-resolution AVHRR imagery) or climatic models (e.g., cell sizes of hundreds of kilometers), they must be scale invariant across spatial resolutions. Because spectral radiance, not NDVI, is measured by a remote sensor, two transformation processes must be examined when NDVI/biophysical parameter relationships developed at a finer scale are to be applied at a coarser scale. The first is the scaling of algorithms for estimating a biophysical parameter from NDVI, and the second is the scaling of equations for calculating NDVI from spectral reflectance (Hall et al., 1992; Friedl et al., 1995).

Assuming a plant biophysical variable, $Y$, is a function, $f$, of NDVI, and NDVI is a function, $g$, of red and near-infrared reflectances, $R_{\text{red}}$ and $R_{\text{nir}}$, at a finer resolution:

\[ Y = f(NDVI) \]

\[ NDVI = g(R_{\text{red}}, R_{\text{nir}}) \]

then the following conditions must be satisfied for spectral reflectances measured at a coarser resolution to be used to estimate $Y$:

\[ \int Y = \int f(NDVI) = f(\int NDVI) \]

and
where the integration occurs in the field of view of the coarse resolution sensor.

According to the definition of NDVI, equation (4.4) is apparently not mathematically valid. The right-hand term, NDVI measured at a coarser resolution, will usually provide lower values than the middle term, the mean NDVI of finer resolution NDVI (Justice et al., 1989). A number of investigations have focused on evaluating the degree of NDVI non-linearity in relation to spectral reflectance across spatial resolutions (e.g., Jasinski, 1990; Hall et al., 1992; Aman et al., 1992; Friedl et al., 1995; De Cola, 1997). Using field-measured data, Hall et al. (1992) and Friedl et al. (1995) observed that the bias induced by non-linearity was minor. Aman et al. (1992) found that when Thematic Mapper (TM) images (resolution=30m) were upsampled to 1-km resolution the root mean square error introduced by non-linearity was 0.0036, which was less than 1% of the mean NDVI value in their tropical savanna and temperate cropland study areas. Similar results were observed by De Cola (1997). Further examinations of more land cover types, however, were suggested (Aman et al., 1992; De Cola, 1997).

Studies such as those cited above have examined only the mean difference between the differently upsampled NDVI values. Additional research is needed to estimate the mean errors caused by non-linearity in different ecological areas, and also to link the error to spatial heterogeneity and spectral properties of the landscape. It is also important to investigate the comparability between NDVI upsampled from fine resolution sensors (e.g., TM and SPOT) to that obtained directly from a coarse resolution sensor (e.g., AVHRR). The principal objectives of this study were: 1) to estimate the bias in spatial
upsampling introduced by non-linearity in NDVI in relation to spectral reflectances for two different areas in Nebraska and throughout the conterminous U.S., and to evaluate the possible effect of bias on land cover classification; 2) to examine the impact of landscape heterogeneity on the degree of NDVI non-linearity across spatial resolutions; 3) to investigate the change in information content of NDVI across spatial resolutions; and 4) to investigate the comparability between coarse resolution NDVI derived from SPOT and TM images and NDVI obtained from AVHRR measurements.

4.2. Data and methods

The fine resolution data used in this study included two SPOT scenes acquired over the Niobrara Valley Preserve and vicinity in north-central Nebraska on June 7 and August 24, 1994, respectively, and one TM image of eastern Nebraska’s Lincoln-Omaha area acquired on August 26, 1991. The images were geo-registered to UTM coordinates by USGS/EROS Data Center (EDC). The SPOT images cover an area of about 4,500 km$^2$ (3,100km$^2$ excluding background), with 20m spatial resolution, and the TM image covers an area of about 10,000km$^2$ with 30m resolution. An AVHRR Maximum Value Composite (MVC) NDVI image for the conterminous U.S. and the red and near-infrared images used to composite the MVC NDVI image were retrieved from CD-ROMs published by EDC. The composite period of the NDVI image is from August 17 to 30, 1990, a period during which most vegetation is in full growth. The AVHRR composite period is in the same month (August), but different year, as the TM image and one of the
SPOT images. The 1990 NDVI image was used because there was a corresponding land cover database for the year. Three additional MVC NDVI images, corresponding to the TM image and two SPOT images, were obtained from the CD-ROMs for TM/AVHRR and SPOT/AVHRR comparison. The composite period corresponding to the TM image is August 16-29, 1991. Those corresponding to the two SPOT images are May 27 - June 9 and August 19 - September 1, 1994, respectively. Ancillary data related to the AVHRR MVC NDVI images in the Niobrara and Lincoln-Omaha areas were also retrieved. These include the exact date and time each MVC pixel was selected and the satellite and solar zenith angles for each pixel.

Land cover maps for the two Nebraska study areas were developed from classification of the TM and SPOT images. Nineteen land cover classes (eleven in rural areas and eight in urban areas) were identified in the Lincoln-Omaha study area. In the Niobrara study area, a twenty-six class land cover map was developed by EDC using the June and August SPOT scenes and two NDVI images calculated from bands 2 and 3 of the two scenes (Wylie, 1995, personal communication). The twenty-six classes were grouped into seventeen general classes. For the conterminous U.S., the 159-class land cover database developed jointly by EDC and the Center for Advanced Land Management Information Technologies (CALMIT), University of Nebraska-Lincoln was used (Loveland et al., 1991, 1995).

Calculation of NDVI for the TM and SPOT images involved several steps: 1) converting digital numbers (DN) into radiance, equations 3.1 and 3.4 respectively; 2) determining at-satellite reflectance from radiance, equation 3.2; and 3) computing NDVI.
from at-satellite reflectance using equation (3.3). The near-infrared and red bands used to calculate NDVI were bands 4 and 3 for the TM image and bands 3 and 2 for the SPOT images. For the AVHRR data set, only coarse resolution NDVI was calculated because full resolution (1km) data were already available from EDC. Two coarse resolution AVHRR NDVI datasets were prepared. The first was the mean NDVI ($M_{\text{NDVI}}$), obtained by averaging the full resolution NDVI images (i.e., the middle term in equation 4.4). The second was simulated NDVI ($S_{\text{NDVI}}$) measured directly by sensors at corresponding resolutions (not considering atmospheric condition and sensor characteristics other than resolution). The latter is obtained by separately averaging the full resolution red and near-infrared bands into the appropriate resolutions first and then calculating NDVI using the averaged data.

A simulation procedure similar to the one used by Jasinski (1990) was utilized to estimate the difference between $M_{\text{NDVI}}$ and $S_{\text{NDVI}}$ at various degrees of fractional vegetation cover. In that procedure, a coarse resolution pixel (AVHRR pixel) is assumed to be composed of 33x33 fine resolution pixels (TM pixel), and that all the fine resolution pixels are either soil or vegetation. The means and standard deviations for the red and near-infrared reflectance of soils are indicated as $M_{\text{red}}^{s}$, $M_{\text{nir}}^{s}$, $\text{STD}_{\text{red}}^{s}$, and $\text{STD}_{\text{nir}}^{s}$. Those for vegetation are indicated as $M_{\text{red}}^{v}$, $M_{\text{nir}}^{v}$, $\text{STD}_{\text{red}}^{v}$, and $\text{STD}_{\text{nir}}^{v}$. $M_{\text{NDVI}}$ and $S_{\text{NDVI}}$ are calculated using:
\[ M_{NDVI} = f \frac{(R_{nir}^v - R_{red}^v)}{(R_{nir}^v + R_{red}^v)} + (1-f) \frac{(R_{nir}^s - R_{red}^s)}{(R_{nir}^s + R_{red}^s)} \]  
\[ S_{NDVI} = [f R_{nir}^v + (1-f) R_{nir}^s] - [f R_{red}^v + (1-f) R_{red}^s] 
\frac{[f R_{nir}^v + (1-f) R_{nir}^s] + [f R_{red}^v + (1-f) R_{red}^s]}{[f R_{nir}^v + (1-f) R_{nir}^s] + [f R_{red}^v + (1-f) R_{red}^s]} \]  

Where \( f \) is the fractional vegetation cover, i.e., ratio of the number of vegetated fine resolution pixels to the total number of fine resolution pixels in the coarse pixel (1089=33x33). \( R_x^y \) is the reflectance of cover \( x \) (either vegetation or soil) at band \( y \) (either red or near infrared).

The number of vegetated fine resolution pixels varies from 0 to 1089, indicating fractional vegetation cover, \( f \), changing from 0.0 to 1.0 in the coarse resolution pixel. At each level of vegetation cover, \( R_x^y \) changes from \( M_x^y-STD_x^y \) to \( M_x^y+STD_x^y \). The simulation procedure produces two NDVI mean and standard deviation values at each \( f \) level, each corresponding to the mean and standard deviation of \( M_{NDVI} \) and \( S_{NDVI} \) over the \( R_x^y \) range. The two mean and standard deviation values of the coarse pixel at each \( f \) level are compared to determine the difference between the two upscaling methods and the effect of fractional vegetation cover on the degree of difference. TM pixels from bare soil and vegetation (soybeans) in the Lincoln study area were used in the simulation, from which the mean and standard deviation of the near-infrared and red reflectances representing
zero (f=0.0) and full (f=1.0) vegetation cover were obtained.

Coarse resolution NDVI values were calculated over a broad range of spatial resolutions: 120, 240, 360, 480, 600, 720, 840, 960, and 1000m for both SPOT and TM images. The SPOT series also included 40 and 80 meter resolution data, while the TM series included 60 and 90 meter resolution data. These included resolutions similar to currently-used and proposed sensor systems (e.g., AVHRR, MODIS). For the AVHRR image, NDVI values were calculated at 2, 4, 8, 16, 24, 32, 40, 48, 56, and 64km resolutions. They represent a continuous range of resolutions coarser than 1 km. The 64km upper limit was selected because 1) it is roughly equivalent to the 0.5 degree cell size used by some ecological models; and 2) the size of the image is reasonably small (45x74 pixels). At each coarse resolution level, two NDVI images were constructed using the two aforementioned upscaling methods from which a difference image between the two was generated for comparison. In addition, a landscape heterogeneity image was constructed corresponding to each NDVI difference image, and the relationship between the difference in the two NDVIs and landscape heterogeneity was examined. Landscape heterogeneity images were derived by overlaying square windows, having the dimension of a coarse resolution pixel, onto the full resolution land cover map and recording the number of land cover types within each window. Thus, the value of each pixel in the landscape heterogeneity images was the number of land cover types included in the coarse pixel.

Finally, the coarse resolution NDVI images derived from SPOT and TM images were compared, in terms of the range of values, mean and standard deviation, and
entropy, to the AVHRR MVC NDVI composited during the period when the SPOT and TM images were acquired. Correlation and regression analyses were performed between the AVHRR NDVI and those derived from SPOT and TM images.

4.3. Results and analyses

4.3.1. Difference in NDVI values upscaled using the two methods (results from simulation)

Figure 4.1 shows the results of the simulation of upscaling TM pixels of two classes (i.e., bare soil and soybean) to AVHRR pixel size. $M_{NDVI}$ calculated using equation (4.5) displays a higher mean value than $S_{NDVI}$ computed from equation (4.6) no matter what the fractional vegetation cover. The result is in agreement with the conclusions of Justice et al. (1989). The mean difference between the two NDVIs is 0.015. The root mean square error (RMSE) is 0.017, which is about 4.6% of minimum-maximum range (i.e., 0.333-0.702) of the coarse resolution pixel and about 3.5% of the average $S_{NDVI}$. The maximum difference between the two, occurring when fractional vegetation cover ranges from 0.48 to 0.59, is about 0.024 NDVI values. This is about 6.5% of the minimum-maximum range and about 4.8% of the average $S_{NDVI}$ from all levels of fractional vegetation cover. The variation caused by varying $R^*$ in each $f$ level is larger in $S_{NDVI}$ than in $M_{NDVI}$ when $f$ is low (<0.6). The pattern reverses after $f$ exceeds 0.6. The maximum difference between the standard deviations of $S_{NDVI}$ and $M_{NDVI}$ is 0.003, which is about 6.5% of the standard deviation of $S_{NDVI}$ at the corresponding $f$ level.
These results indicate that the absolute differences in both NDVI and its variance, caused by the two upscaling methods, are generally not significant. However, the differences are usually about 5% of the possible NDVI range of the coarse pixel and the actual $S_{NDVI}$ (or $M_{NDVI}$) value at corresponding levels of vegetation cover.

4.3.2. Difference in NDVI values upscaled using the two methods (results from real images)

Calculations performed on the TM and SPOT images yielded similar results (fig. 4.2). The overall difference between the two upscaled coarse resolution NDVI series was not significant in most cases (less than 0.02) (figs. 4.3 and 4.4). In the Lincoln/Omaha area, with TM series images, the mean differences for the entire image at 240, 480, and 1000m levels were 0.007, 0.011, and 0.014, respectively, and the RMSE were 0.015, 0.019, and 0.020. The RMSE values were about 2.5% to 3.3% of the mean values and 12.6% to 22.9% of standard deviations of the coarse NDVI images at the corresponding resolutions (fig. 4.2a). Figure 4.2b shows the maximum difference ($\text{MAX}_{\text{diff}}$) observed for each resolution level. The $\text{MAX}_{\text{diff}}$ is the absolute maximum pixel value in an entire difference image $M_{NDVI} - S_{NDVI}$. The maximum differences ranged from 0.20 to 0.38, which were on the same order as the mean NDVI in the area. Although the $\text{MAX}_{\text{diff}}$, which represented the extreme value of a difference image, did not reflect the general difference between the two NDVIs, the existence of large $\text{MAX}_{\text{diff}}$ indicated that the difference could be very significant in some areas. $\text{MAX}_{\text{diff}}$ decreased as resolution became coarser due
to the overall decrease in NDVI contrast among pixels. The percentage of pixels having varying degrees of NDVI difference were calculated at each coarse resolution level. As resolution became coarser, more pixels displayed differences in the two NDVI values. At 240m resolution, 13.4% pixels showed 0.02-0.05 difference and 1.2% displayed greater than 0.05 difference. These two numbers increased to 24.4% and 1.6% at 480m, and to 35.9% and 2.1% at 1000m (fig. 4.5a).

Similar results were observed in the Niobrara study area with SPOT imagery (fig. 4.6). For the August image, the mean differences for the entire image at 240, 480, and 1000m levels were 0.005, 0.006, and 0.007, and the RMSE were 0.009, 0.009, and 0.010. These were about 2% of the mean values and 9.8% to 12.8% of the standard deviations of the coarse NDVI images at corresponding resolutions (fig. 4.6a). The MAX_diff values ranged from 0.011 to 0.049, which were much smaller than those observed in the TM image, but were still several times as large as the mean difference (fig. 4.6b). The percentage of pixels having varying degrees of NDVI differences were shown in Figure 4.5b. Overall, the differences between M_{NDVI} and S_{NDVI} in the Niobrara study area was smaller than those for the Lincoln/Omaha area (figs. 4.2-4.6). This was probably due to the lower NDVI range and lower contrast (less variation) in spectral reflectance among the different land cover types on the Niobrara site where the predominant cover types were native grasses. In the Lincoln/Omaha area, where cropland, urban areas, and large bare fields are intermingled, the contrast in NDVIs among different cover types was much larger.

The mean M_{NDVI} and mean S_{NDVI} averaged for the entire AVHRR image (i.e., the
conterminous U.S.) show almost no difference (fig. 4.7a), although the landscape of the conterminous U.S. is more heterogeneous than that in either of the two Nebraska study areas. The smaller $M_{NDVI}-S_{NDVI}$ in the AVHRR data can be attributed to the stronger averaging effect among the original AVHRR pixel (1km resolution) as compared to the TM and SPOT pixels. When one square kilometer land area is represented by only one measurement, there is an averaging effect, eliminating the contrast among subpixel components which could be detected by finer resolutions such as those of SPOT and TM. This was also reflected in the results obtained from the TM and SPOT data (figs. 4.4a and 4.6a). The gradient of increases in NDVI difference became much smaller after the pixel size exceeded 400m. The off-nadir view, which resulted in the actual pixel size being larger than its nadir field of view (1.1km), also contributed to the reduced contrast between $M_{NDVI}$ and $S_{NDVI}$. The large MAX$_{diff}$ and RMSE values in Figures 4.7b and 4.7a were caused by pixels in the boundary of the conterminous U.S., where pixels with the background filling values (zero) were included in the upscaling procedure.

4.3.3. Relationship between NDVI difference and land cover

In the Lincoln/Omaha area, the within-class standard deviation of the 19 land cover types at 1000m resolution ranged from 0.02 to 0.13 (excluding a single pixel class, "grass/tree") with an average of 0.056 (fig. 4.8). In the Niobrara study area, the within-class standard deviations for the 17 land cover classes were between 0.01 to 0.08 (excluding a single pixel class, "other crops") for both the June and the August scenes at
the 1000m level, with an average of about 0.030. The mean differences between $M_{\text{NDVI}}$ and $S_{\text{NDVI}}$ in both study areas were smaller than the within-class standard deviations of NDVI for all land cover types (except the single pixel classes). At finer resolution levels, the differences in the two NDVIs were smaller while the mean within-class variances were larger. Therefore, the differences in NDVI caused by the differing upscaling methods were not likely to influence land cover classification. Similar observations were also made for the AVHRR image. At 64km resolution, the within-class standard deviation for the 100 classes, each having at least two pixels, ranged from 0.005 to 0.119 with a mean of 0.044, much larger than the difference in the two NDVIs.

The patterns of the $M_{\text{NDVI}} - S_{\text{NDVI}}$ difference were closely related to the spectral reflectances of land cover types. In both of the Nebraska study areas, coarse resolution pixels along the rivers displayed the largest difference between $M_{\text{NDVI}}$ and $S_{\text{NDVI}}$ (figs. 4.2 and 4.3). This seems to be related to large differences in NDVI values among the land cover types near the rivers (water, sandbars, riparian forests, shrubs and grasses). Comparison of $M_{\text{NDVI}} - S_{\text{NDVI}}$ differences with landscape heterogeneity showed that the heterogeneity of a coarse resolution pixel, measured by the number of land cover types contained in the pixel, played an important role in determining the magnitude of the $M_{\text{NDVI}} - S_{\text{NDVI}}$ difference. As landscape heterogeneity increased (as indicated by the number of classes in each pixel), pixels having large differences (0.02) increased significantly (fig. 4.8). Therefore, although the differences in the two differently upscaled NDVIs were generally small, these could be significant in certain areas, depending both on the diversity and spectral properties of the land cover types included in the coarse resolution
pixels under consideration.

### 4.3.4. Change in image information content

Range, mean, standard deviation, and entropy were used to represent the image information content. Range indicates the bounding of the image values and mean represents the average value of an image. Standard deviation depicts the variations of image values around the mean, while entropy is a measure of spread of image values over the available range. Entropy is defined as:

$$H = - \sum_{k=1}^{n} p(k) \log_2 p(k)$$  \hspace{1cm} (4.5)

where $p(k)$ is the probability of a digital value $k$,

$n$ is the number of pixels in the image.

The mean values of both $M_{NDVI}$ and $S_{NDVI}$ were relatively invariant to resolution (<2.6%, relative to the original value at full resolution, for all images at all resolution levels) (fig. 4.9). As expected, the standard deviation and entropy decreased as the resolution became coarser. In the Niobrara area, the standard deviations of the June and August images decreased to 35.15% and 24.22% when resolution changed from 20m to 1km, while the entropy decreased to 12.5% and 6.36%, respectively (fig. 4.9). The results indicated substantial loss in information content, regarding both the variations of NDVI values and the distribution of NDVI within its overall range. In the Lincoln/Omaha area,
the decrease in entropy from 30m to 1km was similar to that in the June Niobrara image, 12.97%, but the standard deviation decreased more substantially (41.39%) (fig. 4.10). This was likely due to the stronger NDVI contrast among different land cover types in the original 30m image as compared with the Niobrara images. The entropy of the AVHRR image decreased by 4.92% while the standard deviation decreased by 16.67% when it was upscaled to 64km (fig. 4.11). The information loss was less than that of the TM and SPOT images, but they were not directly comparable due to differences in resolution ranges. It was clear, however, that information loss could affect land cover classification because the amount of information reduction far exceeded the between-class variances of different land cover classes.

AVHRR MVC NDVI exhibited consistently lower values as compared to the upscaled SPOT and TM NDVI, but the magnitudes of the minimum/maximum, and the mean±STD ranges between the corresponding NDVIs were close, and the spatial patterns between the NDVI images were similar (Table 4.1 and figs. 4.12, 4.13; note that there was only one pixel in the TM image that had a negative NDVI value). The differences in standard deviations between the corresponding AVHRR/SPOT and AVHRR/TM pairs were 0.0%, 7.7%, and 6.9%, and those in entropy were 1.4%, 0.7% and 3.1%. These results demonstrate that the upscaled SPOT and TM NDVI information content is similar to that of the AVHRR MVC NDVI although pixel values are not directly comparable. Pixel to pixel correlation analyses between the AVHRR NDVI and the upscaled NDVI resulted in r values of 0.45 (SPOT, June), 0.79 (SPOT, August) and 0.81 (TM) (figs. 4.14 - 4.16). Regressions between the two NDVIs exhibited varied intercept and slopes.
Table 4.1 The range, mean, standard deviation and entropy of AVHRR MVC NDVI and $S_{\text{NDVI}}$ upscaled from SPOT and TM images (1km resolution)

<table>
<thead>
<tr>
<th>Sensor</th>
<th>Date</th>
<th>Minimum</th>
<th>Maximum</th>
<th>Mean</th>
<th>STD</th>
<th>Entropy</th>
</tr>
</thead>
<tbody>
<tr>
<td>SPOT</td>
<td>6/7/1994</td>
<td>0.38</td>
<td>0.76</td>
<td>0.53</td>
<td>0.06</td>
<td>4.42</td>
</tr>
<tr>
<td>AVHRR</td>
<td>5/27-6/9/1994</td>
<td>0.29</td>
<td>0.66</td>
<td>0.49</td>
<td>0.06</td>
<td>4.48</td>
</tr>
<tr>
<td>SPOT</td>
<td>8/24/1994</td>
<td>0.35</td>
<td>0.73</td>
<td>0.50</td>
<td>0.08</td>
<td>4.82</td>
</tr>
<tr>
<td>AVHRR</td>
<td>8/19-9/1/1994</td>
<td>0.21</td>
<td>0.57</td>
<td>0.36</td>
<td>0.07</td>
<td>4.79</td>
</tr>
<tr>
<td>TM</td>
<td>8/26/1991</td>
<td>-0.08</td>
<td>0.80</td>
<td>0.61</td>
<td>0.09</td>
<td>5.05</td>
</tr>
<tr>
<td>AVHRR</td>
<td>8/16-8/29/1991</td>
<td>0.14</td>
<td>0.74</td>
<td>0.55</td>
<td>0.08</td>
<td>4.89</td>
</tr>
</tbody>
</table>

The relationship between the AVHRR and the upscaled NDVI may be affected by a number of factors, including band widths/locations, viewing/illumination geometry, and atmospheric conditions. About 70% of the pixels in the June AVHRR MVC NDVI image of the Niobrara study site were acquired at off-nadir viewing zenith angles greater than 44°. In the August image, 85% of the pixels were obtained within 20° off-nadir. In the Lincoln/Omaha area, 88% of the MVC NDVI pixels were within 34° off-nadir view. By comparison, the TM pixels were measured at viewing angles within ±7.7° from nadir and the SPOT images were acquired at near nadir view (pushbroom scanner with 4.13° field of view pointed at nadir). To evaluate the effects of off-nadir observation, those pixels corresponding to greater than 20° off-nadir viewing angle were excluded and correlations were performed for the remaining near-nadir viewed pixels. The resultant $r$ values were 0.53, 0.79, and 0.85, respectively. The most significant increase in $r$ value occurred for the June SPOT/AVHRR pair (from 0.45 to 0.53), followed by the AVHRR/TM pair (from 0.81 to 0.85), while that for the August AVHRR/SPOT pair remained unchanged.
result was in agreement with the fact that most of the June AVHRR pixels were acquired at large off-nadir viewing angles and that the majority of the August (Niobrara scene) AVHRR pixels were observed from near-nadir.

Among all three AVHRR/SPOT and AVHRR/TM pairs, only a small percentage of the June AVHRR pixels were obtained on the same day as the SPOT image. Of these, 323 pixels were observed within 20° off-nadir. Correlation between these 323 pixels and their corresponding upscaled SPOT pixels resulted in an $r$ value of 0.55, which was not substantially different from the $r$ value ($r=0.53$) obtained without restricting the observation date. The result may indicate that viewing zenith angle is a more important factor than observation date (within relatively short periods of time, e.g., two weeks) in affecting the comparability between NDVI$s obtained from different sensors. The higher $r$ value obtained from the AVHRR/TM pair (as compared to the AVHRR/SPOT pairs) might be related to illumination conditions. All the AVHRR MVC NDVI pixels in the Niobrara area were obtained from late afternoon satellite observations. The solar zenith angles of the scenes ranged from 48° to 73°, while the SPOT images were acquired when solar zenith angles were less than 34°. In the Lincoln/Omaha area, all MVC NDVI pixels were selected from early afternoon overpasses. The solar zenith angles ranged from 37° to 46°, with more than 93% between 39° and 42°, almost the same as that of the TM image (about 41°). These results indicate the importance of consistency in both solar zenith and viewing zenith angles in NDVI comparison. However, NDVI is subject to contamination by other factors, especially atmospheric conditions. No atmospheric data were obtained for use in this research. Thus, additional studies are needed to determine
the factors influencing the relationships between NDVI upscaled from high resolution sensors such as SPOT and TM and that derived from coarse resolution sensors such as AVHRR.

4.4. Summary and conclusions

Although NDVI is scale variant, the overall difference between NDVI averaged from the high resolution measurements and that directly calculated from coarse resolution reflectance is small. The magnitude of the difference between the two is usually less than 0.02 NDVI, which is less than 6% of the range of the NDVI under consideration. The difference is smaller than that caused by other contaminants such as water vapor, usually reducing NDVI by about 0.02, and aerosol, reducing NDVI by about 0.06 to 0.12 (Holben, 1986). The results are consistent in both study areas and are in agreement with those obtained by other investigators in different ecological regions (e.g., Aman et al., 1992). The difference between the two NDVIs will have little effect on the accuracy of land cover classification because the magnitude of the difference is smaller than the within-class variation of each land cover class under consideration. However, the difference does not evenly distribute throughout the images. It is dependent on both the characteristics and the number of land cover types within the upscaled coarse resolution pixels. In heterogeneous areas, NDVI differences larger than 0.02 are likely to be observed. At 1000m resolution, where there are almost no pure pixels, a significant percentage of pixels in both study areas had differences between $M_{\text{NDVI}}$ and $S_{\text{NDVI}}$ larger
than 0.02. In areas where the landscape is highly heterogeneous and there is significant contrast in spectral reflectances among different land cover components, the bias introduced in upscaling due to non-linearity of NDVI may exceed 0.05, which is usually about 15% of the NDVI range in Nebraska’s predominantly cropland and grassland environment.

The information content (measured by the minimum/maximum range, standard deviation, and entropy) of upscaled NDVI from SPOT and TM images was equivalent to that of the AVHRR MVC NDVI, suggesting that NDVI is generally transferable across spatial scales and that NDVI values derived from the different sensors will provide similar ability to portray green vegetation patterns if they are spatially scaled to the same resolution. Pixel to pixel comparison between AVHRR MVC NDVI and spatially upscaled NDVI is, however, not feasible. In addition to generally lower values of MVC NDVI (shown in the intercepts of the regression lines), the regression slopes also varied. The correlation between MVC NDVI and NDVI upscaled from fine resolution data was subject to the influence of viewing and illumination zenith angles. Substantial increases in r values (from 0.45 to 0.53) were found when correlation was performed on near-nadir (<20°) pixels as compared to that on the mostly off-nadir (>44°) pixels. The varied regression slopes and r values indicated that the relationships between NDVI and biophysical variables observed in SPOT and TM imagery are not directly transferable to AVHRR data. Additional research is needed to assess the effects of other factors, such as atmospheric conditions, on NDVI, its spatial upscaling, and its relationship with plant biophysical parameters.
4.5. References


Figure 4.1 Coarse resolution NDVI upscaled using two different methods.

It was assumed that the coarse resolution pixel was composed of 1089 fine resolution pixels that were either pure soil or pure vegetation. Values of red (RED) and near infrared (NIR) reflectance of bare soil and vegetation (soybean) were taken from the Lincoln/Omaha study area. For bare soil, the values ranged from 0.17 to 0.24 (RED) and from 0.36 to 0.45 (NIR). For vegetation, the values ranged from 0.05 to 0.09 (RED) and from 0.36 to 0.45.
Figure 4.2 Differences in the two NDVI values upscaled differently (TM data, the Lincoln/Omaha study area, August 26, 1991).
(a) Mean difference and root mean square error between the two NDVI values.
(b) Maximum difference in the two NDVI values.
Figure 4.3 Difference between 240m resolution TM NDVI directly averaged from the full resolution (30m) one and that calculated from degraded red and near-infrared reflectances (Lincoln/Omaha area, August 26, 1991).
Figure 4.4 Difference between 240m resolution NDVI directly averaged from full resolution (20m) one and that calculated from degraded red and near-infrared reflectances (Niobrara area, August 24, 1994).
Figure 4.5 Percentage of pixels with varying degrees of Diff(NDVI) across resolutions.
(a) TM data in the Lincoln/Omaha study area, August 26, 1991
(b) SPOT data in the Niobrara study area, August 24, 1994
Figure 4.6 Differences in the two NDVI values upscaled differently (SPOT data, the Niobrara study area, August 24, 1994).

(a) Mean difference and root mean square error between the two NDVI values.
(b) Maximum difference in the two NDVI values.
Figure 4.7 Differences in the two NDVI values upscaled differently (AVHRR data, the conterminous U.S., August 17-30, 1990).

(a) Mean difference and root mean square error between the two NDVI values.

(b) Maximum difference in the two NDVI values.
Figure 4.8 Percent of pixels having greater than 0.01 Diff(NDVI) changes with the number of the land cover types included in coarse resolution pixels. (a) TM data in the Lincoln/Omaha study are, August 21, 1991. (b) SPOT data in the Niobrara study area, August 24, 1994. (c) AVHRR data in the conterminous United States, August 17-30, 1990.
Figure 4.9 Changes in standard deviation and entropy with spatial resolution (SPOT data in the Niobrara study area, June 7, 1994).
Figure 4.10 Changes in standard deviation and entropy with spatial resolution (TM data in the Lincoln/Omaha study area, August 26, 1991).
Figure 4.11 Changes in standard deviation and entropy with spatial resolution (AVHRR data in the contiguous United States, August 17-30, 1990).
Figure 4.12 Comparison between the upscaled TM NDVI (8/26/1991) and the AVHRR MVC NDVI (8/16-29, 1991) at 1km resolution
Figure 4.13 Comparison between the upscaled SPOT NDVI (8/24/1994) and the AVHRR MVC NDVI (8/19-9/1, 1994) at 1km resolution
Figure 4.14 Comparison between AVHRR MVC NDVI and upscaled SPOT NDVI, the Niobrara study area, June 7, 1994. Gray level indicate number of pixels falling within the corresponding NDVI regions, the light gray in the cluster center being the highest (>80 pixels) and the medium gray in the outmost of the cluster the lowest (<5 pixels).

Figure 4.15 Comparison between AVHRR MVC NDVI and upscaled SPOT NDVI, the Niobrara study area, August 24, 1994. Gray level indicate number of pixels falling within the corresponding NDVI regions, the light gray in the cluster center being the highest (>80 pixels) and the medium gray in the outmost of the cluster the lowest (<5 pixels).
Figure 4.16 Comparison between AVHRR MVC NDVI and upscaled TM NDVI, the Lincoln/Omaha study area, August 26, 1991. Gray level indicate number of pixels falling within the corresponding NDVI regions, the light gray in the cluster center being the highest (>80 pixels) and the medium gray in the outmost of the cluster the lowest (<5 pixels).
CHAPTER 5
CHARACTERIZING THE EFFECTS OF SPATIAL RESOLUTION ON LANDSCAPE INDICES: A SIMULATION STUDY

5.1. Introduction

Regional and global scale land cover mapping usually involves the use of coarse resolution remotely sensed data such as that from the Advanced Very High Resolution Radiometer (AVHRR) (e.g., Townshend et al., 1987; Loveland et al., 1991). AVHRR Local Area Cover (LAC) images have 1km spatial resolution and Global Area Cover (GAC) images have 4km resolution. Pixels of such coarse resolution images are not likely to be pure in most cases (Skole and Tucker, 1993; Steininger, 1996). It is desirable to estimate the proportional areas of component land cover types in such coarse pixels. On the other hand, certain eco-climatological models (e.g., global circulation models), use cell sizes of hundreds of kilometers on a side and land cover maps derived from remotely sensed data need to be spatially upscaled for model input (Wessman, 1992; DeFries et al., 1997). In such applications, it is also preferable that proportional areas of different land cover types be estimated so that model parameters can be adjusted according to the land cover composition of each cell.

Estimation of cover composition in mixed-pixels has been performed using linear spectral unmixing models (e.g., Cross et al., 1991; Quarmby et al., 1992; Foody and Cox, 1994; Puyou-Lascassies et al., 1994). Recent studies in scaling issues have shown that proportional areal errors, when one land cover map is spatially rescaled to another, are
closely related to the spatial pattern of the landscape under investigation (Woodcock and Strahler, 1987; Townshend and Justice, 1988; Moody and Woodcock, 1994). Attempts have been made to calibrate quantitatively the proportional errors in coarse resolution land cover maps, either upscaled from fine resolution maps or directly classified from coarse resolution remotely sensed data, using models that include landscape structure indices that measure patch size, perimeter to area ratio, and patch interspersion (e.g., Moody and Woodcock, 1995; Moody, 1996). However, for such models to be used in calibrating a coarse resolution land cover map, one needs first to derive landscape indices from a fine resolution land cover map of the same area. These, unfortunately, are often either not available or not complete. A two-step regression procedure was proposed by Mayaux and Lambin (1995) to directly use landscape indices derived from coarse resolution land cover maps, but the model development and assessment was based on measurements obtained from fine resolution land cover maps. Therefore, generalization of such calibration models over different landscape types and spatial resolutions is necessary before operational applications. As a first step, characteristics of landscape indices themselves over a range of spatial resolutions need to be examined.

Numerous studies in landscape ecology have shown that the spatial structure of landscapes has profound influences on various ecological processes, including population dynamics, nutrition and material flow, and biodiversity (e.g., Forman and Godron, 1986; Franklin and Forman, 1987; Turner, 1987, 1989; Wu, et al., 1993). Spatial scale is one of the most important issues in characterizing landscape structure and there has been an increasing interest in research on scaling and landscape patterns (Turner et al., 1989a,
There are two different but related objectives in scale and pattern research: identifying the scale at which landscape processes work, and investigating the influence of scale on pattern characterization. The first aspect of the scale issue focuses on how processes change across scales and the second aspect deals with how observation scale affects the parameters describing a landscape. The second aspect is more closely related to land cover characterization using remotely sensed data because landscape indices are often derived from maps classified using remotely sensed data (e.g., Simmons et al., 1992; Wichham and Ritters, 1995; Frohn, 1996). Much research has been conducted on the influences of spatial scale on landscape indices (e.g., Nellis and Briggs, 1989; Turner et al., 1989b; Benson and MacKenzie, 1995; Qi and Wu, 1996). Although general patterns of scale effects have been recognized, many important questions, especially the interaction between scale and other factors, remain largely unanswered. Further research is needed to systematically address such issues.

Neutral models, based on percolation theory, have often been used to study spatial patterns of landscapes (e.g., Gardner et al., 1987; O'Neill et al., 1988, 1992; Turner et al., 1989c). In neutral models, land cover classes in a landscape are considered one at a time and all other classes are viewed as one single class. The land cover under consideration is the foreground (e.g., red areas in figure 5.1) while all other classes are background (e.g., blue areas in figure 5.1). This allows examination of the indices derived for a specific cover class. Gustafson and Parker (1992) documented the changes in several landscape indices with proportional area of land cover class using a neutral model.
approach. Benson and MacKenzie (1995) investigated changes in landscape indices calculated from water bodies delineated from images with different spatial resolutions. These studies indicate that landscape indices are affected by both proportional area of land cover and spatial resolution.

In this study, I investigate, in a neutral model context, variations in landscape indices across a range of spatial resolutions with different proportional areas of foreground land cover (i.e., the land cover of interest) and patch properties (i.e., the size and shape of the smallest patches which constitute the land cover). The principal objectives of this study are to characterize the integrated effects of spatial resolution, proportional area, and patch property on landscape indices, and to systematically examine how and why landscape indices are affected simultaneously by various factors.

5.2. Methods

Landscapes with different proportional foreground areas were generated using a computer program. The simulated landscapes were of 1024 by 1024 pixels, assuming a pixel size of 30m. Thus, the area of each landscape was about 31 by 31 km. Each landscape consisted of two land cover types: foreground cover, which occupied $P$, proportion of the whole landscape, and background cover, which represented all other land covers. Landscape metrics were calculated for the foreground land cover. Three different types of basic patch units were used in the generation of the simulated landscapes. The basic patch units in the first landscape were single pixels. That is, the
foreground land cover consisted of mosaics of independent, randomly distributed single pixels. The pixels, however, were allowed to neighbor one another and to constitute larger irregular patches, especially at high $P_1$. The basic patch units for the second landscape were squares. The size of the square was 13x13 pixels, i.e., 390mx390m (close to the size of many agricultural fields in Nebraska). Similar to the first landscape type, the basic patch units (squares) were allowed to neighbor and overlay one another. Thus, complicated patch shapes may result. The basic patch units comprising the foreground land cover in the third landscape were circles with radii of 13 pixels. The size of each circle was approximately the size of most center pivot irrigation systems (i.e., 400m radius circles). The basic patch units were allowed to overlap one another and create irregular patterns. Each of the three landscape types was generated with proportional foreground area, $P_1$, changing from 0.01 to 0.70 at increments of 0.05 (0.04 between the first two, i.e., from 0.01 to 0.05). Thus, there were 15 landscapes in each type. For simplicity, the three sets of landscapes will be referred as DPLS (dot-patch, i.e., single pixel patch, landscape), SPLS (square-patch landscape), and CPLS (circle-patch landscape). It should be noted, however, that the actual shape of patches may be irregular at high percentage of occurrence for all the three types of landscapes. Figures 5.1 to 5.18 show some of the simulated landscapes.

Spatial aggregation was performed on each landscape to create simulated coarse resolution landscapes. A majority rule was used in the creation of the coarse resolution maps. Non-overlapping square windows, each representing one pixel in an output coarse resolution landscape, were overlaid on the original landscapes. Each pixel in the output
landscape was either assigned as foreground or background land cover type, according to the proportional area of each cover type in the window. If each of the two cover types occupied 50% of the window, the output cover type was randomly assigned. Ten coarse resolutions were selected: 60, 90, 120, 240, 360, 480, 600, 720, 840, and 960m. The series included resolutions similar to those of actual remote sensors (e.g., Multispectral Scanner, Moderate Resolution Imaging Spectroradiometer, and AVHRR). Three landscape types with fifteen levels of proportional foreground area and at eleven resolutions resulted in a total of 495 landscapes for the study.

Five landscape indices were used to examine the effects of both spatial resolution and percentage land cover, and the interaction among resolution, percentage cover, and patch type. The five indices included: largest patch index (LPI), mean patch size (MPS), patch size standard deviation (PSSD), landscape shape index (LSI), and mean nearest neighbor distance (MNN). The first three indices were used to describe the patch size characteristics in a landscape while the last two were used to depict the properties of patch shape and spatial location (distance) between patches. Calculation of landscape indices was conducted using Fragstats, a spatial analysis software package (McGarigal and Marks, 1993). Additional descriptions and equations for calculating these indices can be found in Appendix A. It should be noted that some other metrics, such as patch interspersion and contagion, have been proven useful in characterizing a landscape. They were not used in this study because only one foreground land cover type was considered. Examination of those metrics was performed for actual landscapes classified from satellite images in separate studies (see chapters six and seven).
5.3. Results and analyses

5.3.1. Effect of proportional foreground on landscape metrics at full resolution

The indices derived from the three landscapes at the original image resolution are shown in figures 5.19 to 5.21. The curves depict the effects of the property of a basic patch unit and percentage occurrence of a land cover type on landscape metrics. Percolation theory has shown that there is a critical value of land cover proportion in a landscape at which the probability of forming a percolation patch, a patch that connects the opposite edge in a rectangular landscape simulated by a two dimensional array, changes from 0.0 to 1.0 (Gardner et al., 1987). This critical value has been shown to be 0.5928 for infinitely large arrays (Stauffer, 1985). The P_l values corresponding to the formation of predominant patches, shown as the sudden increase in LPI, in the three landscapes reflected the existence of such critical values. But critical values were different among the different landscapes.

In the CPLS landscape, the proportional foreground landcover occupied by the largest patch (indicated by the LPI value) increased linearly from 6% to 95% as P_l changed from 0.55 to 0.70 (fig. 5.19a), indicating the formation of the predominant patch in the landscape. The P_l values between 0.55 and 0.75 can be considered approximately the critical value as shown by the percolation theory. In the SPLS landscape, a predominant patch constituting 77% (i.e., LPI value) of the foreground landcover formed when P_l changed from 0.60 to 0.65 (fig. 5.20a) and the LPI increased to 93% when P_l reached 0.70. The SPLS exhibited a more obvious critical P_l value as compared to the
CPLS. The most prominent critical $P_i$ was observed in the DPLS, where the LPI increased from 0.07% to 94% when $P_i$ changed from 0.40 to 0.45 (fig. 5.21a). In this landscape, patches were composed either of individual pixels or of clumps of a few pixels at low $P_i$ (<0.45), resulting in very small LPI. Due to the randomness in the introduction of new patches (i.e., single pixels), the existing patches tended to be more or less evenly distributed throughout the whole landscape, creating the condition for a sudden, large scale connection among patches. When $P_i$ changed from 0.40 to 0.45, most of the individual patches were connected with each other and they formed a single large patch which constituted almost the entire foreground landscape.

Although critical $P_i$ values existed in all the three landscapes, substantial differences were observed, the DPLS showed both a steeper slope in LPI increase and a lower critical $P_i$, attributed to the fact that there was a higher degree of clumps in the foreground land cover in SPLS and CPLS where the basic patch units themselves were already clumps of pixels. In SPLS and CPLS, the increase in $P_i$ tended to be completed by just a few new patches, which were not likely to link simultaneously all previous unconnected patches. The clump of the existing patches also reduced the randomness in the distribution of occupied foreground landcover in the landscape and thus a higher $P_i$ was needed to create a predominant patch. The size difference of the basic patch units between SPLS and CPLS was also the cause for the difference in $P_i$ values in the two landscapes.

The mean patch sizes (MPS) of the three landscapes behaved similarly through the range of percentage foreground (figs. 5.19b, 5.20b, and 5.21b). There were also critical
P₁ phenomena for changes in MPS but they were not as prominent compared to those shown in LPI. The largest increases in MPS occurred when $P_1$ was between 0.60 to 0.70 for all the three landscapes. The CPLS and SPLS landscapes displayed almost identical patterns of change in MPS, but CPLS showed consistently larger MPS due to its larger basic patch unit. The difference in MPS between the two landscapes increased with the increase of $P_1$. In the DPLS, the critical $P_1$ value for MPS did not correspond to that for LPI because of the existence of a large number of single pixel patches at LPI’s critical $P_1$, i.e., 0.45, which substantially reduced MPS. Although MPS was related to the size of the basic patch unit, the MPS values were similar in SPLS and in DPLS after $P_1$ reached 0.60, due to the fact that the largest patch in DPLS was larger than that in SPLS. These results indicate that the relationship between mean patch size of a landscape and the size of the basic patch unit is not linear when complicated patches form from the basic patch units.

The behavior of patch size standard deviation (PSSD) in the three landscapes was similar to that for MPS (figs. 5.19c, 5.20c, and 5.21c). However, the absolute difference in PSSD among the landscapes was smaller than that in MPS. This is because PSSD is a measure of deviation from the mean value. While increasing the basic patch unit raised the size of patches formed from that unit, the variation between those patches did not correspond to the size of the basic unit in a linear way. The increases in PSSD and in LPI corresponded to one another, because the formation of the largest patch also caused increasing differences in the sizes of patches. This resulted in larger PSSD in the DPLS landscape than in the other two landscapes when $P_1$ was between 0.40 and 0.60, where
the largest patch in DPLS formed.

Landscape shape index (LSI) was inversely related to the size of the basic patch unit in a landscape (figs. 5.19d, 5.20d, and 5.21d). It was smallest for the CPLS landscape and largest for the DPLS, regardless of the $P_1$ value. LSI increased as the percentage of foreground increased and reached its maximum before a dominant patch was formed in a landscape. This was because at low $P_1$ most unit patches had little or no connection with other patches and the landscape was composed primarily of simple patches (i.e., circles, squares, or dots). As more unit patches were introduced into a landscape, more complicated patches emerged due to the high degree of irregular connections. This resulted in an increasing LSI value. As a dominant patch formed in the landscape, newly-introduced patches were more likely to reduce patch shape complexities by filling the gaps than to create more irregular patches. Therefore, the largest LSI tended to appear before a dominant patch emerged. LSI appeared to be a second order function of $P_1$. It can be regressed to a second order polynomial with $r^2$ greater than 0.9 for each landscape (Table 5.1). However, the regression coefficients in

<table>
<thead>
<tr>
<th>Landscape</th>
<th>$a_0$</th>
<th>$a_1$</th>
<th>$a_2$</th>
<th>$r^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>DPLS</td>
<td>145.546</td>
<td>13.401</td>
<td>0.174</td>
<td>0.916</td>
</tr>
<tr>
<td>SPLS</td>
<td>10.536</td>
<td>1.243</td>
<td>-0.013</td>
<td>0.969</td>
</tr>
<tr>
<td>CPLS</td>
<td>6.497</td>
<td>0.804</td>
<td>0.009</td>
<td>0.980</td>
</tr>
</tbody>
</table>

Table 5.1 Regression between LSI and $P_1$ for the three landscapes.

The regressions are in the form of $LSI = a_0 + a_1 P_1 + a_2 P_1^2$. 
the three landscapes differed greatly, indicating that a regression model derived from one landscape is not likely applicable to another landscape.

The mean nearest neighbor distance (MNN) exhibited very similar patterns for the three landscapes (figs. 5.19e, 5.20e, and 5.21e). The most substantial decrease in MNN for all three took place when the percentage foreground was below 0.20, with a 65-70% drop occurring when \( P \) was between 0.05 to 0.10. MNN began to level off after \( P \) reaches 0.20. This was because most new patches at low \( P \) tended to be more or less uniformly distributed into the background area, thus significantly decreasing the distance between patches. After \( P \) reached about 0.20, most newly-introduced patches tended to overlay or adjoin previously existing patches. Therefore, the rate of decrease in patch distance decreased. In each landscape, MNN was linearly related to the reciprocal of percentage foreground. The regression of MNN against \( 1/P \) resulted in \( r^2 \) greater than 0.98 for all three landscapes (Table 5.2). Similar to the models between LSI and \( P \), the regression slopes differed greatly among the different landscapes. Therefore an empirical regression model derived from one landscape cannot be applied directly to another.

Table 5.2 Regression between MNN and the reciprocal of \( P \) for the three landscapes. The regressions are in the form of \( \text{MNN}=a_0+a_1(1/P) \).

<table>
<thead>
<tr>
<th>Landscape</th>
<th>( a_0 )</th>
<th>( a_1 )</th>
<th>( r^2 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>DPLS</td>
<td>28.866</td>
<td>96.896</td>
<td>0.996</td>
</tr>
<tr>
<td>SPLS</td>
<td>71.576</td>
<td>1638.025</td>
<td>0.985</td>
</tr>
<tr>
<td>CPLS</td>
<td>100.423</td>
<td>3522.279</td>
<td>0.989</td>
</tr>
</tbody>
</table>
5.3.2. Integrated effect of spatial resolution and proportional area on landscape metrics

Figures 5.22 to 5.24 show changes in the largest patch index (LPI) across spatial resolutions. Generally, LPI increased with resolution. The changes, however, were closely related both to the property of the basic patch units comprising the landscape and to the percentage of foreground landcover. LPI increased nearly monotonically as resolution coarsened for the CPLS landscape, but the magnitude of change differed for different $P_i$ values (fig. 5.22). The largest increase in LPI occurred when $P_i$ equals 0.50. In this case, the patches were not well connected at the original resolution (30m). As the resolution became coarser, the previously disconnected patches quickly formed dominant patches (at 480m and thereafter), resulting in a significant increase in LPI. For the higher $P_i$, i.e., 0.70, because the original patches were already well connected, there was only limited increase in LPI. Note that the increase in LPI was not as significant for landscapes with $P_i$ lower than 0.50 because the large patches formed at coarse resolutions did not dominate the landscapes. The curve (at $P_i$=0.50) demonstrated the effect of resolution on the critical $P_i$ value at which a predominantly large patch formed. The greatest increase in LPI occurred when $P_i$ is between 0.55 and 0.70 (fig. 5.21a) at 30m resolution while the largest LPI increase occurred when $P_i$ was between 0.30 and 0.50 at coarser resolutions (≥240m). Figure 5.25a shows that the $P_i$ values were actually around 0.40 to 0.50 attributed to the fact that spatial aggregation increased the patch size so that there was a higher probability for a predominant patch to form at relative lower $P_i$. 
Similar results were obtained for the SPLS landscape where critical $P_i$ decreased from the original 0.60 (fig. 5.20a) to between 0.40 and 0.50 (fig. 5.25b) after pixel size exceeded 240m. In SPLS, LPI values for the two smallest $P_i$s (0.01 and 0.05) reached 100% at 720m and 840m resolutions, respectively (fig. 5.23). This was because all patches in these two landscapes are connected at the respective resolutions. However, this did not happen for the CPLS landscape. The basic patch unit in CPLS was larger than that in SPLS. Therefore, fewer patches were needed for a certain $P_i$ level at original resolution. As resolution coarsened, these patches did not completely connect with one another, as occurred in the SPLS, due to a fewer number of patches.

At full resolution (30m), patches in SPLS were scattered more randomly throughout the landscape than those in CPLS were. They were, therefore, more likely to connect with each other when the resolution became coarser, as compared to the patches in the CPLS landscape. These results indicated that the size of the basic patch unit was critical because the patches were more likely to be eliminated at coarse resolutions if the basic patch unit was too small, as was the case of DPLS.

In the DPLS landscape, no foreground landcover patches were left after the pixel size reached 360m when $P_i$ was below 0.50. On the other hand, at $P_i$ of 0.70 and higher, the LPI was always 100% because all patches were connected for the entire range of resolutions (fig. 5.24). Distinctive fluctuations of LPI in DPLS were observed at 60m and 120m pixel resolutions for a $P_i$ of 0.50 because the distribution of original patches (i.e., single pixels) was sparse and the connections between patches happened to be broken at these two resolutions, thus reducing LPI values. After the pixel size reached 240m, the
dominant patch was compact enough to withstand spatial aggregation and there was no substantial decline in LPI.

The pattern of change in mean patch size (MPS) was simpler than that in the LPI (figs. 5.26 to 5.28). In most instances, there were nearly monotonic increases in MPS as resolution coarsened. Higher $P_i$ levels resulted in larger MPS than did low $P_i$ level because the patches were more likely to connect with each other at high $P_i$ even though the size of the basic patch unit was the same in each landscape. In the CPLS and SPLS landscapes, the rate of increase in MPS was also higher for large $P_i$ values but was not significant when $P_i$ was below 0.30 (figs. 5.26 and 5.27). When the resolution was below 600m, there was little or no increase in MPS in the CPLS landscape at $P_i \leq 0.30$. This was because the size increase in some patches was compensated by the elimination of foreground pixels at the edges of other patches. Similar results were observed for the SPLS landscape at resolutions below 480m. In the DPLS landscape, the foreground pixels at low $P_i$ ($\leq 0.30$) were quickly eliminated, due to its small basic patch size, when resolution became coarser (fig. 5.28). At very high $P_i$ (0.70), foreground pixels were connected into a single large patch after resolution reached 120m and the MPS reached its maximum value.

Theoretically, patch size should be in linear relationship to the second order of resolution because linear resolution size was used in this study to indicate pixel size and the area of a pixel was the square of linear resolution (i.e., a 30m resolution pixel has an area of $(30\times30)m^2$). However, no such observation was obtained for any of the simulated landscapes at any $P_i$ levels, and curves showing MPS change were different among
landscapes and $P_i$ levels, although some of them behaved similarly. This suggests that irregular changes in the sizes and spatial relationships among patches during aggregation can not be modeled with a single function.

The patch size standard deviation (PSSD) largely followed the same pattern as MPS (figs. 5.29 to 5.31), because variations in patch size tended to increase as the patch size became larger. It was also likely to increase as spatial resolution and/or percentage foreground increased because coarser resolution and/or higher $P_i$ would result in larger patches. On the other hand, since PSSD was related to the number of patches in a landscape, it could decrease due to the decline in the number of patches at coarser resolutions. In extreme situations (i.e., when the number of patches was less than three or when all patches had the same size), PSSD became zero (see fig. 5.30 -- the four $P_i$ levels 0.01, 0.05, 0.10, and 0.70 in the SPLS landscape).

The landscape shape index (LSI) decreased monotonically as the spatial resolution became coarser in most situations (figs. 5.32 to 5.34). For the CPLS and SPLS landscapes, the change in LSI was mostly linear or near linear (figs. 5.32 and 5.33). Regression of the LSI against pixel resolution showed significant linear relationships in all cases. The $r^2$ values were all greater than 0.9, except when $P_i$ values were equal to 0.01, where $r^2$ values were 0.86 for the CPLS and 0.89 for the SPLS (Tables 5.3 and 5.4). The regression slopes steepened as $P_i$ increased in both these landscapes (Tables 5.3 and 5.4), although the initial (i.e., 30m resolution) LSI was not linearly related to $P_i$ (Table 5.1 and figs. 5.19d and 5.20d). This was perhaps due to the fact that the landscape shapes were relatively simple at low $P_i$ levels (e.g., $P_i<0.1$) and therefore spatial...
Table 5.3  Regression between LSI and spatial resolution for the CPLS landscape at selected $P_i$ levels. The regressions are in the form of $LSI = a_0 + a_1 \text{Res.}$

<table>
<thead>
<tr>
<th>$P_i$</th>
<th>$a_0$</th>
<th>$a_1$</th>
<th>$r^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.01</td>
<td>5.516</td>
<td>-0.003</td>
<td>0.857</td>
</tr>
<tr>
<td>0.05</td>
<td>11.491</td>
<td>-0.006</td>
<td>0.906</td>
</tr>
<tr>
<td>0.10</td>
<td>15.452</td>
<td>-0.008</td>
<td>0.929</td>
</tr>
<tr>
<td>0.30</td>
<td>22.859</td>
<td>-0.012</td>
<td>0.990</td>
</tr>
<tr>
<td>0.50</td>
<td>24.479</td>
<td>-0.015</td>
<td>0.990</td>
</tr>
<tr>
<td>0.70</td>
<td>20.752</td>
<td>-0.017</td>
<td>0.972</td>
</tr>
</tbody>
</table>

Table 5.4 Regression between LSI and spatial resolution for the SPLS landscape at selected $P_i$ levels. The regressions are in the form of $LSI = a_0 + a_1 \text{Res.}$

<table>
<thead>
<tr>
<th>$P_i$</th>
<th>$a_0$</th>
<th>$a_1$</th>
<th>$r^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.01</td>
<td>9.102</td>
<td>-0.001</td>
<td>0.893</td>
</tr>
<tr>
<td>0.05</td>
<td>19.509</td>
<td>-0.021</td>
<td>0.935</td>
</tr>
<tr>
<td>0.10</td>
<td>26.682</td>
<td>-0.028</td>
<td>0.945</td>
</tr>
<tr>
<td>0.30</td>
<td>38.171</td>
<td>-0.032</td>
<td>0.986</td>
</tr>
<tr>
<td>0.50</td>
<td>38.128</td>
<td>-0.032</td>
<td>0.968</td>
</tr>
<tr>
<td>0.70</td>
<td>31.987</td>
<td>-0.035</td>
<td>0.935</td>
</tr>
</tbody>
</table>

aggregation had less effect. As $P_i$ increased, landscapes became increasingly complicated and the spatial aggregation tended to have a greater influence on LSI due to considerable simplification of patch shapes at coarser resolutions. On the other hand, when $P_i$ was greater than 0.50, the spatial aggregation tended to quickly expand the area of the foreground and substantially simplified the shapes, thus greatly reducing the LSI values.
(largest negative slope). The DPLS landscape exhibited a very different pattern in LSI change across spatial resolutions (fig. 5.34). LSI values dropped abruptly when pixel cell size increased because the original landscape was comprised mostly of single pixels. As the spatial resolution became coarser, those single pixel patches (i.e., patches not connecting with other patches) were eliminated, considerably reducing LSI values. At low $P_1 (<0.50)$, there was essentially no foreground land cover at coarse resolutions and thus no LSI ($LSI = 0$). When $P_1$ reached 0.30 and larger, LSI in the DPLS landscape could be modeled with a linear regression against the reciprocal of resolution with $r^2$ greater than 0.95 (Table 5.5).

Table 5.5 Regression between LSI and spatial resolution for the DPLS landscape at selected $P_1$ levels. The regressions are in the form of $LSI = a_0 + a_1(1/\text{Res})$.

<table>
<thead>
<tr>
<th>$P_1$</th>
<th>$a_0$</th>
<th>$a_1$</th>
<th>$r^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.01</td>
<td>-10.987</td>
<td>2673.03</td>
<td>0.753</td>
</tr>
<tr>
<td>0.05</td>
<td>-23.150</td>
<td>5822.77</td>
<td>0.783</td>
</tr>
<tr>
<td>0.10</td>
<td>-30.334</td>
<td>7976.75</td>
<td>0.823</td>
</tr>
<tr>
<td>0.30</td>
<td>-30.063</td>
<td>11928.64</td>
<td>0.974</td>
</tr>
<tr>
<td>0.50</td>
<td>1.693</td>
<td>10994.71</td>
<td>0.998</td>
</tr>
<tr>
<td>0.70</td>
<td>-20.319</td>
<td>8179.34</td>
<td>0.947</td>
</tr>
</tbody>
</table>

The mean nearest neighbor distance ($MNN$) increased as the spatial resolution coarsened, except when there was no foreground land cover due to spatial aggregation or when there was only one patch in a landscape, which resulted zero $MNN$ (figs. 5.35 to 5.37). The rate of increase in $MNN$ was inversely related to the proportion of foreground
land cover, but there was no definite relationship. Generally, MNN grew linearly as resolution coarsened when $P_i$ exceeded 0.1. At low $P_i$, MNN values increased more quickly at resolutions coarser than 480m than at finer resolutions for the CPLS and SPLS landscape because the number of the foreground land cover patches decreased more quickly at the coarser resolutions. In the DPLS landscapes, foreground land cover disappeared at resolutions coarser than 240m except when $P_i$ was around 0.50, where MNN increased near linearly as the resolution became progressively coarser.

5.4. Summary and Conclusions

In this study, five landscape indices representing patch size, shape, and inter-patch distance, were examined across a range of spatial resolutions, proportional landcover areas, and for three sets of landscapes simulated with different basic patch units. The results obtained demonstrate the effects of spatial resolution, proportion of foreground land cover area, and the property of basic patch unit on the five selected landscape metrics in neutral landscapes (i.e., landscapes with only one foreground land cover type considered at a time). More importantly, the results document the integrated effects of resolution, landcover proportion, and basic patch unit. The analyses of the behaviors of the metrics for different types of basic patch units and across different spatial resolutions and $P_i$ levels in neutral landscapes will help the interpretation and understanding of those metrics obtained from real landscapes.

The major findings obtained in this study can be summarized as follows:
1) Changes in the same landscape index across proportional foreground land cover areas were generally similar for the three sets of landscapes at the initial resolution. However, noticeable influences of the property of the basic patch unit on the indices existed and the degree of such influences differed among different indices. A predominant patch, reflected in a sudden increase in LPI, formed in a landscape when proportional foreground land cover, \( P_i \), reached a certain level. The critical \( P_i \) value for such predominant patch formation was most evident, and was also lowest in the dot-patch, i.e., single pixel patch, landscape (DPLS). The value became less obvious in the circular patch landscape (CPLS) which had the largest basic patch unit.

2) Changes in mean patch size (MPS) and patch standard deviation (PSSD) across \( P_i \) largely followed the pattern of largest patch index (LPI) and the critical \( P_i \) values for MPS and PSSD also existed, though they were less prominent as compared to the LPI. Both MPS and PSSD displayed significant increases when \( P_i \) reached 0.60 in all three landscape sets, but there were marked differences in the slopes and in the absolute values of the two indices among the three landscapes. These differences appeared to be related to disparities in the size and spatial distribution of the basic patch units used to generate the landscapes.

3) The landscape shape index can be modeled, at \( r^2 \) values greater than 0.9, with a second order polynomial regression against \( P_i \). However, the regression coefficients were significantly different among the three landscapes, and the model developed for one landscape seemed not applicable to another. LSI was largest for the DPLS landscape and smallest for the CPLS landscape. The maximum value of LSI occurred before a
predominant patch formed, usually 0.10 to 0.15 before the critical $P_1$ for LPI.

4) The patterns of change in mean nearest neighbor distance (MNN) across $P_1$ were very similar among the three landscapes but the absolute magnitudes of MNN values varied. MNN was linearly related, at $r^2$ values greater than 0.98, to the reciprocal of $P_1$ in all three landscapes. The regression models were, however, not intra-applicable among landscapes because the regression coefficients were significantly different, and no relationships among coefficients derived from different landscapes were found.

5) The effect of spatial resolution on landscape indices was related both to the properties of the basic patch units comprising a landscape and to the percentage of foreground land cover, $P_1$. Critical $P_1$ values at which predominant patches formed decreased as the resolution coarsened. LPI generally increased with resolution, but the magnitude differed in different landscapes and at different $P_1$ levels. In the DPLS landscapes, LPI reached its maximum (i.e., 100% at the initial pixel resolution), when $P_1$ values were equal to 0.70 and thus it did not increase with resolution. At low $P_1$ values (e.g., 0.01 and 0.05) the foreground land cover was quickly eliminated as the resolution coarsened and, thus, there was no LPI (LPI=0). By contrast, LPI values for the two smallest $P_1$s (0.01 and 0.05) in the SPLS landscape reached 100% at 720m and 840m resolutions, respectively.

6) MPS generally increased as spatial resolution became coarser. The rate of the increase was related to the $P_1$ value and differed among landscapes. Larger $P_1$ values tended to result in higher rates of increase in the MPS, except when $P_1$ was below 0.30. At very high $P_1$ levels, MPS did not increase if the foreground became a single patch.
7) PSSD tended to the increase as spatial resolution and/or percentage foreground increased because large patches formed at coarser resolutions and/or higher \( P_i \) values, resulting in a large value of the PSSD. However, PSSD could decrease at very coarse resolution due to the decline in the number of patches.

8) The LSI had generally a negative linear relationship with pixel resolution in the SPLS and CPLS landscapes. Regressions between LSI and resolution in these two landscapes resulted in \( r^2 \) greater than 0.9 for all but \( P_i \) values of 0.01, and the slopes of the regression lines increased as \( P_i \) increased. In the DPLS landscape, LSI values decreased abruptly as the pixel resolution coarsened. At moderate and high \( P_i \) values (\( \geq 0.30 \)), LSI could be regressed against the reciprocal of resolution at \( r^2 \) greater than 0.95.

9) The effect of pixel resolution on MNN was influenced greatly by \( P_i \) and the basic patch unit. MNN grew linearly as the resolution coarsened when \( P_i \) was larger than 0.10 in SPLS and CPLS. At lower \( P_i \) levels, it increased near exponentially with resolution.

The following conclusions can be drawn based on the aforementioned findings:

1) There existed a critical value of the proportional land cover area in a landscape at which a predominant patch will form in the landscape. The critical value was lowest, but its effect was most prominent, for the landscape with the smallest initial basic patch unit (i.e., DPLS). The critical \( P_i \) value decreased as the pixel resolution coarsened, but did not fall below 0.40.

2) MPS and PSSD increased with \( P_i \) while the MNN value decreased. The rates and magnitudes of change of these indices across the \( P_i \) range varied among different
landscapes and at different resolutions. MNN possessed a strong linear relationship with the reciprocal of $P_i$ ($r^2 > 0.9$), but a linear regression model derived from one landscape was not applicable to another. LSI has a second order relationship with $P_i$. The largest LSI occurred before a predominant patch appears in a landscape.

3) LPI, MPS, PSSD, and MNN usually increased with spatial resolution, but there were considerable differences in the patterns of change in these indices among different landscapes and $P_i$ values. There seems to be no single relationship (either linear or nonlinear) between these indices and resolution. When $P_i$ exceeded 0.01, LSI may be modeled with a linear regression against resolution for the SPLS and CPLS landscapes. The effect of resolution was stronger for landscapes with higher proportional foreground area. However, there was no functional relationship between the LSI and resolution because of the great differences in models among different landscapes.

This research represents a systematic investigation into the scale dependency of landscape indices. It contributes to a better understanding of the integrated effect of spatial resolution, proportional foreground landcover area, and the properties of the basic patch unit on landscape metrics. Such an understanding is important because landscape metrics are widely used in ecological studies as well as in calibrating landcover maps derived from remotely sensed data. The effects of various factors on landscape metrics must be examined before such metrics can be used in ecological modeling and calibrating remotely sensed data. Although statistical relationships have been sought among some landscape indices, spatial resolution, and proportional area, this work is largely qualitative. Future research will focus on a more quantitative analyses.
5.5. References


Figure 5.1 Circular patches (p=0.05, Res=30m)

Figure 5.2 Circular patches (p=0.05, Res=480m)
Figure 5.3 Circular patches (p=0.55, Res=30m)

Figure 5.4 Circular patches (p=0.55, Res=480m)
Figure 5.5 Circular patches (p=0.65, Res=30m)

Figure 5.6 Circular patches (p=0.65, Res=480m)
Figure 5.7 Square patches ($p=0.05$, Res=30m)

Figure 5.8 Square patches ($p=0.05$, Res=480m)
Figure 5.9  Square patches (p=0.60, Res=30m)

Figure 5.10  Square patches (p=0.60, Res=480m)
Figure 5.11 Square patches (p=0.65, Res=30m)

Figure 5.12 Square patches (p=0.65, Res=480m)
Figure 5.13 Dotted patches (p=0.05, Res=30m)

Figure 5.14 Dotted patches (p=0.05, Res=480m)
Figure 5.15 Dotted patches (p=0.40, Res=30m)

Figure 5.16 Dotted patches (p=0.40, Res=480m)
Figure 5.17 Dotted patches (p=0.45, Res=30m)

Figure 5.18 Dotted patches (p=0.45, Res=480m)
Figure 5.19 Changes of the five landscape indices with the proportional foreground land cover areas, results of the landscapes composed of circular patch units.
Figure 5.20 Changes of the five landscape indices with the proportional foreground land cover areas, results of the landscapes composed of square patch units.
Figure 5.21 Changes of the five landscape indices with the proportional foreground land cover areas, results of the landscapes composed of single pixel patch units.
Figure 5.22 Change of largest patch index (LPI) across the spatial resolutions at selected foreground proportional areas, results of landscapes composed of circular patch units.
Figure 5.23 Change of largest patch index (LPI) across the spatial resolutions at selected foreground proportional areas, results of landscapes composed of square patch units.
Figure 5.24 Change of largest patch index (LPI) across the spatial resolutions at selected foreground proportional areas, results of landscapes composed of single pixel patch units.
Figure 5.25 Change of largest patch index (LPI) across proportional foreground land cover areas at selected spatial resolutions. Results of landscapes composed of (a) circular patch units; (b) square patch units; and (c) single pixel patch units.
Figure 5.26 Change of mean patch size (MPS) across spatial resolutions at selected foreground proportional areas, results from landscapes composed of circular patch units.
Figure 5.27 Change of mean patch size (MPS) across spatial resolutions at selected foreground proportional areas, results from landscapes composed of square patch units.
Figure 5.28 Change of mean patch size (MPS) across spatial resolutions at selected foreground proportional areas, results from landscapes composed of single pixel patch units.
Figure 5.29 Change of patch size standard deviation (PSSD) across spatial resolutions at selected foreground proportional areas, results from landscapes composed of circular patch units.
Figure 5.30 Change of patch size standard deviation (PSSD) across spatial resolutions at selected foreground proportional areas, results from landscapes composed of square patch units.
Figure 5.31 Change of patch size standard deviation (PSSD) across spatial resolutions at selected foreground proportional areas, results from landscapes composed of single pixel patch units.
Figure 5.32 Change of landscape shape index (LSI) across spatial resolutions at selected foreground proportional areas, results from landscapes composed of circular patch units.
Figure 5.33 Change of landscape shape index (LSI) across spatial resolutions at selected foreground proportional areas, results from landscapes composed of square patch units.
Figure 5.34 Change of landscape shape index (LSI) across spatial resolutions at selected foreground proportional areas, results from landscapes composed of single pixel patch units.
Figure 5.35 Change of mean nearest neighbor distance (MNN) across spatial resolutions at selected foreground proportional areas, results from landscapes composed of circular patch units.
Figure 5.36 Change of mean nearest neighbor distance (MNN) across spatial resolutions at selected foreground proportional areas, results from landscapes composed of square patch units.
Figure 5.37 Change of mean nearest neighbor distance (MNN) across spatial resolutions at selected foreground proportional areas, results from landscapes composed of single pixel patch units.
6.1. Introduction

Reliable information on land cover is critical for research on global change. Increasingly, we rely on satellite remote sensing to derive such information (Townshend et al., 1991). Remote sensing observations of land cover have, for example, been incorporated into regional and global scale eco-climatological models to obtain estimates of CO₂ exchange and regional evapotranspiration and photosynthesis (Fung et al., 1987; Running et al., 1989). There are, however, significant unresolved problems in scaling remotely sensed data to work within ecosystem models. Climate models, for example, frequently simulate ecosystem processes on a scale of 100x100 km² (Bolin, 1988). General circulation models (GCM) (e.g., Manabe and Stouffer, 1980) describe processes in the atmosphere and the oceans using horizontal grid cells a few hundred kilometers on a side. Currently, global models that require land cover information as a boundary condition do not have the capability of incorporating pixel level information about land cover (DeFries et al., 1997). In almost all circumstances, remote sensing measurements need to be aggregated to mesoscale grids.

Spatial aggregation may, however, result in substantial differences in prescribed land cover and thus produce significant errors in the estimation of ecoclimatological parameters through modeling (e.g., Oleson et al., 1996). In other applications, such as
the validation of land cover classifications, comparisons between land cover datasets having different spatial resolutions are also needed (e.g., Merchant et al., 1993). It is, therefore, important to understand the characteristics of land cover data represented at different spatial resolutions and to estimate possible biases introduced due to changes in spatial resolution. The purpose of this study is to examine the "representativeness" (i.e., the degree of true representation) of the land cover spatially upscaled from full resolution datasets and to examine the factors influencing such upscaling.

6.2. Data and methods

The major datasets used in this study included: 1) a land cover dataset derived from a TM image of eastern Nebraska’s Lincoln/Omaha area; 2) a land cover dataset derived from two SPOT images of north-central Nebraska’s Niobrara Valley Preserve and vicinity; 3) a conterminous United States land cover database derived from 19 biweekly Maximum Value Composite (MVC) AVHRR NDVI images; 4) three NDVI images corresponding to the three land cover datasets; and 5) an ecoregion map provided by the U.S. Environmental Protection Agency (EPA). Full descriptions of these datasets can be found in chapter 3.

Each full resolution land cover dataset was upscaled to a number of coarse resolution levels using the majority rule method, the most widely used spatial aggregation method (e.g., Turner et al., 1989; Costanza and Maxwell, 1994; Moody and Woodcock, 1994, Oleson et al., 1996). The coarse resolutions of the TM and SPOT derived land
cover datasets included 120, 240, 360, 480, 600, 720, 840, and 960m. In addition, the SPOT series also included 40 and 80 meter data, while the TM series included 60 and 90 meter data. The AVHRR land cover datasets were derived at 2, 4, 8, 16, 24, 32, 40, 48, 56, and 64km grid sizes. Two levels of class groupings (i.e., class aggregation level) were adopted for the AVHRR land cover, one with the original 159-class and the other a 25-class derived from the 159-class.

The EPA ecoregion map was used to stratify the upscaling of the AVHRR land cover datasets. The four ecoregions selected were Omernik's (1987) region #4 -- Great plains, #5 -- Western and southern desserts basins and ranges, #6 -- Northwestern mostly coniferous forested mountains, and #9 -- Eastern temperate forests. The four regions account for about 92% of the area of the conterminous U.S.

At each coarse resolution level, an overall "consistency rate" and a Kappa statistic were used to quantify the agreement between the full resolution and upscaled land cover data. The consistency rate (CR), defined as the percentage of pixels that were labeled exactly the same cover types in both datasets, is the ratio of the sum of the diagonal elements in an error matrix between the full and the upscaled data to the total number of pixels being examined (fig. 6.1). Although CR is the simplest and most frequently used index, it is not without problems because some degree of agreement may occur by chance alone (Monserud and Leemans, 1992). Congalton et al. (1983), therefore, proposed to use Kappa statistics for analysis of error matrices. Kappa statistics are used to evaluate chance agreement by using the marginal totals in an error matrix:

\[ Kappa = \frac{(p_o - p_e)}{(1-p_e)} \]
where $p_0$ is overall agreement and $p_e$ is chance-expected agreement (Cohen, 1960).

Although hypothesis testing is straightforward with Kappa (Fleiss, 1981), it is rarely useful when large sample sizes (e.g., 7,778,854 pixels for the AVHRR-derived 1km conterminous U.S. land cover) are used since these will almost always result in significant differences for any comparison (Monserud and Leemans, 1992). A more useful method is to define the range of Kappa values corresponding to the degree of agreement. In this study, the correspondence between Kappa values and degree of agreement proposed by Monserud and Leemans (1992) was used (Table 6.1).

Table 6.1 Range of the Kappa statistics and degree of agreement

<table>
<thead>
<tr>
<th>Kappa statistics</th>
<th>Degree of agreement</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt; 0.05</td>
<td>No</td>
</tr>
<tr>
<td>0.05 - 0.20</td>
<td>Very poor</td>
</tr>
<tr>
<td>0.20 - 0.40</td>
<td>Poor</td>
</tr>
<tr>
<td>0.40 - 0.55</td>
<td>Fair</td>
</tr>
<tr>
<td>0.55 - 0.70</td>
<td>Good</td>
</tr>
<tr>
<td>0.70 - 0.85</td>
<td>Very good</td>
</tr>
<tr>
<td>0.85 - 0.99</td>
<td>Excellent</td>
</tr>
<tr>
<td>0.99 - 1.00</td>
<td>Perfect</td>
</tr>
</tbody>
</table>

The consistency rate and Kappa statistics describe the overall agreement between full and upscaled land cover datasets. They do not provide information about the degree of class mixture within an aggregated coarse pixel. In this study, the number of land cover types in each upscaled coarse pixel was used to indicate the heterogeneity of the coarse pixel (fig. 6.1). The percentage of coarse pixels having different numbers of land
cover types in an upscaled land cover dataset reflect the overall heterogeneity of the coarse resolution land cover dataset. The heterogeneity index, however, still does not demonstrate completely the characteristics of a coarse pixel. For example, if the sixteen 20m resolution pixels in an 80m coarse pixel belong to two different land cover types, the distribution of the pixel numbers among the two cover types may be 8-8, 9-7, 10-6, until 15-1. Although the heterogeneity index is always equal to two in these combinations, the different distribution of pixel numbers represent significantly different characteristics of the coarse pixel. Another index, degree of dominance, was therefore utilized to indicate the actual proportional area of a majority cover type in a coarse pixel. The index was defined as the ratio of the number of full resolution pixels of a majority class to the total number of full resolution pixels in a coarse resolution pixel (fig. 6.1). The index portrays how dominant a majority class is in a coarse resolution pixel. By definition, a 100% dominance indicates that all the full resolution pixels in a coarse pixel are of the same cover type. That is, the coarse pixel is a "pure pixel", assuming each full resolution pixel is a pure pixel.

The first step in land classification is to categorize the pixel using usually statistical clustering of image DN values, reflectance, NDVI, or other spectral indices. For example, the 159-class conterminous U.S. land cover database (Loveland et al., 1991) was developed from an initial 70-class NDVI clustering. Statistical clustering (or other classification algorithms) is mostly based on the variance of the DN (or reflectance, NDVI, etc.) values of an image. The larger the image variance, the more classes will be clustered. When a high resolution image is spatially upscaled to a coarse resolution one,
the variance decreases as the resolution coarsens (i.e., the image information content decreases). In turn, the image’s ability to accurately portray land cover classes decreases, given the condition that other ancillary data are the same. Therefore, there is an association between the image information content and the accuracy of land cover classification. In this study, the variations in standard deviation (a measure of image information content) of the NDVI across resolutions were compared to the degree of agreement between the original and upscaled land cover databases to explore the potential relationship between image information content and the accuracy of upscaled land cover.

6.3. Results and analyses

6.3.1. SPOT-derived land cover in the Niobrara Preserve and vicinity

The consistency rate and Kappa statistics computed between the original SPOT-derived land cover and the upscaled coarse resolution data displayed similar logarithmic patterns (figs. 6.2a and 6.2b). The most significant decrease in consistency between the original and the aggregated data occurred when resolution changed from 20m to about 240m (note that the consistent rate was 100% at 20m, i.e., no aggregation). At 40m level, about 78 percent of the pixels were labeled to the same class as at the original resolution. The substantial decrease in consistency rate between 20m and 40m was due to the existence in the original classification map of many single-pixel patches. The consistency rate dropped to about 63% at 120m. Nevertheless, the agreement between the full and the coarse resolution data, according to Kappa values, was considered good or very good
(Table 6.1) when resolution size was below 120m. The slopes of the consistency and Kappa curves leveled off after resolution reached 240m. There was little change in the consistency rate and the Kappa value after 480m and the agreement could be characterized as poor (Kappa<0.4).

Figure 6.3 shows the heterogeneity of upscaled land cover at selected resolutions. Although the overall consistency rate at 40m resolution was about 80%, there were only about 44% "pure" coarse pixels. Another 43% were composed of two different land cover types. About 0.5% pixels were composed of four land cover types, indicating that there were no dominant cover types and the land cover types for these coarse pixels were assigned randomly to one of the component types. At 80m resolution, "pure" coarse resolution pixels dropped to only about 15% and most coarse resolution pixels were composed of two or three land cover types. As resolution became more coarse, the peaks of the curves decreased and moved toward the more heterogeneous direction. The peak values of the curves followed logarithmic patterns similar to those shown in the consistency rate and Kappa statistics.

Figure 6.4 shows percentages of pixels having different degrees of dominance at selected resolutions. At 40m resolution, about 44 percent (i.e., 56 subtracted from 100 in the X-axis of the graph) of the upscaled pixels possessed 100% dominance (i.e., they were pure pixels), while about 28 percent of the pixels were assigned land cover types that comprised only 50% of the area of the coarse resolution pixels. When the pixel size reached 120m, about one-half of the upscaled pixels were represented by classes that comprised less than 58% of the pixel. After resolution exceeded 240m, degree of
dominance for most coarse resolution pixels (more than half) decreased to below 50%. These results suggest that when resolution cell size reached a certain level (240m in this case), a majority cover type that represents a coarse resolution pixel may actually comprise a minority proportional area of the pixel. The degree of dominance, of course, is also related to the total number of land cover classes and spatial heterogeneity of the landscape in the study area.

6.3.2. TM-derived land cover in the Lincoln/Omaha study area

Figures 6.5a and 6.5b portray the consistency rate and Kappa statistics for the original TM-derived land cover and the aggregated coarse resolution pixels in the Lincoln/Omaha study area. The overall patterns of the two statistics were similar to those observed for the SPOT classification in the Niobrara study area. Both consistency rate and Kappa statistics exhibited logarithmic curves across all resolutions. However, the curves derived from the TM upscaling started to level off at about 480 to 600 meters, while those obtained from the SPOT upscaling began to level off at about 240 to 480 meters. This was likely due to the fact that the original TM classification map had a larger pixel size (30m) as compared with that of the SPOT classification (20m). At 60m resolution, about 76 percent of the pixels were labeled as the same class as at the original 30m resolution. The overall consistency rate dropped to about 65% at 120m. Kappa statistics indicated that the agreement between the full and the coarse resolution data was good at 120m and finer resolutions. These results were also very similar to those
obtained in the Niobrara study area. The Kappa values at more coarse resolutions, i.e., >360m, were, however, lower in the Lincoln/Omaha study area. This was probably due to the high degree of interspersion of the land cover types in Lincoln/Omaha’s urban areas and to the more general cover type definitions in the Niobrara study area.

The heterogeneity of upscaled land cover at selected resolutions is shown in Figure 6.6. About 41 percent of the pixels at 60m resolution were "pure" pixels. This decreased to about 22 percent at 90m where the majority of the pixels were either two- or three-class mixed-pixels. At the 120m level, more than 68 percent (sum of the percentages for x>2 in the graph) of pixels were mixed-pixels of three or more classes. As the cell size reached 480m and larger, most pixels contained six or more classes, indicating significant mixture of land cover types. There was little change in pixel heterogeneity at cell sizes between 480m and 960m, which was different from the results obtained for the Niobrara study area where the peak heterogeneity changed from 8 to 11 between these two resolutions. This was because the 19 land cover classes in the Lincoln/Omaha study area were actually located in two different areas, i.e., urban and rural areas. A coarse pixel could contain, even at the coarsest resolution (960m), at most 11 classes (the total number of rural classes), except for a very few pixels at the urban/rural boundary. This was also reflected by the very low (close to zero) percentage of pixels that contained 11 or more classes.

Figure 6.7 shows the percentages of pixels having different degrees of dominance at selected resolutions in the Lincoln/Omaha study area. The percentage of pixels having 100% dominance (i.e., "pure" pixels) decreased from 41 (i.e., 59 subtracted from 100 at
the X-axis) at the 60m resolution to about 14 at the 120m level, and to only about 3 at 240m. There were essentially no "pure" pixels after the resolution reached 480m. At 60m resolution, about one-third (37 percent) of the pixels were assigned the land cover types that comprised only 50% of the area of the coarse pixels. After resolution exceeded 240m, degree of dominance for most pixels decreased to below 50%.

These results were similar to those observed in the Niobrara study area with the SPOT classification. They again suggested that the areal errors introduced by assigning a majority class to a coarse resolution pixel may well exceed 50% of the total area of the coarse pixel, depending on the size of the pixel. It is interesting to note that although the upscaled land cover datasets in the Niobrara study area seemed more heterogeneous (fig. 6.3) than those in the Lincoln/Omaha study area (fig. 6.6), a comparison between figures 6.4 and 6.7 demonstrated that the overall degree of dominance was higher in the Niobrara area, especially at resolutions coarser than 240m. This is due to the more detailed class definition and interspersion of different classes in the Lincoln/Omaha study area. These results suggest that the representativeness of a coarse resolution land cover dataset upscaled from a finer resolution should be examined from more than one perspective.

6.3.3. AVHRR-derived 159-class land cover of the conterminous U.S.

The consistency rates and Kappa statistics for the original 1km 159-class AVHRR-derived conterminous U.S. land cover dataset and the aggregated coarse resolution products (from 2 to 64km) are shown in figure 6.8. Both curves displayed logarithmic
patterns similar to those in the TM and SPOT classifications of the Nebraska study areas.

The consistency rate and the Kappa statistics had essentially the same values (by a factor of 100). The largest decrease in the Kappa statistics (and also the consistency rate) occurred when pixel size was below 16km. The curve began to level off at 16km. The agreement between the full and the coarse resolution land cover datasets could be categorized as good and very good, when the pixel size was 8km or smaller. The Kappa value dropped to below 0.40 (poor agreement) when the pixel resolution reached 64km.

The relationships between the Kappa values and spatial resolution were not directly comparable to those observed in the upscaling of the TM- and SPOT-derived land cover data due to differences in pixel sizes. However, the Kappa values can be compared by the levels of aggregation, i.e., the ratio of upscaled pixel size to the full resolution pixel size. Using this method, the Kappa value at 2km for the AVHRR data can be compared to the Kappa value at 60m for TM data and 40m for SPOT data, and so on. Table 6.2 summarizes the Kappa values of the three datasets. Values of the missing aggregation levels, e.g., AVHRR classification at 3km, were interpolated from neighboring levels.

<table>
<thead>
<tr>
<th>Agg. level</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>6</th>
<th>8</th>
<th>12</th>
<th>24</th>
<th>32</th>
<th>48</th>
<th>64</th>
</tr>
</thead>
<tbody>
<tr>
<td>SPOT</td>
<td>0.75</td>
<td>0.67</td>
<td>0.62</td>
<td>0.56</td>
<td>0.53</td>
<td>0.47</td>
<td>0.40</td>
<td>-0.37</td>
<td>0.34</td>
<td>NA</td>
</tr>
<tr>
<td>TM</td>
<td>0.71</td>
<td>0.67</td>
<td>0.57</td>
<td>0.51</td>
<td>0.45</td>
<td>0.38</td>
<td>0.27</td>
<td>0.24</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>AVHRR</td>
<td>0.78</td>
<td>0.73</td>
<td>0.67</td>
<td>0.63</td>
<td>0.58</td>
<td>0.55</td>
<td>0.46</td>
<td>0.44</td>
<td>0.41</td>
<td>0.38</td>
</tr>
</tbody>
</table>
The results indicate that the AVHRR series data had consistently higher Kappa values as compared with those of the TM and the SPOT data, although the total number of classes in the AVHRR land cover was much larger than those in the other two datasets. This was probably due to the large size of the original AVHRR pixel (i.e., 1km), and the fact that each pixel contained more mixed classes.

The heterogeneity of the upscaled land cover datasets at selected resolutions is shown in figure 6.9. About 44 percent of the pixels at 2km resolution were "pure" pixels (assuming the 1km pixels are pure) and another 40 percent were composed of two classes. These results were similar to those observed in the 2x2-pixel aggregations of the TM and SPOT land cover datasets. "Pure" pixels dropped to only about 15 percent when pixel size increased to 4km. At 8km level, about 58 percent of pixels contained five or fewer classes, among which about 32 percent consisted of either four or five classes. The heterogeneity of coarse resolution pixels increased rapidly as resolution further coarsened. At the 32km level, the numbers of land cover types included in coarse pixels ranged from 1 to over 40 and about half of these pixels were composed of fifteen or more land cover types. At the 64km level, most pixels contained more than twenty land cover types.

Degree of dominance was used to further examine the characteristics of the coarse resolution pixels (fig. 6.10). At 2km resolution, about 28 percent of the pixels were represented by a class having 50% or less dominance. This was similar to results from the SPOT dataset, but was better than the TM results, where about 35 percent pixels at 60m were represented by 50% or less dominance. The number of "pure" pixels dropped from 44 percent at 2km level to about 15 percent at 4km and to only 4 percent at 8km.
There were essentially no "pure" pixels after resolution exceeded 8km and most pixels had 35%-50% dominance (i.e., a land cover type constituting only 35%-50% of the area of a coarse pixel).

6.3.4. AVHRR-derived 25-class land cover in the conterminous U.S.

Figure 6.11 portrays the consistency rate and Kappa statistics for a 25-class dataset for the conterminous U.S. These classes were grouped from the original 159-class dataset (see Table 3.8). The patterns of the two curves largely followed those of the original 159-class data except for the higher consistency rate and Kappa values. The results were expected because errors introduced by spatial aggregation are related to the definition of classes in the full resolution data. On the other hand, the results indicated that increase in the Kappa value (as compared with the 159-class data) was not a constant across the spatial resolutions. The differences in the Kappa values between the 25-class and 159 class datasets were larger at coarser resolutions (e.g., Δ=0.18 at 64km) than at finer resolutions (e.g., Δ=0.07 at 2km). This was because the number of candidate classes available for selection at a coarser resolution was much larger than at a finer resolution for the original 159-class data (e.g., a 12km resolution pixel might contain 144 different land cover classes while a 4km pixel can contain only 16 classes at most). For the 25-class dataset, the maximum possible number of classes would not exceed 25, the total number of land cover classes, no matter how coarse a pixel is. Such comparisons, however, must be based on data for the same initial resolution (e.g., 1km). When land
cover data derived from different original spatial resolutions were upscaled, higher Kappa values could result for the data with more total classes (see the comparison between the TM classification and the 159-class AVHRR classification).

The pixel heterogeneity in the upscaled land cover datasets at selected resolutions is shown in figure 6.12. About 63 percent of pixels at the 2km resolution were "pure" pixels and 32 percent were composed of two land cover types. As the pixel resolution coarsened to 4km, "pure" pixels constituted 34 percent and the number of two-class mixed-pixels was about 38 percent. Most coarse pixels contained fewer than five land cover types when the pixel resolution was finer than 32km. Even at the 64km resolution, about half of the coarse pixels consisted of fewer than eight cover types. These results were much better than those obtained in the upscaling of the 159-classes data, where a coarse pixel typically contained 15 to 30 cover types at 64km resolution. The degree of dominance (fig. 6.13) indicated that the majority class generally gave good representation of the original land cover in terms of proportional area. For resolutions up to 16km, more than 80 percent of the coarse resolution pixels were represented by majority classes that had proportional areas larger than 50%. Even at the 64km pixel size, about 38% of the pixels were represented by truly majority classes, i.e., classes constituting more than half of the pixel areas. On the other hand, it should be pointed out that, although the 25-class database reflected substantial simplification of the original 159-class data, improvements in the degree of dominance were less than two-fold even at the 64km resolution, at which a coarse resolution pixel contained 4096 full resolution pixels and the total number of classes could have substantial impact on the degree of dominance. The results suggest
that a simplification in classification scheme in land cover mapping may not necessarily result in a similar degree of reduction in proportional areal errors when the land cover dataset is upscaled.

6.3.5. Stratification by ecoregion

Spatial upscaling was conducted using both the 159-class and the grouped 25-class AVHRR land cover databases in four of the Omernik level I ecoregions. Overall consistency rate, Kappa statistics, heterogeneity and degree of dominance were calculated for both land cover datasets in each ecoregion. Figures 6.14 and 6.15 depict the consistency rate and Kappa values for the 159-class and 25-class land cover products, respectively. Pixel heterogeneity and degree of dominance are shown in figures 6.16 to 6.19 for the 159-class database and in figures 6.20 to 6.23 for the 25-class database. Note that here the percentage pixels in the heterogeneity graphs (figs. 6.16, 6.18, 6.20, and 6.22) are indicated with cumulative values to better compare regions. Results from two of the regions, #4 and #6, are shown for illustration purposes.

Although the overall patterns of the curves seemed similar, differences were observed among the ecoregions. For both the 25-class and the 159-class databases, results obtained from region #4 displayed the highest Kappa values while region #6 exhibited the lowest after the pixel resolution exceeded 8km. The relative magnitudes in Kappa values for regions #5 and #9 were also consistent between the two 159- and 25-class datasets. Since upscaling was conducted for only one area within each of the four regions (i.e., the
whole area of each ecoregion), no statistical method could be used to test if the
differences among the four ecoregions were significant. Qualitative methods were utilized
to examine if the observed differences were just randomly introduced. According to the
method of Monserud and Leemans (1992) (Table 6.1), when the pixel size exceeded 4km,
Kappa values from the four regions fell into at least two categories for both the 159-class
and the 25-class datasets. For example, for the 25-class land cover, the value from region
#6 at 4km resolution, 0.69, fell into the "good" range while values from the other three
regions were ranked "very good". At the 64km level, region #4 still displayed "good"
agreement and region #9 was "fair", while regions #5 and #6 exhibited "poor" agreement.
The 25-class land cover data resulted in consistently better agreement between the
upscaled and the full resolution datasets than the 159-class dataset. Because the 25-class
database was a significant generalization of the 159-class database in terms of both class
numbers and class definitions, the differences in Kappa values obtained from the two
databases should represent, to a certain extent, the consequential contrast in the upscaling
of the databases. The magnitudes of regional differences among the Kappa values were
therefore compared to those between the two datasets to examine if they were of the same
order. Tables 6.3 and 6.4 list the average differences in Kappa values found in the two
datasets and those observed among the four ecoregions.

Table 6.3 Differences in Kappa values between the 25-class and the 159-class land
cover, averaged from the four ecoregions in the conterminous United States.

<table>
<thead>
<tr>
<th></th>
<th>2km</th>
<th>4km</th>
<th>8km</th>
<th>16km</th>
<th>32km</th>
<th>64km</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>0.06</td>
<td>0.09</td>
<td>0.11</td>
<td>0.12</td>
<td>0.13</td>
<td>0.14</td>
</tr>
</tbody>
</table>
Table 6.4 Differences in Kappa values among the four ecoregions, averaged from pairwise comparisons.

<table>
<thead>
<tr>
<th></th>
<th>2km</th>
<th>4km</th>
<th>8km</th>
<th>16km</th>
<th>32km</th>
<th>64km</th>
</tr>
</thead>
<tbody>
<tr>
<td>159-class data</td>
<td>0.05</td>
<td>0.07</td>
<td>0.06</td>
<td>0.08</td>
<td>0.09</td>
<td>0.09</td>
</tr>
<tr>
<td>25-class data</td>
<td>0.03</td>
<td>0.05</td>
<td>0.06</td>
<td>0.08</td>
<td>0.12</td>
<td>0.15</td>
</tr>
</tbody>
</table>

The data in the two Tables indicated that the magnitudes of the differences in Kappa values observed among different ecoregions were similar to those that resulted from the 25- and 159-class datasets, especially at resolutions coarser than 16km. The results, although not statistically tested, seemed to suggest that the effects of regional differences on upscaling were significant.

Because fewer cover types were likely to result in higher agreement, other conditions being equal, the number of land cover types in each of the four regions was examined to see if the different Kappa values were caused by differences in class numbers. For the 159-class database, the two regions with higher Kappa values, regions #4 and #9, had fewer classes (122 and 105, respectively) than the other two regions, #5 and #6, which had 137 and 135 cover types. However, for the 25-class database, all four regions had almost the same numbers of cover types, 17, 16, 15 and 16 for regions #4, #5, #6, and #9 (excluding very small classes with less than 0.1% of the areas of corresponding regions). Since there were larger differences in Kappa values among regions in the 25-class data than in the 159-class data, the number of classes seemed not to be a factor in determining the agreement level.
The regional differences observed were likely caused by differences in spatial characteristics of the landscapes among the regions. Region #6 (Northwestern mostly coniferous forested mountains) which displayed the lowest Kappa values, includes steep, rugged mountains as high as 4,300m. The land surface is more fragmented than in other regions. Although there was a single large class in the 25-class AVHRR dataset (i.e., western coniferous forest), the other cover types in this region were small and scattered. By contrast, region #4, the Great Plains, which showed the highest Kappa values, is characterized by rolling plains and tablelands of moderate relief. The land surface is very flat, only occasionally dissected by valleys and canyons. The relatively homogeneous landscape in this area apparently contributed to the higher degree of agreement between the original and the upscaled land cover datasets. The complexity of the landscape is also reflected in the heterogeneity graphs (figs. 6.16, 6.18, 6.20, and 6.22). For example, the curves obtained for region #6 were located toward the higher values on the X-axis (the number of classes) as compared with curves from region #4. This indicates that the average number of classes contained in coarse pixels were greater in region #6 than in region #4.

6.3.6. Comparison between the Kappa values and image information content

Figures 6.24 and 6.25 show the relationship between NDVI STD and spatial resolution for the four level I ecoregions and for the Lincoln/Omaha and the Niobrara study areas. Percentages of STD decrease across resolutions, defined as the ratio of the
absolute difference between the coarse and full resolution STDs to the full resolution
STD, multiplied by 100, were plotted. The percentage decrease exhibited more clearly
the change in STD relative to the original STD magnitude. The observed patterns of STD
change across resolutions were very similar to the change in Kappa values (see figs. 6.2,
6.5, 6.8, and 6.11). As resolution became coarser, NDVI STD decreased logarithmically.
However, the relative magnitudes of the decrease in NDVI STD values for the four
regions were contrary to the relative magnitudes of the Kappa values after the pixel size
reached 32km. The larger the Kappa value, the smaller the decrease of NDVI STD (see
figures 6.8, 6.11, and 6.24). This was also observed for the relative magnitudes in the
decrease of NDVI STD and the Kappa values obtained in the Lincoln/Omaha and the
Niobrara study area (see figures 6.2, 6.5, and 6.25). These results suggested a possible
inverse relationship between the magnitude of NDVI STD decrease (from fine to coarse
resolutions) and the degree of agreement (between fine and coarse resolution land cover
data). Scatter plots between the NDVI STD and Kappa values were, therefore, used to
explore the possible relationship between the two (figs. 6.26 and 6.27). Each sample
point in the scatter plots represented one Kappa/NDVI STD correspondence at one coarse
resolution level. Linear regressions for the sample points were performed for each area
and for both datasets. Tables 6.5 - 6.7 summarize the results of these regressions.
Table 6.5 Summary of linear regression between Kappa values and NDVI standard deviations for the 159-class AVHRR dataset.

<table>
<thead>
<tr>
<th>Reg_#</th>
<th>Coeff.</th>
<th>Std_err</th>
<th>t</th>
<th>r²</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>#4</td>
<td>-3.47</td>
<td>0.23</td>
<td>-15.02 &lt; p_{0.001}</td>
<td>0.97</td>
<td>299.49 &lt; p_{0.001}</td>
</tr>
<tr>
<td>slope</td>
<td>28.88</td>
<td>1.67</td>
<td>17.31 &lt; p_{0.001}</td>
<td></td>
<td></td>
</tr>
<tr>
<td>#5</td>
<td>-1.12</td>
<td>0.11</td>
<td>-11.14 &lt; p_{0.001}</td>
<td>0.97</td>
<td>249.42 &lt; p_{0.001}</td>
</tr>
<tr>
<td>slope</td>
<td>16.41</td>
<td>1.04</td>
<td>15.79 &lt; p_{0.001}</td>
<td></td>
<td></td>
</tr>
<tr>
<td>#6</td>
<td>-0.93</td>
<td>0.09</td>
<td>-10.28 &lt; p_{0.001}</td>
<td>0.96</td>
<td>216.19 &lt; p_{0.001}</td>
</tr>
<tr>
<td>slope</td>
<td>14.26</td>
<td>0.97</td>
<td>14.70 &lt; p_{0.001}</td>
<td></td>
<td></td>
</tr>
<tr>
<td>#9</td>
<td>-1.07</td>
<td>0.08</td>
<td>-13.48 &lt; p_{0.001}</td>
<td>0.98</td>
<td>386.02 &lt; p_{0.001}</td>
</tr>
<tr>
<td>slope</td>
<td>15.44</td>
<td>0.79</td>
<td>19.65 &lt; p_{0.001}</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 6.6 Summary of linear regression between Kappa values and NDVI standard deviations for the 25-class AVHRR dataset.

<table>
<thead>
<tr>
<th>Reg_#</th>
<th>Coeff.</th>
<th>Std_err</th>
<th>t</th>
<th>r²</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>#4</td>
<td>-2.10</td>
<td>0.15</td>
<td>-14.45 &lt; p_{0.001}</td>
<td>0.98</td>
<td>361.49 &lt; p_{0.001}</td>
</tr>
<tr>
<td>slope</td>
<td>19.94</td>
<td>1.05</td>
<td>19.01 &lt; p_{0.001}</td>
<td></td>
<td></td>
</tr>
<tr>
<td>#5</td>
<td>-0.99</td>
<td>0.07</td>
<td>-14.47 &lt; p_{0.001}</td>
<td>0.99</td>
<td>518.26 &lt; p_{0.001}</td>
</tr>
<tr>
<td>slope</td>
<td>15.41</td>
<td>0.68</td>
<td>22.77 &lt; p_{0.001}</td>
<td></td>
<td></td>
</tr>
<tr>
<td>#6</td>
<td>-0.86</td>
<td>0.05</td>
<td>-18.24 &lt; p_{0.001}</td>
<td>0.99</td>
<td>826.23 &lt; p_{0.001}</td>
</tr>
<tr>
<td>slope</td>
<td>14.46</td>
<td>0.50</td>
<td>28.74 &lt; p_{0.001}</td>
<td></td>
<td></td>
</tr>
<tr>
<td>#9</td>
<td>-0.739</td>
<td>0.05</td>
<td>-15.62 &lt; p_{0.001}</td>
<td>0.99</td>
<td>806.01 &lt; p_{0.001}</td>
</tr>
<tr>
<td>slope</td>
<td>13.28</td>
<td>0.47</td>
<td>28.39 &lt; p_{0.001}</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 6.7 Summary of linear regression between Kappa values and NDVI standard deviations for the TM dataset in Lincoln/Omaha area and for the SPOT dataset in the Niobrara area.

<table>
<thead>
<tr>
<th></th>
<th>Coeff.</th>
<th>Std_err</th>
<th>$t$</th>
<th>$r^2$</th>
<th>$F$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lincoln</td>
<td>const. -0.49</td>
<td>0.04</td>
<td>-11.47 $&lt; p_{0.001}$</td>
<td>0.98</td>
<td>461.15 $&lt; p_{0.001}$</td>
</tr>
<tr>
<td></td>
<td>slope</td>
<td>8.01</td>
<td>0.37</td>
<td>21.47 $&lt; p_{0.01}$</td>
<td></td>
</tr>
<tr>
<td>Niobrara</td>
<td>const. -0.43</td>
<td>0.07</td>
<td>-6.09 $&lt; p_{0.001}$</td>
<td>0.94</td>
<td>361.49 $&lt; p_{0.001}$</td>
</tr>
<tr>
<td>(June)</td>
<td>slope</td>
<td>13.55</td>
<td>1.05</td>
<td>12.87 $&lt; p_{0.001}$</td>
<td></td>
</tr>
<tr>
<td>Niobrara</td>
<td>const. -0.98</td>
<td>0.17</td>
<td>-5.61 $&lt; p_{0.001}$</td>
<td>0.92</td>
<td></td>
</tr>
<tr>
<td>(August)</td>
<td>slope</td>
<td>16.38</td>
<td>1.61</td>
<td>10.14 $&lt; p_{0.001}$</td>
<td></td>
</tr>
</tbody>
</table>

The $r^2$, $t$ and $F$ values indicated that the degree of agreement (as measured by the Kappa statistics) between the full and the coarser resolution land cover datasets could be significantly related to changes in the NDVI standard deviation across spatial resolutions, regardless of the differences in the study areas and in the land cover classification scheme. The results, however, should not be interpreted to mean there is a direct relationship between NDVI standard deviation and classification accuracy because the mapping of land cover involved not only the clustering of images but also a great deal of ancillary information. Nevertheless, the comparison between the Kappa values and the NDVI standard deviation indicated that there was a significant association between land cover representation and image information content across spatial resolutions.

6.4. Summary and conclusions

In this study, the characteristics of the majority rule based spatial aggregation, the most widely used upscaling method, were examined and factors influencing the
performance of this aggregation technique were assessed. Spatial aggregations were conducted for land cover databases developed from images acquired by sensors with different spatial resolutions. The aggregations were performed in different ecological regions and on land cover classes with different degrees of generalization.

Kappa statistics indicated good or very good overall agreement between full and aggregated land cover databases when upscaling was performed to pixel sizes four times as large as the original resolutions. As resolution increased to eight times the original resolution, although most coarse resolution land cover products showed good agreement with the corresponding full resolution datasets, some gave only fair representations (i.e., $0.40 < \text{Kappa} < 0.55$). When aggregations were carried out to a resolution level sixteen times the original pixel size, the 25-class AVHRR dataset still exhibited good representation of the original data because of its general definition of the land cover types while the TM dataset in the Lincoln/Omaha study area displayed poor agreement, likely due to its detailed rural area classification and complex cover type interspersion in the urban areas. As aggregation proceeded to even coarser resolutions, only the 25-class AVHRR land cover data showed fair to good agreement with the original 1km data. At the 64km resolution, which was 64 times the original resolution, even the 25-class general land cover classes were poorly represented by the upscaling in some ecoregions.

These results suggest that it may be problematic to compare land cover databases derived from data with very different resolutions (e.g., TM and SPOT vs. AVHRR). Such observations have significant implications for global land cover mapping, especially for
the validation of global datasets using finer resolution data. This is in agreement with Merchant et al. (1993) who suggested re-examination of conventional techniques for accuracy assessment of continental scale land cover databases. The poor representations of AVHRR land cover at 64km suggest that additional assessment is needed when satellite-derived data are to be used in regional/global ecoclimatological models such as GCMs. The findings are also in agreement with those observed by Oleson et al. (1996) who found significant differences in land cover proportions upscaled from two satellite-derived land cover datasets and one standard map set to two GCM cells.

The degree of agreement between the full and the coarse resolution land cover datasets was dependent on the definition of land cover classes, the degree of mixture in the original pixel, and landscape characteristics. The grouped 25-class AVHRR-derived land cover dataset exhibited the highest Kappa values due to its general definition of cover types. Compared to the TM- and SPOT-derived land covers, the 159-class AVHRR-derived land cover, though having much larger number of classes, displayed consistently a higher degree of agreement, given the same level of aggregation (i.e., ratio of coarse to fine resolutions). This was because most 1km AVHRR pixels themselves were mixed-pixels. The TM-derived land cover, although it had a larger original pixel size, exhibited lower Kappa values than those of the SPOT-derived datasets due to its more specific class definitions and complicated interspersion of cover types in the study area. Regional differences were also observed in the upscaling. Poor representation was more common in regions with complex land-surface forms because landscapes in such areas tend to be more heterogeneous.
The majority (also called dominant) rule method is the most widely used procedure for aggregation upscaling of land cover because it selects a land cover class having the largest proportional area to represent a coarse resolution pixel and thus introduces the lowest areal error. The term "majority" (or "dominant"), however, is misleading because, although the selected majority class in a coarse resolution pixel has a larger proportional area than any other class in the same pixel, it may not necessarily comprise a majority portion of the pixel. The measurement of "degree of dominance" proposed in this study provides a useful assessment of the representation of a majority class for a coarse, mixed-pixel. It portrays the real proportional area of a "majority" class in a coarse pixel and thus gives an accurate estimation of areal errors for upscaled land cover databases.

Kappa values, representing the degree of agreement between coarse and fine resolution land cover databases, could be linearly regressed against NDVI standard deviation with significant $r^2$ for all the study areas. Although the accuracy of land cover classification is not likely to be associated with the NDVI standard deviation using a simple linear relationship, the results suggested that there was a significant association between land cover representation and image information content across spatial resolutions.

This study investigated the characteristics of coarse resolution land cover databases upscaled using the spatial aggregation method. It provided a better understanding on the upscaling representation of land cover datasets. The assessment of spatially upscaled land cover databases should be based on different measures (i.e., overall agreement, pixel heterogeneity, degree of dominance) and under different conditions. The conditions
include the definition of full resolution land cover types, the dimensions of the full and
the coarse resolution pixels, and structures of the landscapes under investigation. The
study, however, concentrated on the overall comparison between land cover representation
at full and coarse resolutions, and only one specific upscaling method, i.e., majority rule
aggregation, was examined. The patterns of areal changes in individual classes across
resolutions and characteristics of different upscaling approaches are examined in a
separate study presented in the next chapter.
6.5. References


Merchant, J. W., Yang, L., and Yang, W., 1993, Validation of continental-scale land cover data bases developed from AVHRR data, in *Proceeding of the Pecora XII Symposium*, pp. 63-72, Sioux, Falls, South Dakota.


The error matrix

Fine resolution (20m) subimage

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>3</th>
<th>3</th>
<th>3</th>
<th>2</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>2</td>
<td>4</td>
<td>4</td>
<td>5</td>
<td>6</td>
<td>3</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>3</td>
<td>4</td>
<td>4</td>
<td>6</td>
<td>6</td>
<td>4</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>4</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>5</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>6</td>
<td>5</td>
<td>5</td>
<td>2</td>
<td>2</td>
<td>3</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>7</td>
<td>5</td>
<td>5</td>
<td>2</td>
<td>2</td>
<td>4</td>
<td>4</td>
<td>4</td>
</tr>
</tbody>
</table>

Numbers 1 to 6 represent six different land cover classes. Total number of pixels = 64

Consistency rate between the fine and the coarse resolution subimage:

\[ \text{sum along diagonal in the error matrix} = 25 \]
\[ \text{total number of pixels in the subimage} = 64 \]
\[ \text{Consistency rate} = \frac{25}{64} \times 100\% = 39\% \]

Kappa statistics:

\[ k = \frac{\left( p_o - p_e \right)}{1 - p_e} = \frac{N \sum x_{ii} + \sum x_i \cdot x_j - x_{ii}}{N^2 - \sum x_i \cdot x_j} \]
\[ = \frac{348}{334} = 0.25 \]

where \( x_{ii} \) is element (i,i) in the error matrix, \( x_i \), the sum of row i, \( x_j \) the sum of column j, and \( N \) the total number of pixels in the fine resolution subimage.

Fine resolution (20m) subimage

<table>
<thead>
<tr>
<th></th>
<th>4</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Coarse resolution (80m) subimage

<table>
<thead>
<tr>
<th></th>
<th>4</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>5</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>3</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>4</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>5</td>
<td>5</td>
<td>0</td>
</tr>
<tr>
<td>6</td>
<td>6</td>
<td>0</td>
</tr>
</tbody>
</table>

Numbers 1 to 6 represent the land cover types in the fine resolution subimage while the first row (b) represents those in the coarse resolution subimage.

Consistency rate between the fine and the coarse resolution subimage:

\[ \text{sum along diagonal in the error matrix} = 100\% = 39\% \]

Kappa statistics:

\[ k = \frac{\left( p_o - p_e \right)}{1 - p_e} = \frac{N \sum x_{ii} + \sum x_i \cdot x_j - x_{ii}}{N^2 - \sum x_i \cdot x_j} \]
\[ = \frac{348}{334} = 0.25 \]

where \( x_{ii} \) is element (i,i) in the error matrix, \( x_i \), the sum of row i, \( x_j \) the sum of column j, and \( N \) the total number of pixels in the fine resolution subimage.

Consistency rate between the fine and the coarse resolution subimage:

\[ \text{sum along diagonal in the error matrix} = 25 \]
\[ \text{total number of pixels in the subimage} = 64 \]
\[ \text{Consistency rate} = \frac{25}{64} \times 100\% = 39\% \]

Kappa statistics:

\[ k = \frac{\left( p_o - p_e \right)}{1 - p_e} = \frac{N \sum x_{ii} + \sum x_i \cdot x_j - x_{ii}}{N^2 - \sum x_i \cdot x_j} \]
\[ = \frac{348}{334} = 0.25 \]

where \( x_{ii} \) is element (i,i) in the error matrix, \( x_i \), the sum of row i, \( x_j \) the sum of column j, and \( N \) the total number of pixels in the fine resolution subimage.

Consistency rate between the fine and the coarse resolution subimage:

\[ \text{sum along diagonal in the error matrix} = 100\% = 39\% \]

Kappa statistics:

\[ k = \frac{\left( p_o - p_e \right)}{1 - p_e} = \frac{N \sum x_{ii} + \sum x_i \cdot x_j - x_{ii}}{N^2 - \sum x_i \cdot x_j} \]
\[ = \frac{348}{334} = 0.25 \]

where \( x_{ii} \) is element (i,i) in the error matrix, \( x_i \), the sum of row i, \( x_j \) the sum of column j, and \( N \) the total number of pixels in the fine resolution subimage.

Consistency rate between the fine and the coarse resolution subimage:

\[ \text{sum along diagonal in the error matrix} = 100\% = 39\% \]

Kappa statistics:

\[ k = \frac{\left( p_o - p_e \right)}{1 - p_e} = \frac{N \sum x_{ii} + \sum x_i \cdot x_j - x_{ii}}{N^2 - \sum x_i \cdot x_j} \]
\[ = \frac{348}{334} = 0.25 \]

where \( x_{ii} \) is element (i,i) in the error matrix, \( x_i \), the sum of row i, \( x_j \) the sum of column j, and \( N \) the total number of pixels in the fine resolution subimage.

Consistency rate between the fine and the coarse resolution subimage:

\[ \text{sum along diagonal in the error matrix} = 100\% = 39\% \]

Kappa statistics:

\[ k = \frac{\left( p_o - p_e \right)}{1 - p_e} = \frac{N \sum x_{ii} + \sum x_i \cdot x_j - x_{ii}}{N^2 - \sum x_i \cdot x_j} \]
\[ = \frac{348}{334} = 0.25 \]

where \( x_{ii} \) is element (i,i) in the error matrix, \( x_i \), the sum of row i, \( x_j \) the sum of column j, and \( N \) the total number of pixels in the fine resolution subimage.

Consistency rate between the fine and the coarse resolution subimage:

\[ \text{sum along diagonal in the error matrix} = 100\% = 39\% \]

Kappa statistics:

\[ k = \frac{\left( p_o - p_e \right)}{1 - p_e} = \frac{N \sum x_{ii} + \sum x_i \cdot x_j - x_{ii}}{N^2 - \sum x_i \cdot x_j} \]
\[ = \frac{348}{334} = 0.25 \]

where \( x_{ii} \) is element (i,i) in the error matrix, \( x_i \), the sum of row i, \( x_j \) the sum of column j, and \( N \) the total number of pixels in the fine resolution subimage.
Figure 6.2 Changes in consistency rate (a) and Kappa statistics (b) across different spatial resolutions, results of the SPOT derived land cover data in the Niobrara study area.
Figure 6.3 Pixel heterogeneity at selected spatial resolutions. The curves indicate average number of land cover classes in each coarse resolution pixel. Results of the SPOT derived land cover data in the Niobrara study area.
Figure 6.4 Degree of dominance of the majority class, results of the SPOT derived land cover in the Niobrara study area.

A point at x=28, y=50 indicates that 28 percent of the coarse pixels having 50% or less dominance.
Figure 6.5 Changes in consistency rate (a) and Kappa statistics (b) across different spatial resolutions, results of the TM derived land cover data in the Lincoln/Omaha study area.
Figure 6.6 Pixel heterogeneity at selected spatial resolutions. The curves indicate average number of land cover classes in each coarse resolution pixel. Results of the TM derived land cover data in the Lincoln/Omaha study area.
Figure 6.7 Degree of dominance of the majority class, results of the TM derived land cover in the Lincoln/Omaha study area.
A point at x=59, y=99 indicates that 59 percent of the coarse pixels having 99% or less dominance.
Figure 6.8 Changes in consistency rate (a) and Kappa statistics (b) across different spatial resolutions, results of the 159-class AVHRR derived land cover in the conterminous U.S.
Figure 6.9 Pixel heterogeneity at selected spatial resolutions. The curves indicate average number of land cover classes in each coarse resolution pixel. Results of the 159-class AVHRR derived land cover in the conterminous U.S.
Figure 6.10 Degree of dominance of the majority class, results of the 159-class AVHRR derived land cover in the conterminous U.S.
A point at x=85, y=99 indicates that 85 percent of the coarse pixels having 99% or less dominance.
Figure 6.11 Changes in consistency rate (a) and Kappa statistics (b) across different resolutions, results of the 25-class AVHRR derived land cover in the conterminous U.S.
Figure 6.12 Pixel heterogeneity at selected spatial resolutions. The curves indicate average number of land cover classes in each coarse resolution pixel. Results of the 25-class AVHRR derived land cover in the conterminous U.S.
Figure 6.13 Degree of dominance of the majority class, results of the 25-class AVHRR derived land cover in the conterminous U.S.

A point at x=20, y=50 indicates that 20 percent of the coarse pixels having 50% or less dominance.
Figure 6.14 Changes in consistency rate (a) and Kappa statistics (b) across different resolutions in the four level I ecoregions, results of the 159-class AVHRR derived land cover in the conterminous U.S.
Figure 6.15 Changes in consistency rate (a) and Kappa statistics (b) across different resolutions in the four level I ecoregions, results of the 25-class AVHRR derived land cover in the conterminous U.S.
Figure 6.16 Heterogeneity of coarse pixels at selected resolutions. Curves portray the cumulative percentage of pixels having different number of land cover types in each pixel. Results of the 159-class AVHRR derived land cover in ecoregion #4.

The point at x=1, y=46 (the 2km curve) indicates that 46% of pixels contain one land cover type and the point at x=2, y=86 indicates that 86% of pixels consist of either one or two land cover classes.
Figure 6.17  Degree of dominance of the majority classes. Results of the 159-class AVHRR derived land cover in ecoregion #4.
A point at x=60, y=73 indicates that 60 percent of pixels having 73% or less degree of dominance.
Figure 6.18 Heterogeneity of coarse pixels at selected resolutions. Curves portray the cumulative percentage of pixels having different number of land cover types in each pixel. Results of the 159-class AVHRR derived land cover in ecoregion #6. The point at x=1, y=46 (the 2km curve) indicates that 46% of pixels contain one land cover type and the point at x=2, y=86 indicates that 86% of pixels consist of either one or two land cover classes.
Figure 6.19 Degree of dominance of the majority classes. Results of the 159-class AVHRR derived land cover in ecoregion #6.
A point at x=60, y=73 indicates that 60 percent of pixels having 73% or less degree of dominance.
Figure 6.20 Heterogeneity of coarse pixels at selected resolutions. Curves portray the cumulative percentage of pixels having different number of land cover types in each pixel. Results of the 25-class AVHRR derived land cover in ecoregion #4. The point at x=1, y=64 (the 2km curve) indicates that 64% of pixels contain one land cover type and the point at x=2, y=95 indicates that 95% of pixels consist of either one or two land cover classes.
Figure 6.21 Degree of dominance of the majority classes. Results of the 25-class AVHRR derived land cover in ecoregion #4.
A point at x=53, y=92 indicates that 53 percent of pixels having 92% or less degree of dominance.
Figure 6.22 Heterogeneity of coarse pixels at selected resolutions. Curves portray the cumulative percentage of pixels having different number of land cover types in each pixel. Results of the 25-class AVHRR derived land cover in ecoregion #6.

The point at $x=1, y=65$ (the 2km curve) indicates that 65% of pixels contain one land cover type and the point at $x=2, y=94$ indicates that 94% of pixels consist of either one or two land cover classes.
Figure 6.23 Degree of dominance of the majority classes. Results of the 25-class AVHRR derived land cover in ecoregion #6.
A point at x=34, y=99 indicates that 34 percent of pixels having 99% or less degree of dominance.
Figure 6.24 Change in standard deviation of AVHRR MVC NDVI with spatial resolutions for the four Level I ecoregions. Results of the 8/17-8/30, 1990 composite period. (a) NDVI standard deviation. (b) percent decrease in NDVI standard deviation. calculated as: abs(STD(i)-STD(1))/STD(1), where STD(i) is the NDVI standard deviation at resolution i km.
Figure 6.25 Change in NDVI standard deviation with spatial resolutions. (a) Results of the Lincoln/Omaha study area; (b) results of the Niobrara study area.

- Percent decrease in NDVI standard deviation was calculated as: \( \text{abs}(\text{STD}(i) - \text{STD}(1))/\text{STD}(1) \), where \( \text{STD}(i) \) is the NDVI standard deviation at resolution \( i \) km.
Figure 6.26 Relationships between Kappa statistics and NDVI standard deviation of in the four Level I ecoregions: (a) region 4, (b) region 5, (c) region 6, and (d) region 9.

- For Kappa values obtained from the 159-class land cover
- For Kappa values obtained from the grouped 25-class land cover.
Figure 6.27 Relationships between Kappa statistic and NDVI standard deviation. Results of the Lincoln/Omaha study area and the Niobrara study area.
CHAPTER 7
SPATIAL UPSCALING AND LAND COVER AREA ESTIMATION: LINKING LAND COVER TO LANDSCAPE STRUCTURE

7.1. Introduction

The study presented in chapter 6 demonstrated that spatial upscaling of land cover data derived from remote sensing typically contain mixed-pixels. Significant areal errors may be introduced when such mixed-pixels are assigned a single class type by using the majority rule. The degree of land cover mixture, and the areal error, in a coarse resolution pixel is dependent on complex interrelationships among a number of factors including the difference between the full and coarse pixel sizes, landscape heterogeneity, and changes in image information content. Further research is required to examine how the scaling-induced errors are associated with individual land cover types and to what extent such errors can be estimated and corrected.

Recent studies have demonstrated that relationships between the areal errors in spatially upscaled land cover data and landscape structure metrics could be established for specific study areas using regression-based methods (Moody and Woodcock, 1995; Moody, 1996). The operational use of the regression models, however, requires that they be extensible across landscapes because landscape structure metrics, measured at high resolution, are usually not available for areas where such corrections are needed. Mayaux and Lambin (1995) suggested a two-step regression procedure: one between the areal error
and fine resolution landscape metrics and the other between fine and coarse resolution landscape metrics. They showed that if significant relationship was found between the landscape metrics at fine and coarse resolutions, the areal errors could be estimated directly from coarse resolution land cover data. Turner et al. (1989), however, suggested that although simple relationships between landscape parameters measured at different scales might be identified, the exact relationship will vary across different landscapes and may not permit extrapolation from one region to another. The results of chapter 5 demonstrated that there are significant regressional relationships between some landscape metrics and spatial resolutions in neutral models, but the regression coefficients changed with different landscapes. Clearly, additional research on such relationships for real landscapes is needed to determine if they are generalizable for different datasets and over different ecological regions.

Coarse resolution land cover data can be developed using different methods. In many applications, land cover databases having pixel sizes between 1 to 16km were often classified directly from coarse resolution data (e.g., AVHRR measurements and their degraded products), rather than aggregated from finer resolution land cover datasets (e.g., Malingreau, 1986; Townshend et al., 1987; James and Kalluri, 1994; Loveland et al., 1995). When land cover data are used in global modeling, spatial aggregation is a common method (e.g., Oleson et al., 1996; DeFries et al., 1997). It is important, therefore, to investigate the relationships between the spatially aggregated land cover databases and those derived directly from coarse resolution measurements or from data degraded from full resolution measurements.
The principal objective of this study were: 1) to determine the relationships between landscape structure and areal errors in upscaled land cover data and their extensibility over different datasets and over ecological regions; 2) to examine relationships between landscape structure indices and spatial resolution for different land cover datasets; and 3) to investigate the effects of different upscaling techniques on the land cover representation and on the relationships between land cover area and landscape structure.

7.2. Data and methods

The major datasets used in this study included four land cover databases: 1) SPOT-derived land cover data for the Niobrara Valley Preserve and vicinity in northern Nebraska; 2) TM-derived land cover data in the area of Lincoln/Omaha; 3) AVHRR-derived land cover data for the conterminous U.S. (159 land cover classes); and 4) a 25-class AVHRR land cover dataset generalized (i.e., grouped) from the 159-class AVHRR data. For the Lincoln/Omaha study area, the original six-band TM image with 30m resolution was also used. These datasets were used to create a series of multiresolution land cover databases. The complete procedure for the creation of the multiresolution databases are described in Chapter 3. Specifically, the 20m SPOT-derived land cover data were upscaled to eleven coarse resolutions and the 30m TM classification was upscaled to ten coarse resolutions, using a spatial aggregation method based on a majority rule. The coarse resolutions were 60, 120, 240, 360, 480, 600, 720, 840, and 960 meters. The
SPOT series also included 40 and 80 meters, while the TM series included 90 meters. Both the 159-class and 25-class AVHRR land cover were upcaled to ten coarse resolutions: 2, 4, 8, 16, 24, 32, 40, 48, 56, and 64 kilometers.

Because coarse resolution land cover databases between 1 to 16 kilometers are often derived from either AVHRR measurements or their degraded products, a "re-classification" method was used to create another series of coarse resolution land cover datasets in the Lincoln/Omaha study area. The original six-band TM data were degraded, using a spatial averaging method, into "coarse resolution TM images" and these coarse resolution TM images were classified and labeled, using the same spectral signatures and reference data, into coarse resolution land cover datasets. The aggregated and re-classified land cover data were used to examine the effects of different upscaling approaches on land cover representation and to investigate if the two approaches would result in different relationships between land cover and landscape structure.

Landscape indices were calculated for individual classes as well as for the entire landscape using Fragstats, a spatial analysis software package (McGarigal and Marks, 1993). Fragstats computes forty class level indices and forty-six landscape level indices. Due to the large computation load of Fragstats (e.g., 15 to 20 hours, depending on specific machines, on a SUN SPARC 20 for the 25-class conterminous U.S. land cover), the landscape indices were not calculated for all the coarse resolution AVHRR-derived databases. The resolution levels for which landscape indices were calculated included 1, 2, 4, 8, 16, 32, and 64 kilometers. For the TM-derived and the SPOT-derived land cover databases, the indices were computed for the full and all coarse resolution levels.
A class neighbor matrix, which was not part of the Fragstats indices, was used to assist the interpretation of class area pattern change with resolution. Each element in the class neighbor matrix, \( C_{ij} \), was defined as: 
\[
C_{ij} = 100 \times N_{ij} / N_{i},
\]
where \( N_{ij} \) is the number of two-pixel pairs in which one pixel belongs to cover class \( i \) and the other to class \( j \), and \( N_i \) is the number of neighboring two-pixel pairs in which one pixel belongs to class \( i \) while the other may belong to any class. The class neighbor matrix is not symmetric because \( N_{ij} \) equals to \( N_{ji} \) while \( N_i \) may not necessarily equal to \( N_j \). The matrix reflects the spatial connection among classes. A high value of \( C_{ij} \) indicates that class \( i \) is usually a neighbor of class \( j \).

Changes in individual class areas across different resolutions were first qualitatively examined and their association with the spatial characteristics of corresponding classes were identified. Statistical analyses (regressions) were then performed to explore relationships between land cover areas and landscape parameters. The regressions were conducted between class area and selected class level landscape indices. Areas of land classes are considered as class level indices in Fragstats. In this study, only the metrics describing the spatial characteristics of a class, such as shape and patch size, were used as predictor variables. The selection of indices was primarily based on the correlation among indices, because significant association among predictor variables would cause multicollinearity problems in regression analyses.

Pairwise Pearson-product-moment correlations were conducted among the class level indices for each land cover dataset. As a rule of thumb, variables with \( r^2 > 0.9 \) would cause serious multicollinearity and those with \( r^2 > 0.75 \) also warrant investigation (Glantz
Indices having $r^2>0.75$ were, therefore, excluded. The eight indices that appeared to be least correlated were retained. These were: the largest patch index (LPI), number of patches (NP), mean patch size (MPS), patch size standard deviation (PSSD), mean shape index (MSI), area-weighted mean shape index (AWMSI), double log fractal (DLF), and patch interspersion and juxtaposition (IJI). Descriptions and equations for calculating the eight indices can be found in Appendix A.

Class areas obtained from the full resolution of each land cover dataset (i.e., 20m, 30m, and 1km for the SPOT, TM and AVHRR datasets, respectively) were considered as the base areas ($A_0$) and those from the coarse resolution maps as estimations of corresponding base areas ($A_e$). The proportional error (following the definition by Moody and Woodcock, 1995) was the ratio of the difference between estimated area and base area, to the base area, i.e., $e=(A_e-A_0)/A_0$. Proportional error is, in fact, the normalized difference between estimated and base areas. An over-estimation results in a positive error while an under-estimation causes a negative error. Forward stepwise multi-linear regression was used to explore relationships among estimated and base areas, proportional errors and indices of landscape structure. The dependent variable was either the coarse resolution area ($A_e$) or proportional error ($e$). The predictor variables used were landscape indices with pairwise correlation coefficients, $r^2$, less than 0.75, and the base area ($A_0$) and spatial resolution.

In each regression, the dependent variables, $A_e$ or $e$, were calculated for each class at all resolution levels, while the predictor variables, except for spatial resolution, were all obtained from the original full resolution level. For example, there were 10 coarse
resolution levels and 19 cover types (i.e., 60, 90, ..., 960m) for the TM dataset. Thus, there were 190 (i.e., 19 land cover classes at 10 resolutions) measurements for each dependent variable, but the number of measurements for the predictor variables (except for resolution which was 10) was 19, i.e., one measurement from each cover type. Some small classes were prone to disappear at coarse resolutions for some databases (e.g., both the 159-class and 25-class AVHRR land cover datasets). Therefore, the number of samples used in the regression could be less than the product of class number and resolution level. The data organization of the regressions is summarized in Table 7.1.

Table 7.1. Data organization for regressions of coarse resolution class areas and proportional errors against landscape indices, original areas, and spatial resolutions.

<table>
<thead>
<tr>
<th>Dependent Variables</th>
<th>Predictor Variables</th>
</tr>
</thead>
<tbody>
<tr>
<td>$A_c^{(1,1)}$ or $e^{(1,1)}$</td>
<td>$X_1^{(1,0)}$, $X_2^{(1,0)}$, ..., $X_n^{(1,0)}$, $R_1$, $A_0(1)$</td>
</tr>
<tr>
<td>$A_c^{(1,2)}$ or $e^{(1,2)}$</td>
<td>$X_1^{(2,0)}$, $X_2^{(2,0)}$, ..., $X_n^{(2,0)}$, $R_1$, $A_0(1)$</td>
</tr>
<tr>
<td>$A_c^{(1,C)}$ or $e^{(1,C)}$</td>
<td>$X_1^{(C,0)}$, $X_2^{(C,0)}$, ..., $X_n^{(C,0)}$, $R_1$, $A_0(1)$</td>
</tr>
<tr>
<td>$A_c^{(2,1)}$ or $e^{(2,1)}$</td>
<td>$X_1^{(1,0)}$, $X_2^{(1,0)}$, ..., $X_n^{(1,0)}$, $R_2$, $A_0(2)$</td>
</tr>
<tr>
<td>$A_c^{(2,2)}$ or $e^{(2,2)}$</td>
<td>$X_1^{(2,0)}$, $X_2^{(2,0)}$, ..., $X_n^{(2,0)}$, $R_2$, $A_0(2)$</td>
</tr>
<tr>
<td>$A_c^{(2,C)}$ or $e^{(2,C)}$</td>
<td>$X_1^{(C,0)}$, $X_2^{(C,0)}$, ..., $X_n^{(C,0)}$, $R_2$, $A_0(2)$</td>
</tr>
<tr>
<td>$A_c^{(N,1)}$ or $e^{(N,1)}$</td>
<td>$X_1^{(1,0)}$, $X_2^{(1,0)}$, ..., $X_n^{(1,0)}$, $R_N$, $A_0(C)$</td>
</tr>
<tr>
<td>$A_c^{(N,2)}$ or $e^{(N,2)}$</td>
<td>$X_1^{(2,0)}$, $X_2^{(2,0)}$, ..., $X_n^{(2,0)}$, $R_N$, $A_0(C)$</td>
</tr>
<tr>
<td>$A_c^{(N,C)}$ or $e^{(N,C)}$</td>
<td>$X_1^{(C,0)}$, $X_2^{(C,0)}$, ..., $X_n^{(C,0)}$, $R_N$, $A_0(C)$</td>
</tr>
</tbody>
</table>
where

\( N \) is the number of coarse resolution levels;

\( C \) is the number of land cover types;

\( A_e(i,j) \) is the area of class \( j \) at resolution \( i \);

\( e(i,j) \) is the proportional error for class \( j \) at resolution \( i \);

\( X_k(i,0) \) is landscape indices \( k \) of class \( i \) at original resolution; and

\( A_0(i) \) is the area of class \( i \) at original resolution.

Measurements used in the regressions might not be truly independent because class areas at one resolution level tended to be related to those at another level and all class proportions summed to 1 at any given resolution. Nevertheless, the dataset was judged acceptable because the primary emphasis of the study was focused on identifying the role that landscape structure played in land cover upscaling, and not a quantitative inversion. Such regression methods were also used by other investigators in similar studies in other ecological areas (e.g., Moody and Woodcock, 1995).

All variables were standardized prior to the regression to mitigate potential multicollinearity problems. The forward regressions were performed using SigmaStat, a statistical analysis software package in which the \( F \)-value at the end of each step was checked and variables with \( F \)-values below a threshold (3.9 in all regressions in this study) were removed. Highly correlated variables were, therefore, excluded from the regression models.

The aforementioned regression analyses were used to explore relationships between
landscape indices and the proportional class area estimated from upscaled land cover datasets. Regressions of landscape level indices against spatial resolution were performed to examine relationships between spatial resolution and spatial characteristics of a landscape. Landscape level indices and class level indices were the same if only one land cover type was considered as foreground and everything else as background. The computation procedures for landscape level indices and class level indices were similar (see Appendix A and B). For example, the mean patch size at the class level was computed for patches of the same cover type, while it was computed for all patches (all cover types) at the landscape level. The approaches used in regressing landscape indices against spatial resolutions are similar to those described in Chapter 5. The difference was that in the current study observed data from different sensors and study areas were utilized.

7.3. Results and analyses

7.3.1. Change in class area across resolutions: the Niobrara dataset

Changes in class area across spatial resolutions in the Niobrara study area are shown in Figures 7.1 to 7.3. Different scales for the Y-axis are used and classes with similar areas are plotted together for better comparison. The results indicate that class areas at coarser resolutions were predominantly determined by the initial area at full resolution (20m). Generally, areas of classes having large initial areas (>5%) tended to increase while those having small initial areas (<5%) tended to decrease. However, the
effects of landscape structure, indicating the spatial interspersion among classes, on the pattern of area change was evident. The three largest classes, each possessing more than 15 percentage of the total study area, were highly intermingled (the values of the corresponding row/column in Table 7.2 are large, ≥15, as compared to most of the other non-diagonal values). They competed with each other over the resolutions and changes in area were dependent on their patch characteristics and the spatial relationship among these classes. Class 4, the class having the largest patch size (MPS=1.25 hectare), showed the greatest increase in area. The area of class 10 was stable across all

Table 7.2. Neighboring relationship among different land cover types in the SPOT-derived land cover in Niobrara. The first column and row are cover types. Numbers in the table, \(C_{ij}\), indicate the percentages of one cover type neighboring another.

<table>
<thead>
<tr>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
<th>13</th>
<th>14</th>
<th>15</th>
<th>16</th>
<th>17</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>53</td>
<td>2</td>
<td>2</td>
<td>5</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>23</td>
<td>3</td>
<td>0</td>
<td>6</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>5</td>
<td>21</td>
<td>2</td>
<td>8</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>34</td>
<td>13</td>
<td>0</td>
<td>9</td>
<td>1</td>
<td>4</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>3</td>
<td>3</td>
<td>1</td>
<td>31</td>
<td>31</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>17</td>
<td>7</td>
<td>0</td>
<td>5</td>
<td>1</td>
<td>2</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>4</td>
<td>1</td>
<td>1</td>
<td>5</td>
<td>49</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>16</td>
<td>15</td>
<td>2</td>
<td>2</td>
<td>0</td>
<td>6</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>5</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>11</td>
<td>29</td>
<td>14</td>
<td>0</td>
<td>0</td>
<td>19</td>
<td>13</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>6</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>6</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>7</td>
<td>19</td>
<td>30</td>
<td>0</td>
<td>0</td>
<td>14</td>
<td>7</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>19</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>7</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>91</td>
<td>0</td>
<td>2</td>
<td>3</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>8</td>
<td>8</td>
<td>0</td>
<td>2</td>
<td>5</td>
<td>1</td>
<td>0</td>
<td>2</td>
<td>25</td>
<td>19</td>
<td>16</td>
<td>2</td>
<td>12</td>
<td>1</td>
<td>1</td>
<td>7</td>
<td>0</td>
</tr>
<tr>
<td>9</td>
<td>5</td>
<td>4</td>
<td>3</td>
<td>15</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>38</td>
<td>16</td>
<td>1</td>
<td>6</td>
<td>1</td>
<td>3</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>10</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>17</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>19</td>
<td>42</td>
<td>6</td>
<td>3</td>
<td>0</td>
<td>3</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>11</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>11</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>11</td>
<td>37</td>
<td>33</td>
<td>1</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>12</td>
<td>3</td>
<td>3</td>
<td>2</td>
<td>5</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>16</td>
<td>5</td>
<td>0</td>
<td>52</td>
<td>5</td>
<td>4</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>13</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>1</td>
<td>15</td>
<td>5</td>
<td>0</td>
<td>25</td>
<td>41</td>
<td>5</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>14</td>
<td>0</td>
<td>2</td>
<td>1</td>
<td>22</td>
<td>4</td>
<td>10</td>
<td>0</td>
<td>0</td>
<td>12</td>
<td>10</td>
<td>1</td>
<td>7</td>
<td>2</td>
<td>29</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>15</td>
<td>11</td>
<td>0</td>
<td>1</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>16</td>
<td>12</td>
<td>4</td>
<td>1</td>
<td>11</td>
<td>0</td>
<td>1</td>
<td>32</td>
<td>0</td>
</tr>
<tr>
<td>16</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>3</td>
<td>37</td>
<td>7</td>
<td>0</td>
<td>0</td>
<td>8</td>
<td>4</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>2</td>
<td>1</td>
<td>33</td>
</tr>
<tr>
<td>17</td>
<td>5</td>
<td>1</td>
<td>3</td>
<td>4</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>4</td>
<td>1</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>1</td>
<td>6</td>
<td>3</td>
</tr>
</tbody>
</table>
resolutions. This was due to its small patch size (0.76 hectares), smaller than all classes that have greater than 5 percentage initial area. Thus, it lost more area to, than it gained from, the other two classes. This was reflected in the mixture matrix of classes between coarse and fine resolutions. For example, at 960m resolution 41% of the pixels of this class were classified into the other two classes (row 10, columns 4 and 9 in Table 7.3). It gained only 27% pixels from the same two classes (column 10, rows 4 and 9 in Table 7.4). On the other hand, because it had a larger initial area and patch size when

Table 7.3. Mixture-matrix between the 20m and 960m SPOT-derived land cover in the Niobrara study area. Percentage numbers were based on the 20m resolution classes.

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
<th>13</th>
<th>14</th>
<th>15</th>
<th>16</th>
<th>17</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>54.75</td>
<td>68</td>
<td>8.81</td>
<td>0.03</td>
<td>0.04</td>
<td>3.52</td>
<td>0</td>
<td>20.69</td>
<td>5.29</td>
<td>0.2</td>
<td>5.08</td>
<td>0.31</td>
<td>0.06</td>
<td>0.55</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>13.44</td>
<td>0.54</td>
<td>22.87</td>
<td>0.36</td>
<td>0.57</td>
<td>4.96</td>
<td>0</td>
<td>37.75</td>
<td>7.81</td>
<td>0.26</td>
<td>10.18</td>
<td>0.48</td>
<td>0.82</td>
<td>0.25</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>6.34</td>
<td>6.62</td>
<td>47.41</td>
<td>0.07</td>
<td>0.06</td>
<td>3.76</td>
<td>0</td>
<td>21.33</td>
<td>7.88</td>
<td>0.15</td>
<td>5.76</td>
<td>0.18</td>
<td>0.27</td>
<td>0.16</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>2.94</td>
<td>0.66</td>
<td>59.99</td>
<td>0.32</td>
<td>0.27</td>
<td>2.79</td>
<td>0</td>
<td>17.16</td>
<td>10.78</td>
<td>0.55</td>
<td>3.15</td>
<td>0.13</td>
<td>1.16</td>
<td>0.1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>1.76</td>
<td>0.17</td>
<td>33.12</td>
<td>6.32</td>
<td>4.55</td>
<td>3.4</td>
<td>0</td>
<td>32.18</td>
<td>11.43</td>
<td>0.77</td>
<td>3.82</td>
<td>0.18</td>
<td>2.27</td>
<td>0.02</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>0.79</td>
<td>0.11</td>
<td>26.03</td>
<td>3.59</td>
<td>12.29</td>
<td>2.3</td>
<td>0</td>
<td>30.38</td>
<td>10.96</td>
<td>0.66</td>
<td>3.75</td>
<td>0.15</td>
<td>8.98</td>
<td>0.03</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>0.32</td>
<td>0.02</td>
<td>2.21</td>
<td>0</td>
<td>84.25</td>
<td>0</td>
<td>3.26</td>
<td>6.78</td>
<td>0.12</td>
<td>3.61</td>
<td>0.3</td>
<td>0.09</td>
<td>0.03</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>5.34</td>
<td>0.16</td>
<td>8.46</td>
<td>0.52</td>
<td>0.33</td>
<td>10.57</td>
<td>0.01</td>
<td>35.28</td>
<td>21.65</td>
<td>0.54</td>
<td>16.27</td>
<td>0.37</td>
<td>0.21</td>
<td>0.31</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>7.44</td>
<td>0.56</td>
<td>22.38</td>
<td>0.55</td>
<td>0.82</td>
<td>4.18</td>
<td>0</td>
<td>44.76</td>
<td>10.54</td>
<td>0.23</td>
<td>7.33</td>
<td>0.42</td>
<td>0.69</td>
<td>0.12</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>2.76</td>
<td>0.24</td>
<td>20.95</td>
<td>0.4</td>
<td>0.28</td>
<td>6.77</td>
<td>0</td>
<td>20.49</td>
<td>40.31</td>
<td>0.96</td>
<td>5.78</td>
<td>0.30</td>
<td>0.69</td>
<td>0.05</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>1.73</td>
<td>0.06</td>
<td>16.25</td>
<td>0.3</td>
<td>0.11</td>
<td>8.91</td>
<td>0</td>
<td>11.97</td>
<td>42.78</td>
<td>11.23</td>
<td>5.74</td>
<td>0.4</td>
<td>0.49</td>
<td>0.04</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>1.45</td>
<td>0.01</td>
<td>5.96</td>
<td>0.15</td>
<td>0.18</td>
<td>5.22</td>
<td>0</td>
<td>23.34</td>
<td>8.74</td>
<td>0.39</td>
<td>5.36</td>
<td>0.72</td>
<td>0.19</td>
<td>0.06</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>2.18</td>
<td>0</td>
<td>4.79</td>
<td>0.02</td>
<td>0.11</td>
<td>14.13</td>
<td>0</td>
<td>25.48</td>
<td>6.94</td>
<td>0.2</td>
<td>34.97</td>
<td>10.45</td>
<td>0.52</td>
<td>0.23</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>0.91</td>
<td>0.13</td>
<td>36.69</td>
<td>0.76</td>
<td>2.78</td>
<td>2.23</td>
<td>0</td>
<td>24.66</td>
<td>10.65</td>
<td>0.59</td>
<td>9.97</td>
<td>0.37</td>
<td>10.21</td>
<td>0.05</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>5.91</td>
<td>0.54</td>
<td>9.26</td>
<td>0.78</td>
<td>0.65</td>
<td>6.4</td>
<td>0</td>
<td>46.60</td>
<td>10.02</td>
<td>0.3</td>
<td>18.5</td>
<td>0.31</td>
<td>0.54</td>
<td>0.21</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>4.47</td>
<td>0.25</td>
<td>33.85</td>
<td>3.87</td>
<td>5.03</td>
<td>3.29</td>
<td>0</td>
<td>28.48</td>
<td>13.11</td>
<td>0.12</td>
<td>4.74</td>
<td>0</td>
<td>2.17</td>
<td>0.62</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>48.98</td>
<td>1.18</td>
<td>11.53</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>15.76</td>
<td>3.96</td>
<td>1.72</td>
<td>0</td>
<td>0</td>
<td>14.87</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* The first column of the table indicates cover types at the 20m resolution while the first row indicates cover types at 960m resolution. The numbers in each row indicate the percentages of the corresponding cover type at 20m resolution which were labeled as other cover types at the 960m resolution. For example, the row for class 1 shows that, at 960m resolution, 54.75% of original class 1 were labeled as class 1, 0.68% were labeled as class 3, 8.81% were labeled as class 4, and so on. Each row sums to 100. Note that class 2 and 15 did not appear at 960m resolution.
Table 7.4. Mixture-matrix between the 20m and 960m SPOT-derived land cover in the Niobrara study area. Percentage numbers were based on the 960m resolution classes.

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
<th>13</th>
<th>14</th>
<th>17</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>44.34</td>
<td>6.22</td>
<td>1.67</td>
<td>0.22</td>
<td>0.23</td>
<td>1.66</td>
<td>0</td>
<td>4.34</td>
<td>1.75</td>
<td>1.36</td>
<td>2.6</td>
<td>3.3</td>
<td>0.24</td>
<td>18.98</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>4.24</td>
<td>1.94</td>
<td>1.69</td>
<td>1.22</td>
<td>1.35</td>
<td>0.91</td>
<td>0</td>
<td>3.06</td>
<td>1.01</td>
<td>0.71</td>
<td>2.03</td>
<td>2</td>
<td>1.26</td>
<td>2.68</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>3.13</td>
<td>36.94</td>
<td>5.49</td>
<td>0.39</td>
<td>0.24</td>
<td>1.08</td>
<td>0</td>
<td>2.73</td>
<td>1.59</td>
<td>0.63</td>
<td>1.79</td>
<td>1.15</td>
<td>0.64</td>
<td>3.39</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>9.36</td>
<td>23.68</td>
<td>44.88</td>
<td>11.1</td>
<td>6.51</td>
<td>5.18</td>
<td>0</td>
<td>14.18</td>
<td>14.09</td>
<td>15.14</td>
<td>6.34</td>
<td>5.62</td>
<td>18.13</td>
<td>13.14</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>0.82</td>
<td>0.91</td>
<td>3.63</td>
<td>31.78</td>
<td>15.88</td>
<td>0.93</td>
<td>0</td>
<td>3.9</td>
<td>2.19</td>
<td>3.1</td>
<td>1.13</td>
<td>1.1</td>
<td>5.2</td>
<td>0.39</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>0.28</td>
<td>0.45</td>
<td>2.17</td>
<td>13.69</td>
<td>32.55</td>
<td>0.48</td>
<td>0</td>
<td>2.79</td>
<td>1.59</td>
<td>2</td>
<td>0.84</td>
<td>0.68</td>
<td>15.62</td>
<td>0.41</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>0.4</td>
<td>0.33</td>
<td>0.65</td>
<td>0.05</td>
<td>0</td>
<td>61.51</td>
<td>0</td>
<td>1.06</td>
<td>3.49</td>
<td>1.26</td>
<td>2.06</td>
<td>4.99</td>
<td>0.54</td>
<td>1.57</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>0.69</td>
<td>0.24</td>
<td>0.26</td>
<td>0.72</td>
<td>0.31</td>
<td>0.79</td>
<td>71.43</td>
<td>1.18</td>
<td>1.14</td>
<td>0.59</td>
<td>1.32</td>
<td>0.63</td>
<td>0.13</td>
<td>1.7</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>7.28</td>
<td>7.17</td>
<td>12.94</td>
<td>11.46</td>
<td>5.58</td>
<td>10.38</td>
<td>14.29</td>
<td>13.98</td>
<td>43.51</td>
<td>21.64</td>
<td>9.61</td>
<td>10.4</td>
<td>8.92</td>
<td>5.96</td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>0.7</td>
<td>0.3</td>
<td>1.54</td>
<td>1.32</td>
<td>0.33</td>
<td>2.1</td>
<td>0</td>
<td>1.25</td>
<td>7.09</td>
<td>38.9</td>
<td>1.46</td>
<td>2.12</td>
<td>0.96</td>
<td>0.71</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>1.95</td>
<td>0.12</td>
<td>1.88</td>
<td>2.16</td>
<td>1.8</td>
<td>4.09</td>
<td>0</td>
<td>8.15</td>
<td>4.82</td>
<td>4.48</td>
<td>45.58</td>
<td>12.67</td>
<td>1.27</td>
<td>3.56</td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>0.55</td>
<td>0.01</td>
<td>0.28</td>
<td>0.05</td>
<td>0.2</td>
<td>2.08</td>
<td>0</td>
<td>1.67</td>
<td>0.72</td>
<td>0.43</td>
<td>5.57</td>
<td>34.55</td>
<td>0.64</td>
<td>2.51</td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>0.65</td>
<td>1.06</td>
<td>6.09</td>
<td>5.79</td>
<td>14.69</td>
<td>0.92</td>
<td>0</td>
<td>4.52</td>
<td>3.09</td>
<td>3.59</td>
<td>4.46</td>
<td>3.39</td>
<td>35.4</td>
<td>1.39</td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>0.33</td>
<td>0.34</td>
<td>0.12</td>
<td>0.47</td>
<td>0.27</td>
<td>0.21</td>
<td>0</td>
<td>0.68</td>
<td>0.23</td>
<td>0.14</td>
<td>0.66</td>
<td>0.22</td>
<td>0.15</td>
<td>0.51</td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>0.13</td>
<td>0.08</td>
<td>0.24</td>
<td>1.24</td>
<td>1.11</td>
<td>0.06</td>
<td>0</td>
<td>0.22</td>
<td>0.16</td>
<td>0.03</td>
<td>0.09</td>
<td>0</td>
<td>0.32</td>
<td>0.79</td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>1.99</td>
<td>0.54</td>
<td>0.11</td>
<td>0</td>
<td>0</td>
<td>0.05</td>
<td>0</td>
<td>0.17</td>
<td>0.07</td>
<td>0</td>
<td>0.04</td>
<td>0</td>
<td>25.92</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* The first column indicates cover types at the 20m resolution while the first row indicates cover types at 960m resolution. The numbers in each column indicate the percentages of original 20m resolution cover type contained in the corresponding cover type at 960m resolution. For example, the column for class 1 shows that the 44.34% of class 1 at 960m were actually original class 1, 4.24% of it were actually original class 3, and so on. Each column sums to 100.

compared to other non-dominant classes (e.g., class 11 -- other grasses/low biomass), it gained area from those classes. Thus, its area did not decrease substantially. Among classes whose initial percentage areas were between 5% to 10%, class 7, corn, increased its area most significantly. The corn class had the largest patch size in all classes (MPS=64.3 Hectares) and was also most agglomerated (see Table 7.2). It was most likely to gain area from other neighboring classes. Although the other two classes in this range, class 1 (forest) and class 12 (subirrigated hay), had about 3.5% difference in initial areas,
their mean patch sizes and degree of conglomerate were similar (MPS=1.23 and 1.21 hectares, respectively). The two classes had almost the same amount of areal increase across the resolutions. Areas of all small classes, except class 17 -- water, decreased monotonically due to their small initial percentage areas and patch sizes. The area of water increased when pixel size was smaller than 480m and then decreased when the cell size further increased. The water class, although it had only 0.26% initial area, had the second largest mean patch size (MPS=4.56) among all classes. It tended to absorb other classes when pixel size increased. On the other hand, patches in this class, though large in size, were elongated in shape. Thus, they were eventually absorbed into other classes when pixel size exceeded the width of the patches.

7.3.2. Statistical analysis for the Niobrara 17 land cover classes

Seven landscape metrics with pairwise $r^2$ values less than 0.75 (Table 7.5), together with the initial area and spatial resolution, were used in the regression with

Table 7.5. Pairwise *Pearson-product-moment* correlations, $r^2$, among class indices. SPOT-derived land cover in the Niobrara study area.

<table>
<thead>
<tr>
<th></th>
<th>NP</th>
<th>MPS</th>
<th>PSSD</th>
<th>MSI</th>
<th>AWMSI</th>
<th>DLFD</th>
<th>IJI</th>
</tr>
</thead>
<tbody>
<tr>
<td>LPI</td>
<td>0.228</td>
<td>0.060</td>
<td>0.149</td>
<td>0.155</td>
<td>0.744</td>
<td>0.001</td>
<td>0.034</td>
</tr>
<tr>
<td>NP</td>
<td>0.080</td>
<td>0.049</td>
<td>0.078</td>
<td>0.264</td>
<td>0.373</td>
<td>0.001</td>
<td></td>
</tr>
<tr>
<td>MPS</td>
<td>0.974</td>
<td>0.169</td>
<td>0.006</td>
<td>0.594</td>
<td>0.106</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PSSD</td>
<td>0.219</td>
<td>0.005</td>
<td>0.540</td>
<td>0.124</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MSI</td>
<td></td>
<td>0.175</td>
<td>0.003</td>
<td>0.442</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AWMSI</td>
<td></td>
<td></td>
<td>0.075</td>
<td>0.028</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DLFD</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.003</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 7.6. Regression summary for the forward multi-linear regression between the nine predictor variables and coarse resolution class area. SPOT-derived land cover in the Niobrara study area.

<table>
<thead>
<tr>
<th>Var.</th>
<th>Coeff.</th>
<th>Std_err</th>
<th>( t )-ratio</th>
<th>( p )</th>
<th>( r^2 )</th>
<th>( r^2_{\text{inc.}} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( A_0 )</td>
<td>1.301</td>
<td>0.026</td>
<td>45.948</td>
<td>&lt;0.001</td>
<td>0.969</td>
<td>0.969</td>
</tr>
<tr>
<td>NP</td>
<td>-0.318</td>
<td>0.022</td>
<td>-13.005</td>
<td>&lt;0.001</td>
<td>0.982</td>
<td>0.013</td>
</tr>
<tr>
<td>LPI</td>
<td>-0.052</td>
<td>0.014</td>
<td>-3.278</td>
<td>&lt;0.001</td>
<td>0.986</td>
<td>0.004</td>
</tr>
</tbody>
</table>

coarse resolution class area and proportional error (PSSD was excluded because of its close correlation with MPS). In the forward stepwise regression with the coarse resolution class area as the dependent variable, a significant \( r^2 \) (0.986) is obtained (Table 7.6). The three predictor variables included in the regression model were: the original area, the number of patches, and the largest patch index. The original area was the first selected by the model (contributing 0.969 \( r^2 \)). The results of the regression conformed with the pattern of area changes shown in figures 7.1 - 7.3. That is, class area at a coarse resolution was primarily controlled by the initial area at full resolution (i.e., 20m). The number of patches, which reflected the mean patch size, was statistically significant in influencing the coarse resolution area. The negative regression coefficient of NP indicates that, for a given initial area, more patches would cause a decrease in area. The coefficient for the other predictor variable, LPI, was also negative. Large LPI for a class indicates the existence of a single large patch for the class, which could have two opposite effects on class areas at coarse resolutions. On one hand, it may result in a class area increase because the single large patch tends to expand as the resolution coarsens. On the other hand, the larger the dominant single patch the smaller the other patches, for
a given total class area and number of patches. The small patches tend to be absorbed when resolution coarsens, resulting in a decrease in the class area. The negative coefficient of LPI was due to the latter situation.

Table 7.7. Regression summary for the forward multi-linear regression between the nine predictor variables and proportional error. SPOT-derived land cover in the Niobrara study area.

<table>
<thead>
<tr>
<th>Var</th>
<th>Coeff.</th>
<th>Std_err</th>
<th>$t_{ratio}$</th>
<th>$p$</th>
<th>$r^2$</th>
<th>$r^2_{inc.}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>MSI</td>
<td>0.457</td>
<td>0.046</td>
<td>9.93</td>
<td>&lt;0.001</td>
<td>0.389</td>
<td>0.389</td>
</tr>
<tr>
<td>AWMSI</td>
<td>0.692</td>
<td>0.102</td>
<td>6.784</td>
<td>&lt;0.001</td>
<td>0.533</td>
<td>0.144</td>
</tr>
<tr>
<td>Res</td>
<td>-0.292</td>
<td>0.042</td>
<td>-6.952</td>
<td>&lt;0.001</td>
<td>0.621</td>
<td>0.088</td>
</tr>
<tr>
<td>DLFD</td>
<td>-0.266</td>
<td>0.052</td>
<td>-5.115</td>
<td>&lt;0.001</td>
<td>0.659</td>
<td>0.038</td>
</tr>
<tr>
<td>LPI</td>
<td>-0.223</td>
<td>0.097</td>
<td>-2.398</td>
<td>0.023</td>
<td>0.668</td>
<td>0.009</td>
</tr>
</tbody>
</table>

In the regression with proportional error as the dependent variable, a moderate $r^2$, 0.668, was obtained (Table 7.7). The first variable selected was the mean shape index, which contributed 0.389 $r^2$, followed by area weighted mean shape index, resolution, double log fractal, and largest patch index. Original area was not selected, due probably to the normalization in calculating $e$. The regression model demonstrated that patch shape indices, especially MSI and AWMSI, are important to the proportional error. The positive regression coefficients of AWMSI and MSI suggest that classes having a more complicated patch shape tended to result in a positive proportional error. This seems inconsistent with a general intuitive understanding since classes consisting of large, round homogeneous patches should have small perimeter to area ratios and, thus, have small
MSI and AWMSI values. The result obtained here was perhaps unique to the land cover maps classified from remote sensing data, lacking cartographic generalization. Examination of the original classification (fig. 3.14) reveals that it has an obvious salt-and-pepper appearance. Many patches consisted of only one pixel or a few pixels, resulting in a low shape index values. The MSI of a single pixel patches is 1, which is the minimum possible value for the index. These patches are most likely to be lost when resolution coarsens. The negative coefficient for resolution was due to the loss of boundary pixels at coarse resolutions.

7.3.3. Area Change for the Aggregated 25 Class US Land Cover

Results of the 25 class conterminous U.S. land cover dataset exhibited patterns similar to those displayed in the SPOT data. Class areas at coarse resolutions were predominantly determined by those at the original 1km resolution (figures 7.4 - 7.7). The percentage areas of the 25 land cover classes at the full resolution (1km) can be divided into following ranges: less than one percent, one to five percent, five to ten percent, and over ten percent. After being upscaled, percentage areas of most classes remained in the same ranges as at the original 1km resolution. There were, however, significant variations in the magnitude and direction of area change for classes having similar initial areas. Among the three largest classes (dryland cropland, woodland/cropland, and shrubland/grassland), only shrubland/grassland, the smallest, displayed sharp and near-monotonic increase in area when resolution became more coarse. Areas of the other two
classes stayed relatively steady (note that the abrupt change at 2 km resolution most likely resulted from random selection of the majority class; see also the 40m resolution for the SPOT-derived land cover). This could be explained by the characteristics of the spatial distribution of these classes. The patches of shrubland/grassland (MPS=63\text{km}^2) were much larger than those of the dryland cropland and woodland/cropland (MPS=29 and 26\text{km}^2, respectively) classes. Spatially, dryland cropland (class #1) and woodland/cropland (class #5) frequently neighbored each other (figure 7.8 and Table 7.8).

Table 7.8. Neighboring relationship among different land cover types in the 25-class AVHRR land cover of the conterminous U.S. The first column and row are cover types. Numbers in the table, \( C_{ij} \), indicate the percentages of one cover type neighboring another.

|    | 1  | 2  | 3  | 4  | 5  | 6  | 7  | 8  | 9  | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 | 22 | 23 | 24 | 25 |
|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|
| 1  | 64 | 1  | 2  | 6  | 14 | 3  | 0  | 1  | 0  | 4  | 1  | 0  | 0  | 0  | 0  | 0  | 0  | 1  | 0  | 0  | 0  | 0  | 0  | 0  | 0  |
| 2  | 1  | 1  | 4  | 4  | 2  | 7  | 8  | 5  | 1  | 2  | 2  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 1  | 0  | 0  | 0  | 0  | 0  |
| 3  | 24 | 22 | 23 | 21 | 22 | 10 | 3  | 0  | 6  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 1  | 0  | 0  | 0  | 0  |
| 4  | 12 | 1  | 1  | 4  | 4  | 13 | 0  | 6  | 0  | 4  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 1  | 0  |
| 5  | 14 | 0  | 1  | 2  | 53 | 1  | 0  | 0  | 0  | 0  | 1  | 6  | 5 | 0  | 4  | 1  | 0  | 4  | 5  | 0  | 0  | 0  | 0  | 0  |
| 6  | 5  | 0  | 0  | 9  | 26 | 1  | 1  | 1  | 0  | 0  | 0  | 0  | 0  | 0  | 2  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  |
| 7  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 2  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  |
| 8  | 1  | 0  | 0  | 3  | 0  | 9  | 13 | 69 | 0  | 2  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  |
| 9  | 1  | 4  | 0  | 1  | 1  | 11 | 0  | 0  | 56 | 4  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 17 | 5  | 0  | 0  | 0  | 0  |
| 10 | 17 | 1  | 0  | 8  | 5  | 2  | 1  | 6  | 0  | 55 | 0  | 0  | 0  | 0  | 0  | 2  | 1  | 0  | 0  | 0  | 0  | 0  | 0  | 0  |
| 11 | 5  | 0  | 0  | 0  | 0  | 0  | 22 | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 49 | 11 | 0  | 0  | 0  | 0  | 0  |
| 12 | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 7  | 60 | 0  | 0  | 0  | 0  | 0  | 18 | 0  | 0  | 0  | 0  | 0  | 0  |
| 13 | 3  | 1  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 38 | 0  | 26 | 1  | 1  | 0  | 29 | 0  | 0  | 0  | 0  | 0  | 0  |
| 14 | 3  | 0  | 0  | 0  | 26 | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 33 | 0  | 0  | 0  | 0  | 0  |
| 15 | 1  | 1  | 0  | 4  | 2  | 1  | 0  | 0  | 1  | 1  | 0  | 0  | 0  | 1  | 76 | 6  | 0  | 0  | 0  | 4  | 0  | 0  | 0  | 0  |
| 16 | 1  | 1  | 0  | 12 | 1  | 6  | 4  | 9  | 1  | 1  | 0  | 0  | 0  | 0  | 16 | 43 | 0  | 0  | 4  | 0  | 0  | 0  | 0  | 0  |
| 17 | 2  | 0  | 0  | 0  | 17 | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  |
| 18 | 1  | 0  | 0  | 0  | 8  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  |
| 19 | 1  | 1  | 0  | 3  | 1  | 2  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 5  | 0  | 28 | 9  | 0  | 0  | 50 | 0  | 0  | 0  |
| 20 | 2  | 0  | 2  | 1  | 0  | 1  | 3  | 7  | 0  | 1  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 2  | 0  | 58 | 19 | 0  | 0  | 5  |
| 21 | 1  | 21 | 0  | 1  | 3  | 4  | 0  | 3  | 0  | 3  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 15 | 10 | 36 | 0  | 0  | 0  |
| 22 | 22 | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  |
| 23 | 23 | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  |
| 24 | 0  | 0  | 0  | 1  | 0  | 4  | 3  | 9  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 15 | 2  | 0  | 0  | 0  | 0  | 0  |
| 25 | 1  | 0  | 0  | 1  | 4  | 11 | 2  | 2  | 4  | 0  | 1  | 3  | 3  | 0  | 0  | 1  | 4  | 1  | 4  | 3  | 0  | 1  | 1  | 2  | 0  | 0  | 43
Because they had similar initial areas and patch sizes, neither one was especially likely to be merged into another. Furthermore, most of the other classes neighboring these two classes had similar or larger patch sizes, e.g., grassland (MPS=33km$^2$), southeastern deciduous forest (MPS=25km$^2$), southeastern mixed forest (MPS=42km$^2$), and northern mixed forests (MPS=36km$^2$). Thus, the areas of these two classes did not increase significantly as pixel size increased. On the other hand, their areas were not likely to decrease because they were the two largest classes.

For classes falling in the second largest group (i.e., initial percent area > 5%) (figure 7.4), the most significant increase in area was observed for the western coniferous forest (class #15). The spatial distribution of this class was highly concentrated (see row 15 column 15 in Table 7.8) and it had a very large mean patch size (MPS=72km$^2$). Its increased area came mostly from two of its neighboring classes: western woodland (class #16) and grassland cropland (class #4), both of which had smaller patch sizes and initial class areas (Tables 7.8 and 7.9). Another class in this group showing considerable area increase was southeastern mixed forest (class #18), which was also concentrated and had a MPS of 42km$^2$. The area of the largest class in this group, grassland (class #6), increased but its magnitude was not as substantial as those of the other two classes. This class had the smallest mean patch size in this group (MPS=33km$^2$). It was widespread in the north central Great Plains, but was rather sparse in western and southwestern areas of the U.S. Therefore, the loss of area in the latter regions might have offset, to a certain extent, its areal gain in the northern Great Plains.
Table 7.9. Mixture-matrix between the 1km and 64km 25-class AVHRR-derived land cover of the conterminous U.S. The percentages are based on the 64km classes*.

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
<th>13</th>
<th>14</th>
<th>15</th>
<th>16</th>
<th>17</th>
<th>18</th>
<th>19</th>
<th>20</th>
<th>21</th>
<th>22</th>
<th>25</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>69</td>
<td>8</td>
<td>19</td>
<td>15</td>
<td>13</td>
<td>6</td>
<td>0</td>
<td>1</td>
<td>4</td>
<td>13</td>
<td>2</td>
<td>2</td>
<td>8</td>
<td>2</td>
<td>0</td>
<td>3</td>
<td>3</td>
<td>0</td>
<td>8</td>
<td>25</td>
<td>0</td>
<td>7</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>0</td>
<td>35</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>11</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>2</td>
<td>24</td>
<td>3</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>6</td>
<td>7</td>
<td>21</td>
<td>48</td>
<td>2</td>
<td>9</td>
<td>1</td>
<td>6</td>
<td>1</td>
<td>6</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>6</td>
<td>10</td>
<td>0</td>
<td>9</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>12</td>
<td>3</td>
<td>17</td>
<td>57</td>
<td>2</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>5</td>
<td>20</td>
<td>9</td>
<td>24</td>
<td>2</td>
<td>1</td>
<td>17</td>
<td>11</td>
<td>1</td>
<td>3</td>
<td>8</td>
<td>0</td>
<td>6</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>2</td>
<td>15</td>
<td>8</td>
<td>17</td>
<td>1</td>
<td>63</td>
<td>1</td>
<td>8</td>
<td>18</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>3</td>
<td>11</td>
<td>0</td>
<td>5</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>0</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>57</td>
<td>15</td>
<td>2</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>3</td>
<td>2</td>
<td>4</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>20</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>1</td>
<td>7</td>
<td>1</td>
<td>5</td>
<td>0</td>
<td>8</td>
<td>26</td>
<td>55</td>
<td>2</td>
<td>5</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>4</td>
<td>10</td>
<td>0</td>
<td>7</td>
<td>4</td>
<td>0</td>
<td>8</td>
<td>13</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>0</td>
<td>6</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>35</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>3</td>
<td>2</td>
<td>0</td>
<td>3</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>2</td>
<td>3</td>
<td>64</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>6</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>58</td>
<td>6</td>
<td>0</td>
<td>0</td>
<td>11</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>5</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>11</td>
<td>64</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>8</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>4</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>43</td>
<td>0</td>
<td>0</td>
<td>9</td>
<td>0</td>
<td>1</td>
<td>6</td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>0</td>
<td>4</td>
<td>3</td>
<td>3</td>
<td>0</td>
<td>3</td>
<td>1</td>
<td>4</td>
<td>13</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>65</td>
<td>14</td>
<td>0</td>
<td>19</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>0</td>
<td>4</td>
<td>6</td>
<td>3</td>
<td>0</td>
<td>3</td>
<td>3</td>
<td>5</td>
<td>8</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>6</td>
<td>31</td>
<td>0</td>
<td>11</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>7</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>65</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>4</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>17</td>
<td>20</td>
<td>0</td>
<td>0</td>
<td>66</td>
<td>0</td>
<td>5</td>
<td>11</td>
<td>0</td>
<td>4</td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>19</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>4</td>
<td>11</td>
<td>0</td>
<td>39</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>48</td>
<td>8</td>
<td>0</td>
<td>3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>21</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>15</td>
<td>40</td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>22</td>
<td>0</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>9</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>68</td>
<td>5</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>23</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>24</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>25</td>
<td>1</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>3</td>
<td>1</td>
<td>8</td>
<td>1</td>
<td>1</td>
<td>40</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* The first column indicates cover types at the 1 km resolution while the first row indicates cover types at 64 km resolution. The numbers in each column indicate the percentages of original 20m resolution cover type contained in the corresponding cover type at 960m resolution. For example, the column for class 15 shows that the 6% of class 15 at 64 km were actually original class 4, another 6% were actually original class 16, and so on. Each column sums to 100.

It is noted that the area of the class having the second largest patch size (MPS=51km²) in this group, desert shrubland (class #7), not only did not increase as resolution became coarser, but actually decreased slightly. This was due to the fact that
in most areas this class was spatially intermingled with another major class, shrubland/grassland (class #8) (figure 7.9 and Table 7.8), which has both larger initial area and larger mean patch size. About 14% of the shrubland/grassland was merged into desert shrubland while 32% of desert shrubland was combined into shrubland/grassland. The results suggest that, in addition to the initial class area, mean patch size and other spatial properties, the class neighbor relationships are also important.

Among classes having about five percent initial percentage areas, the grassland/cropland class displayed the most significant decrease in area, i.e., more than 2% (or 43% relative to its original area). This was apparently attributable to the fact that this class is sparsely distributed (MPS=15km²) as compared with two of its major neighboring classes, dryland cropland and grassland, which also had much larger initial areas. By comparison, the area of southeastern deciduous forest, which had almost the same initial area as grassland/cropland, but was more aggregated in spatial distribution (MPS=25km²), was rather stable across the resolutions. Similar observations are made for other classes having similar initial areas (i.e., about 3-5%). For example, the area of northern deciduous forest (class #11) decreased due to its dispersed distribution (MPS=13km²), while the area of its neighboring class, northern mixed forest (class #17), increased substantially due to its larger patches (MPS=36km²) and concentrated distribution (Table 7.8).

Areas of small classes (percentage area < 3%) decreased in most instances as resolution coarsened. However, class spatial characteristics still showed important effects on the pattern of area change. For example, the initial percentage area of barren land
(class #22) (mean patch size 144km$^2$), was only 2.3%, yet its class area increased until resolution reached 32km. The drop of area between 32km and 64km resolutions was likely due to the loss of some scattered patches in this class. Water (class #25) and western mixed forest (class #19) had similar initial areas (about 1%), the area of water decreased almost monotonically due to its small patch size (MPS=5km$^2$) and scattered distribution while the area of western mixed forest remained stable across resolutions due to its relatively large patch size (MPS=20km$^2$) (figure 7.7).

7.3.4. Statistical Analysis for the 25 class U.S. land cover data

The base area, resolution and seven landscape metrics (LPI was not used) with pairwise $r^2$ values less than 0.75 (Table 7.10) were used in the regression with class area. In the forward stepwise regression with coarse resolution class area as the dependent variable, a significant $r^2$, 0.990, was obtained (Table 7.11). Only two predictor variables were selected by the regression model, the original area and number of patches. The

<table>
<thead>
<tr>
<th>NP</th>
<th>MPS</th>
<th>PSSD</th>
<th>MSI</th>
<th>AWMSI</th>
<th>DLFD</th>
<th>IJI</th>
</tr>
</thead>
<tbody>
<tr>
<td>LPI</td>
<td>0.314</td>
<td>0.136</td>
<td>0.786</td>
<td>0.006</td>
<td>0.824</td>
<td>0.001</td>
</tr>
<tr>
<td>NP</td>
<td>0.006</td>
<td>0.114</td>
<td>0.007</td>
<td>0.426</td>
<td>0.225</td>
<td>0.067</td>
</tr>
<tr>
<td>MPS</td>
<td>0.538</td>
<td>0.078</td>
<td>0.080</td>
<td>0.412</td>
<td>0.243</td>
<td></td>
</tr>
<tr>
<td>PSSD</td>
<td>0.000</td>
<td>0.644</td>
<td>0.089</td>
<td>0.041</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MSI</td>
<td>0.017</td>
<td>0.103</td>
<td>0.025</td>
<td>0.004</td>
<td></td>
<td></td>
</tr>
<tr>
<td>AWMSI</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.087</td>
<td></td>
</tr>
</tbody>
</table>
Table 7.11. Regression summary for the forward multi-linear regression between the nine predictor variables and coarse resolution class area. Results of the 25-class AVHRR-derived land cover in the conterminous United States.

<table>
<thead>
<tr>
<th>Var</th>
<th>Coeff.</th>
<th>Std_err</th>
<th>t_ratio</th>
<th>p</th>
<th>$r^2$</th>
<th>$r^2_{inc.}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$A_0$</td>
<td>1.070</td>
<td>0.0138</td>
<td>2.308</td>
<td>&lt;0.001</td>
<td>0.986</td>
<td>0.986</td>
</tr>
<tr>
<td>NP</td>
<td>-0.097</td>
<td>0.013</td>
<td>-7.462</td>
<td>&lt;0.001</td>
<td>0.990</td>
<td>0.004</td>
</tr>
</tbody>
</table>

original area contributed 0.986 to the $r^2$ and the number of patches contributed 0.012. The results of regression agreed with previous analyses indicating that class area at a coarse resolution was predominantly determined by the initial area at full resolution (i.e., 1km). The landscape index that had the largest influence on coarse resolution area was the number of patches, which also represented patch size when combined with initial area. Its negative regression coefficient indicated that class area tended to decrease as the number of patches increased. The reason that the mean patch size index was not selected was perhaps because after the initial area (the most important predictor variable) was selected, the contribution from the number of patches was more than that from the mean patch size. When both the initial area and the number of patches were selected, the mean patch size was no longer needed because it is the ratio of the two.

In the regression with proportional error as the dependent variable, a relatively low $r^2$, 0.399, resulted (Table 7.12). The largest contribution of $r^2$ was from the area weighted mean shape index, 0.220, followed by resolution, mean patch size, and mean shape index. Original area was not selected by the regression due perhaps to the fact that it was
Table 7.12. Regression summary for the forward multi-linear regression between the nine predictor variables and proportional error. Results of the 25-class AVHRR-derived land cover in the conterminous United States.

<table>
<thead>
<tr>
<th>Var</th>
<th>Coeff.</th>
<th>Std_err</th>
<th>t_ratio</th>
<th>p</th>
<th>r²</th>
<th>r²_inc.</th>
</tr>
</thead>
<tbody>
<tr>
<td>AWMSI</td>
<td>0.135</td>
<td>0.020</td>
<td>6.75</td>
<td>&lt;0.001</td>
<td>0.220</td>
<td>0.220</td>
</tr>
<tr>
<td>Res</td>
<td>-0.174</td>
<td>0.034</td>
<td>-5.118</td>
<td>&lt;0.001</td>
<td>0.312</td>
<td>0.092</td>
</tr>
<tr>
<td>MPS</td>
<td>0.074</td>
<td>0.021</td>
<td>3.524</td>
<td>&lt;0.001</td>
<td>0.385</td>
<td>0.072</td>
</tr>
<tr>
<td>MSI</td>
<td>0.039</td>
<td>0.020</td>
<td>1.95</td>
<td>0.047</td>
<td>0.399</td>
<td>0.014</td>
</tr>
</tbody>
</table>

Normalized in the calculation of the proportional area. The results indicated that patch shape and size indices were important to proportional error. The positive regression coefficient for mean patch size indicated that increasing patch size tended to result in a positive proportional error. The negative coefficient for resolution was perhaps due to a decrease in the total area at coarse resolutions, which was attributed to the loss of some edge pixels during aggregation (i.e., windows containing more than 99% background pixels were not used). Similar to results observed from the Niobrara SPOT dataset, coefficients of AWMSI and MSI were positive, indicating classes with complicated shapes tended to increase in proportional error. The positive coefficients might also be related to the fact that those classes having large patches in the dataset usually exhibited extremely complicated boundaries (e.g., shrubland/grassland, desert shrubland, western coniferous forest, and southeastern deciduous forest) and thus had high shape index values. The smaller $r^2$ value, as compared to that in the class area regression, indicated that proportional error might not be accurately predicted by a simple multi-linear
regression model using landscape metrics. Nonetheless, the variables selected by the regression model still reflected the effect of landscape structure on class area variation as spatial resolution changed.

7.3.5. Statistical Analysis for the 159-class U.S. land cover data

Class areas and proportional estimation error at ten coarse resolutions were regressed against seven class level indices (Table 7.13, PSSD not used), the base area and spatial resolution. In the regression with the coarse resolution area as the dependent variable, the resultant $r^2$ was 0.972 (Table 7.14). All landscape indices were selected and the only predictor variable that was not included in the model was resolution. However, the base area ($A_0$) alone contributed 0.956 $r^2$ and the only two landscape indices that

Table 7.13. Pairwise Pearson-product-moment correlations, $r^2$, among class indices. AVHRR-derived 159-class land cover in the conterminous United States.
Table 7.14. Regression summary for the forward multi-linear regression between the nine predictor variables and coarse resolution class area. Results of the 159-class AVHRR-derived land cover in the conterminous United States.

<table>
<thead>
<tr>
<th>Var.</th>
<th>Coeff.</th>
<th>Std_err</th>
<th>r²</th>
<th>r²_inc.</th>
</tr>
</thead>
<tbody>
<tr>
<td>%A₀</td>
<td>1.056</td>
<td>0.013</td>
<td>0.956</td>
<td>0.956</td>
</tr>
<tr>
<td>NP</td>
<td>-0.168</td>
<td>0.010</td>
<td>0.969</td>
<td>0.013</td>
</tr>
<tr>
<td>AWMSI</td>
<td>0.100</td>
<td>0.010</td>
<td>0.971</td>
<td>0.002</td>
</tr>
<tr>
<td>IJI</td>
<td>0.034</td>
<td>0.006</td>
<td>0.971</td>
<td>0.000</td>
</tr>
<tr>
<td>MPS</td>
<td>0.071</td>
<td>0.010</td>
<td>0.972</td>
<td>0.000</td>
</tr>
<tr>
<td>DLFD</td>
<td>0.029</td>
<td>0.006</td>
<td>0.972</td>
<td>0.000</td>
</tr>
<tr>
<td>LPI</td>
<td>-0.063</td>
<td>0.014</td>
<td>0.972</td>
<td>0.000</td>
</tr>
<tr>
<td>MSI</td>
<td>-0.015</td>
<td>0.006</td>
<td>0.972</td>
<td>0.000</td>
</tr>
</tbody>
</table>

contributed to $r^2$ were the number of patches (NP) and the area weighted mean shape index (AWMSI), which together added 0.015 $r^2$. The results were similar to those observed from the 25-class AVHRR land cover dataset, where only $A_0$ and NP contributed to the $r^2$. Thus, original area was a predominant factor controlling the class area at coarse resolutions. The regression coefficient of NP was negative, indicating that more patches, for a given area, would result in a decrease in class area at a coarse resolution. The results were also similar to those obtained from the SPOT-derived land cover in the Niobrara area, where NP and LPI were the only two landscape indices included in regression model, with NP contributing to a much higher $r^2$ than the LPI.
Table 7.15. Regression summary for the forward multi-linear regression between the nine predictor variables and proportional error. Results of the 159-class AVHRR-derived land cover in the conterminous United States.

<table>
<thead>
<tr>
<th>Var.</th>
<th>Coeff.</th>
<th>Std_err</th>
<th>$r^2$</th>
<th>$r^2_{\text{inc.}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>MSI</td>
<td>0.229</td>
<td>0.026</td>
<td>0.153</td>
<td>0.153</td>
</tr>
<tr>
<td>AWMSI</td>
<td>0.815</td>
<td>0.049</td>
<td>0.256</td>
<td>0.103</td>
</tr>
<tr>
<td>DLFD</td>
<td>-0.064</td>
<td>0.027</td>
<td>0.275</td>
<td>0.020</td>
</tr>
<tr>
<td>IJI</td>
<td>0.233</td>
<td>0.026</td>
<td>0.293</td>
<td>0.017</td>
</tr>
<tr>
<td>NP</td>
<td>-0.059</td>
<td>0.028</td>
<td>0.302</td>
<td>0.010</td>
</tr>
<tr>
<td>Res</td>
<td>-0.097</td>
<td>0.021</td>
<td>0.310</td>
<td>0.007</td>
</tr>
<tr>
<td>MPS</td>
<td>0.475</td>
<td>0.047</td>
<td>0.317</td>
<td>0.007</td>
</tr>
<tr>
<td>LPI</td>
<td>0.229</td>
<td>0.026</td>
<td>0.361</td>
<td>0.044</td>
</tr>
</tbody>
</table>

In the regression with proportional error as the dependent variable, the resultant $r^2$ was 0.361 (Table 7.15), which was significantly lower than that obtained when using coarse resolution area as a dependent variable. The $r^2$ was very similar to the one obtained from the 25-class AVHRR land cover. The indices included in the regression and the relative contribution of the selected indices, however, were different. The most important landscape metrics in this regression model was mean shape index (MSI), followed by area weighted mean shape index (AWMSI). These two variables together contributed 0.256 to the $r^2$. In the 25-class dataset, AWMSI alone contributed 0.22 $r^2$ while MSI was the least important variable selected. The results were similar to the Niobrara SPOT dataset, where MSI and AWMSI were the two most important variables, but the order of relative importance from other indices varied among the three regressions.
7.3.6. Effects of upscaling techniques on land cover representation and on the relationship between land cover area and landscape structure indices

The results presented and discussed in the previous sections indicated that changes in land cover area across spatial resolution were related to landscape structure indices and some indices, such as the number of patches and area weighted mean shape index, appeared to be important in determining class area at coarse resolutions. In this section, similar analyses were performed for TM-derived land cover in the Lincoln/Omaha study area. In addition to the aggregation method used in the SPOT and AVHRR datasets, the re-classification method described in chapter 3 was utilized to derived coarse resolution TM land cover datasets.

a. Overall consistency of the two upscaling methods

Figure 7.10 shows the overall consistencies and Kappa statistics between the full and coarse resolution land covers obtained by the two methods. The aggregation method resulted in higher consistency between full and coarse land covers as compared to the re-classification method. The agreement between full and coarse resolution land cover could be considered, according to Table 6.1 (Monserud and Leemans, 1992), as good or very good when the pixel size was 120m or smaller for the aggregation method while only the 60m dataset could be put into this category for the re-classification method. This was because the aggregation method selected the class that was the largest in a coarse pixel and thus set the upper limit for the overall consistency measured by pixel to pixel
comparison. The agreement between the two methods was higher than that between full and coarse resolution datasets. At 120m or finer resolutions, the two methods produced *good or very good agreement*. The pixel to pixel agreement was over 70%. At more coarse resolutions, however, the consistency dropped to below 60% and there was only *fair (poor at 960m)* agreement between the two. Since the re-classification approach simulated land cover mapping using coarse resolution remote sensing data, while the aggregation approach simulated map generalization, the great discrepancy between the results obtained by the two upscaling methods at resolutions coarser than 120m indicated that there were different factors governing the change of land cover types and areas across resolutions. The results also suggested that it might be inappropriate to compare directly land cover maps derived from coarse resolution data (e.g., AVHRR images) with maps aggregated from fine resolution (e.g., TM images) classification.

*b. Change of class area with resolution: aggregation method*

The pattern of area variation differed greatly by land cover class (fig. 7.11 - 7.14). As observed in other study areas, areal change was determined by initial area and patch characteristics. Classes covering large areas and having larger a MPS relative to other classes at the 30m level were more likely to be retained at coarse resolutions. In rural areas, soybeans and corn, the two classes with the largest percent of initial areas (22% and 26%) and MPS (5.8 and 2.3 hectares), exhibited monotonic increases in area as the resolution became coarser. Grass, another spatially extensive class, on the other hand, decreased in area as resolution coarsened. This appeared to stem from the fact that grass
occurred in scattered, rather small patches (MPS=0.5 hectares) as compared with large-grain crops like corn. The areas of the two classes with the third and fourth largest MPS (wheat/oat and forest, MPS=1.2 and 1.6 hectares, respectively) increased at first and then decreased as cell size continued to increase. Areas of all other classes in the rural area decreased as resolution became more coarse. In urban areas (figs. 7.13 and 7.14), the two largest classes, grass and trees, displayed a monotonic increase in area as resolution became more coarse. Their MPS ranked second and third in area. The area of light color pavement, a class with small total area but the largest MPS (1.5 hectares) in the urban area, remained relatively stable when resolution size was smaller than 360m, and decreased as resolution became coarser. The areas of other urban classes decreased as the resolution coarsened due to relatively small initial areas and mean patch sizes.

The pairwise correlation coefficients, $r^2$, among the eight landscape indices were all less than 0.75 and thus all the indices were utilized, together with the base area and resolution, as predictor variables in the regressions. When coarse resolution class area was used as the dependent variable, six predictor variables were selected (Table 7.16).

Table 7.16. Pairwise Pearson-product-moment correlations, $r^2$, among class indices. TM-derived land cover in the Lincoln/Omaha study area.

<table>
<thead>
<tr>
<th></th>
<th>NP</th>
<th>MPS</th>
<th>PSSD</th>
<th>MSI</th>
<th>AWMSI</th>
<th>DLFD</th>
<th>III</th>
</tr>
</thead>
<tbody>
<tr>
<td>LPI</td>
<td>0.002</td>
<td>0.441</td>
<td>0.570</td>
<td>0.726</td>
<td>0.602</td>
<td>0.073</td>
<td>0.003</td>
</tr>
<tr>
<td>NP</td>
<td>0.048</td>
<td>0.071</td>
<td>0.027</td>
<td>0.039</td>
<td>0.031</td>
<td>0.056</td>
<td></td>
</tr>
<tr>
<td>MPS</td>
<td>0.746</td>
<td>0.555</td>
<td>0.144</td>
<td>0.439</td>
<td>0.110</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PSSD</td>
<td>0.579</td>
<td>0.423</td>
<td>0.271</td>
<td>0.148</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MSI</td>
<td>0.612</td>
<td>0.042</td>
<td>0.005</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AWMSI</td>
<td>0.002</td>
<td>0.016</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DLFD</td>
<td>0.081</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 7.17. Regression summary for the forward multi-linear regression between the ten predictor variables and coarse resolution class area. Results of the TM-derived land cover in the Lincoln/Omaha study area upscaled using the aggregation method.

<table>
<thead>
<tr>
<th>Var.</th>
<th>Coeff.</th>
<th>Std_err</th>
<th>t_ratio</th>
<th>p</th>
<th>r^2</th>
<th>r^2_inc.</th>
</tr>
</thead>
<tbody>
<tr>
<td>A_0</td>
<td>1.074</td>
<td>0.0422</td>
<td>5.57</td>
<td>&lt;0.001</td>
<td>0.897</td>
<td>0.897</td>
</tr>
<tr>
<td>NP</td>
<td>-0.314</td>
<td>0.022</td>
<td>-14.27</td>
<td>&lt;0.001</td>
<td>0.964</td>
<td>0.067</td>
</tr>
<tr>
<td>DLF</td>
<td>-0.163</td>
<td>0.030</td>
<td>-5.4</td>
<td>&lt;0.001</td>
<td>0.969</td>
<td>0.005</td>
</tr>
<tr>
<td>LPI</td>
<td>0.221</td>
<td>0.033</td>
<td>6.70</td>
<td>&lt;0.001</td>
<td>0.972</td>
<td>0.003</td>
</tr>
<tr>
<td>MSI</td>
<td>-0.138</td>
<td>0.029</td>
<td>-4.76</td>
<td>&lt;0.001</td>
<td>0.974</td>
<td>0.002</td>
</tr>
<tr>
<td>AWMSI</td>
<td>0.085</td>
<td>0.041</td>
<td>2.07</td>
<td>0.041</td>
<td>0.974</td>
<td>0.001</td>
</tr>
</tbody>
</table>

They were, in descending order of r^2 increment: original area, the number of patches, double log fractal, largest patch index, mean shape index, and area weighted mean shape index. Among these variables, original area alone contributed 0.897 to r^2 and number of patches contributed 0.067 (Table 7.17). The results were similar to those observed in the SPOT and AVHRR datasets. That is, the area of land cover at a coarse resolution was mainly determined by its initial area at the full resolution level. The mean patch size was not included in the regression model because the first two variables selected already took into account the mean patch size. There were conflicts in the signs of regression coefficients for similar variables. For example, both MSI and AWMSI indicated shape complexity but the two variables showed opposite signs (negative for MSI and positive for AWMSI). Because the relative contributions of the last four variables were very small, the conflicts were not significant to the regression.
Table 7.18. Regression summary for the forward multi-linear regression between the ten predictor variables and proportional error. Results of the TM-derived land cover in the Lincoln/Omaha study area upscaled using the aggregation method.

<table>
<thead>
<tr>
<th>Var.</th>
<th>Coeff.</th>
<th>Std_err</th>
<th>t_ratio</th>
<th>p</th>
<th>r^2</th>
<th>r^2_inc.</th>
</tr>
</thead>
<tbody>
<tr>
<td>MSI</td>
<td>0.445</td>
<td>0.082</td>
<td>5.43</td>
<td>&lt;0.001</td>
<td>0.622</td>
<td>0.622</td>
</tr>
<tr>
<td>Res</td>
<td>-0.193</td>
<td>0.030</td>
<td>-6.43</td>
<td>&lt;0.001</td>
<td>0.659</td>
<td>0.037</td>
</tr>
<tr>
<td>IJI</td>
<td>0.225</td>
<td>0.037</td>
<td>6.08</td>
<td>&lt;0.001</td>
<td>0.692</td>
<td>0.033</td>
</tr>
<tr>
<td>NP</td>
<td>-0.247</td>
<td>0.056</td>
<td>-4.41</td>
<td>&lt;0.001</td>
<td>0.737</td>
<td>0.046</td>
</tr>
<tr>
<td>DLF</td>
<td>-0.356</td>
<td>0.039</td>
<td>-13.54</td>
<td>&lt;0.001</td>
<td>0.769</td>
<td>0.032</td>
</tr>
<tr>
<td>AWMSI</td>
<td>0.707</td>
<td>0.097</td>
<td>7.29</td>
<td>&lt;0.001</td>
<td>0.801</td>
<td>0.032</td>
</tr>
<tr>
<td>LPI</td>
<td>-0.631</td>
<td>0.072</td>
<td>-5.26</td>
<td>&lt;0.001</td>
<td>0.830</td>
<td>0.029</td>
</tr>
<tr>
<td>A_0</td>
<td>0.323</td>
<td>0.111</td>
<td>2.91</td>
<td>0.004</td>
<td>0.834</td>
<td>0.004</td>
</tr>
</tbody>
</table>

In the regression with proportional error (e) as the dependent variable, the largest contributor to \( r^2 \) was mean shape index (MSI), followed by number of patches (NP), resolution (Res), and interspersion (IJI) (Table 7.18). MSI was also identified as the most important variable in the SPOT-derived and the AVHRR-derived 159-class land cover datasets. Although the largest contributor to \( r^2 \) was AWMSI in the AVHRR-derived 25-class dataset, it is of the same category as the MSI. These results suggest that patch spatial characteristics quantified by the shape index (either MSI or AWMSI) are the most important factor affecting the proportional error at coarse resolutions.

c. Change of class area with resolution: re-classification method

The change of class area across the range of spatial resolutions using the re-classification method exhibited different patterns from those observed using the
aggregation method, although large classes remained dominant and small classes remained small (fig. 7.15 - 7.18). For example, classes having distinctive spectral properties tended to have low commission errors but high omission errors in coarse resolution land covers, resulting in a decreased area as resolution became coarser. Three major classes in the rural area -- sorghum, soybeans and corn -- although intermixed, behaved very differently in terms of area change. The areas of sorghum and corn increased significantly, while the area of soybeans decreased sharply. This seemed to be related to the fact that soybeans had a very high near-infrared to red reflectance ratio as compared to sorghum and corn. The averaging of radiance in mixed-pixels at coarse resolutions resulted in a decrease in the near-infrared to red ratio and, thus, mixed-pixels were less likely to be classified as soybeans when the "pure" soybean spectral properties were used as training seeds. In the urban area, the most significant area increase occurred in the house/grass/tree class, itself a mixed-class at 30m resolution. As the resolution became coarser, more pixels became mixed-pixels with spectral characteristics similar to the mixed-class, and thus, were put into this class. The results suggested that with the re-classification method the spectral characteristics of the classes played a more important role in controlling changes in class area at coarse resolutions than was true in the aggregation method.

The regression model indicated that the initial class area contributed 0.929 $r^2$ increment to the prediction of coarse resolution class area (Table 7.19). Two other variables that significantly influenced the regression were the double log fractal (DLF) and the mean shape index (MSI). Using the re-classification method, the coarse
Table 7.19. Regression summary for the forward multi-linear regression between the ten predictor variables and coarse resolution class area. Results of the TM-derived land cover in the Lincoln/Omaha study area upscaled using the re-classification method.

<table>
<thead>
<tr>
<th>Var</th>
<th>Coeff.</th>
<th>Std_err</th>
<th>t_ratio</th>
<th>p</th>
<th>r²</th>
<th>r²_inc.</th>
</tr>
</thead>
<tbody>
<tr>
<td>A₀</td>
<td>1.060</td>
<td>0.021</td>
<td>49.64</td>
<td>&lt;0.001</td>
<td>0.929</td>
<td>0.929</td>
</tr>
<tr>
<td>DLF</td>
<td>-0.288</td>
<td>0.033</td>
<td>-8.74</td>
<td>&lt;0.001</td>
<td>0.945</td>
<td>0.016</td>
</tr>
<tr>
<td>MSI</td>
<td>0.260</td>
<td>0.027</td>
<td>9.63</td>
<td>&lt;0.001</td>
<td>0.952</td>
<td>0.007</td>
</tr>
</tbody>
</table>

Resolution class area was more dependent on the initial area and less on the landscape structure indices as compared to the aggregation method. In the regression with proportional error as the dependent variable, a $r^2$ value of 0.483 was obtained (Table 7.20). About 0.042 was contributed by the initial class area and the rest by four landscape indices: double log fractal (DLF), mean shape index (MSI), area weighted mean shape index (AWMSI), and interspersion (IJI). Comparing the $r^2$ with that obtained in

Table 7.20. Regression summary for the forward multi-linear regression between the ten predictor variables and proportional error. Results of the TM-derived land cover in the Lincoln/Omaha study area upscaled using the re-classification method.

<table>
<thead>
<tr>
<th>Var</th>
<th>Coeff.</th>
<th>Std_err</th>
<th>t_ratio</th>
<th>p</th>
<th>r²</th>
<th>r²_inc.</th>
</tr>
</thead>
<tbody>
<tr>
<td>DLF</td>
<td>0.878</td>
<td>0.078</td>
<td>11.26</td>
<td>&lt;0.001</td>
<td>0.331</td>
<td>0.331</td>
</tr>
<tr>
<td>A₀</td>
<td>0.191</td>
<td>0.066</td>
<td>2.89</td>
<td>0.004</td>
<td>0.374</td>
<td>0.042</td>
</tr>
<tr>
<td>AWMSI</td>
<td>-0.378</td>
<td>0.066</td>
<td>5.73</td>
<td>&lt;0.001</td>
<td>0.409</td>
<td>0.036</td>
</tr>
<tr>
<td>MPS</td>
<td>0.558</td>
<td>0.089</td>
<td>6.27</td>
<td>&lt;0.001</td>
<td>0.435</td>
<td>0.025</td>
</tr>
<tr>
<td>IJI</td>
<td>0.302</td>
<td>0.059</td>
<td>5.12</td>
<td>&lt;0.001</td>
<td>0.483</td>
<td>0.048</td>
</tr>
</tbody>
</table>
the aggregation method, it was obvious that the ability of landscape indices to predict proportional error decreased for the re-classified coarse resolution land cover maps.

7.3.7 Change of landscape level indices across resolutions

The relationship between landscape level indices (i.e., indices computed from the whole landscape without distinguishing individual cover types) and spatial resolution was presented in chapter 5 using simulated neutral landscapes. In this study, the dependence of seven landscape indices, derived from real landscapes, on spatial resolution were examined for different study areas and over a wide range of resolution levels. The seven landscape indices selected included: largest patch index (LPI), mean patch size (MPS), patch standard deviation (PSSD), landscape shape index (LSI), double log fractal dimension (DLF), patch interspersion and juxtaposition index (IJI), and contagion (CONTAG). Among these indices, four were previously examined in the simulated neutral landscapes (i.e., LPI, MPS, PSSD, and LSI). The other three were either not available (for IJI and CONTAG) for the neutral landscape with only one foreground land cover type, or were considered not reliable (for DLF) when the number of patches in a landscape was small (e.g., simulated landscape with \( P_i < 0.01 \)). The mean nearest neighbor distance (MNN) was not calculated in the real landscapes because of the large computational load.

Figures 7.19 -7.20 show changes in the landscape indices calculated from the Niobrara land cover maps with spatial resolutions ranging from 20m to 960m. Figures
7.21 - 7.22 are seven indices derived from the land cover maps of Lincoln/Omaha area. The indices at coarse resolutions (i.e., from 60m to 960m) were calculated from both the aggregated maps and re-classified maps. Figures 7.23 - 7.24 and 7.25 - 7.26 are results obtained from the 159-class and the grouped 25-class conterminous U.S. land cover maps, derived from AVHRR MVC NDVI images.

Generally, as the spatial resolution became coarser, indices describing the size of patches increased and those indicating the complexity of landscape decreased. The increase in MPS was due both to the size increase in the smallest possible patches (i.e., single pixel patches) and to the combining of small patches into larger patches as resolution coarsened. As MPS increased, the standard deviation among patches also increased at a similar rate. Increases in LPI appeared to stem from percentage area growth of the largest patches in the landscapes when the resolution coarsened. The decrease in landscape shape complexity at coarser resolutions was due both to the smoothing of patch boundaries and to the elimination of small patches. These results were in agreement with those observed using the simulated neutral landscapes. The three indices describing patch shape property and interspersion among patches (DLF, IJI, and CONTAG) showed less regularity. It was expected that DLF and IJI would decrease as the resolution became coarser, while CONTAG would increase. However, increases in DLF and IJI were observed for some landscapes and CONTAG decreased in one instance (the 25-class AVHRR data). The exact reason for this is not clear. One of the possible reasons is that errors might be introduced when variables (e.g., class areas, number of patches, and length of patch edges) in the equations used to calculate the indices have
very small values (see Appendix B).

Although the general trend of changes in most indices was as expected, the pattern of change differed among different indices, landscapes, degrees of class aggregation, and methods of upscaling. Regression analyses were utilized to determine if significant statistical relationships existed between the landscape indices and spatial resolution. The relationship between LPI and resolution was logarithmic in most instances, but was more linear in the TM dataset. Logarithmic regression resulted in significant $r^2$ values in all cases (Table 7.21). The relationship, however, was not transferable among different datasets due to considerable variations in regression slope. Mean patch size and the square of resolution was linearly associated (Table 7.22). This was because the resolution was expressed using pixel edge length but not the pixel area (e.g., 20 meters instead of 400 square meters). The slopes changed significantly among datasets, indicating the relationships were site/data specific. PSSD, linearly regressed against resolution, also displayed substantial differences in regression coefficients among datasets (Table 7.23). LSI had a reciprocal relationship with resolution (Table 7.24) and CONTAG displayed a similar relationship with resolution except in the case of the 25-class AVHRR dataset (Table 7.25). These results suggested that, although landscape indices could be regressed against resolution with significant $r^2$ in each dataset, the relationships were, almost without exception, data/area specific. These observations were also in agreement with those obtained using the simulated neutral landscapes. It seemed impossible to directly apply a regression model established in one area to another.
Table 7.21. Regressions between largest patch index and spatial resolution. The regressions were in the form of $LPI = a_0 + a_1 \ln(Res)$. $Res$ was in meters.

<table>
<thead>
<tr>
<th>Dataset</th>
<th>$a_0$</th>
<th>$a_1$</th>
<th>$r^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>SPOT</td>
<td>-5.588</td>
<td>2.373</td>
<td>0.978</td>
</tr>
<tr>
<td>TM (aggregation)</td>
<td>-24.526</td>
<td>6.157</td>
<td>0.853</td>
</tr>
<tr>
<td>TM (re-classification)</td>
<td>-17.082</td>
<td>4.366</td>
<td>0.899</td>
</tr>
<tr>
<td>AVHRR (159-class)</td>
<td>-1.233</td>
<td>0.384</td>
<td>0.912</td>
</tr>
<tr>
<td>AVHRR (25-class)</td>
<td>4.798</td>
<td>0.568</td>
<td>0.937</td>
</tr>
</tbody>
</table>

Table 7.22. Regressions between mean patch size and spatial resolution. The regressions were in the form of $MPS = a_0 + a_1 (Res)^2$. $MPS$ was in hectares and $Res$ was in meters.

<table>
<thead>
<tr>
<th>Dataset</th>
<th>$a_0$</th>
<th>$a_1$</th>
<th>$r^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>SPOT</td>
<td>7.554</td>
<td>0.00069</td>
<td>0.999</td>
</tr>
<tr>
<td>TM (aggregation)</td>
<td>-16.000</td>
<td>0.00087</td>
<td>0.992</td>
</tr>
<tr>
<td>TM (re-classification)</td>
<td>-0.220</td>
<td>0.00039</td>
<td>0.999</td>
</tr>
<tr>
<td>AVHRR (159-class)</td>
<td>40776.1</td>
<td>0.000196</td>
<td>0.962</td>
</tr>
<tr>
<td>AVHRR (25-class)</td>
<td>79333.781</td>
<td>0.00114</td>
<td>0.994</td>
</tr>
</tbody>
</table>

Table 7.23. Regressions between patch size standard deviation and spatial resolution. The regressions were in the form of $PSSD = a_0 + a_1 Res$. $PSSD$ was in hectares and $Res$ in meters.

<table>
<thead>
<tr>
<th>Dataset</th>
<th>$a_0$</th>
<th>$a_1$</th>
<th>$r^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>SPOT</td>
<td>-133.297</td>
<td>2.632</td>
<td>0.990</td>
</tr>
<tr>
<td>TM (aggregation)</td>
<td>-1105.818</td>
<td>9.128</td>
<td>0.939</td>
</tr>
<tr>
<td>TM (re-classification)</td>
<td>-311.542</td>
<td>3.248</td>
<td>0.966</td>
</tr>
<tr>
<td>AVHRR (159-class)</td>
<td>32008.988</td>
<td>35.466</td>
<td>0.993</td>
</tr>
<tr>
<td>AVHRR (25-class)</td>
<td>51426.391</td>
<td>195.917</td>
<td>0.997</td>
</tr>
</tbody>
</table>
Table 7.24. Regressions between landscape shape index and spatial resolution. The regressions were in the form of $LSI = a_0 + a_1(1/Res)$. Res was in meters.

<table>
<thead>
<tr>
<th>Dataset</th>
<th>$a_0$</th>
<th>$a_1$</th>
<th>$r^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>SPOT</td>
<td>14.550</td>
<td>9419.087</td>
<td>0.981</td>
</tr>
<tr>
<td>TM (aggregation)</td>
<td>20.700</td>
<td>18872.493</td>
<td>0.994</td>
</tr>
<tr>
<td>TM (re-classification)</td>
<td>38.575</td>
<td>18839.316</td>
<td>0.981</td>
</tr>
<tr>
<td>AVHRR (159-class)</td>
<td>24.489</td>
<td>467497.0</td>
<td>0.989</td>
</tr>
<tr>
<td>AVHRR (25-class)</td>
<td>13.089</td>
<td>293027.0</td>
<td>0.992</td>
</tr>
</tbody>
</table>

Table 7.25. Regressions between contagion and spatial resolution. The regressions were in the form of $CONTAG = a_0 + a_1(1/Res)$. Res was in meters.

<table>
<thead>
<tr>
<th>Dataset</th>
<th>$a_0$</th>
<th>$a_1$</th>
<th>$r^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>SPOT</td>
<td>62.329</td>
<td>-567.786</td>
<td>0.899</td>
</tr>
<tr>
<td>TM (aggregation)</td>
<td>68.126</td>
<td>-913.325</td>
<td>0.882</td>
</tr>
<tr>
<td>TM (re-classification)</td>
<td>62.215</td>
<td>-714.396</td>
<td>0.887</td>
</tr>
<tr>
<td>AVHRR (159-class)</td>
<td>54.985</td>
<td>-13950.0</td>
<td>0.950</td>
</tr>
<tr>
<td>AVHRR (25-class)</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
</tbody>
</table>

7.4. Summary and conclusions

Variations in land cover areas estimated from upscaled land cover datasets were examined both qualitatively and statistically. The absolute area of a land cover class at coarse resolutions was predominantly determined by the initial area in the full resolution data. Classes with small initial area tended to disappear as the resolution became coarser.
However, the role of landscape structure was apparent. Areas of classes having large patch sizes tended to increase, even when the initial area was not particularly large. In the regression of landscape indices to coarse resolution areas, the number of patches (NP), which reflected the mean patch size (MPS), was consistently included in all models as a variable having a significant contribution to the $r^2$ (not including the TM re-classification upscaling). When proportional errors were used as a dependent variable, either mean shape index (MSI) or area weighted mean shape index (AWMSI) was the most important parameter.

The signs of the regression coefficients for variables that contributed most to the $r^2$ (e.g., NP, MSI, and AWMSI) were consistent with different datasets. The magnitude of the coefficients, however, differed among all models, suggesting that the regression models were not transferable across different datasets. The consistent inclusion of NP, MSI and AWMSI in the regression models might be related to the formulation of the landscape indices. NP was simply the count of patch numbers for each class. MSI and AWMSI were computed from the ratio of patch perimeters to the square root of patch area. They were simple indices as compared with others such as the double log fractal. The fact that MSI and AWMSI had more contribution to the regression models than other indices indicated that the simple landscape indices were perhaps more useful in predicting proportional errors than the complicatedly formulated indices. Indices involving more advanced computations (e.g., DLFD which employs regressions) might produce unpredictable errors due to the complicated mosaic of patches. Actually, indices involving more complicated computation were significantly more noisy than simple ones.
(e.g., LSI vs. IJI). The pattern of change in the seven landscape level indices was generally as expected although there were unpredicted variations in DLFD, IJI, and CONTAG. The four simplest indices, -- LPI, MPS, PSSD, and LSI -- could be significantly regressed against resolution, and each with its own functional form, i.e., linear, logarithmic, and reciprocal. However, the regressions were still not directly comparable for different datasets.

Land cover datasets upscaled using different techniques exhibited significant differences. The agreement between the aggregated and re-classified TM datasets could only be categorized fair to poor when pixel size exceeded 120m. There were different rules governing the area change in land cover classes. Land cover areas were more dependent on their spectral properties than on spatial properties if coarse resolution land cover was derived from degraded spectral data. Areas with moderate spectral radiance values tended to increase as resolution became coarser, while those with distinctive spectral features decreased. In the aggregation method, classes possessing large initial areas and mean patch sizes tended to increase their areas. These observations have significant implications in real applications because land cover datasets having around 1 to 16 km resolutions are usually derived using AVHRR measurements or their degraded products. Validation of such land cover databases using high resolution data (e.g., TM) needs to take scaling issues into consideration. At coarser resolution levels (e.g., GCM models), aggregation or other generalization methods seemed to be more common than re-classification methods (e.g., Oleson et al., 1996; DeFries et al., 1997). Verification and calibration of those datasets probably needs to put more emphases on characteristics of
landscape structure.

Most landscape level indices, which included LPI, MPS, PSSD, LSI, and CONTAG, could be regressed against spatial resolution at very significant $r^2$ values. The regression forms (i.e., linear, reciprocal linear, etc) were the same for same indices, which indicate that the general relationships were consistent among different dataset. The regression coefficients, however, varied greatly from one dataset to another and the relationship seemed not to be extensible. The results are in agreement with those observed from simulated neural landscape and with the conclusion of Turner et al. (1989).

This study contributed to a better understanding of the role of landscape structure on spatial upscaling of land cover datasets and of landscape index patterns across spatial resolutions. The substantial variations in coefficient observed from the regressions either between class area and landscape indices or between landscape indices and resolution, however, suggested that considerable additional research is needed to derive generalized models.
7.5. References


Figure 7.1 Change of class area across spatial resolutions, results of the SPOT derived land cover in the Niobrara study area.
Figure 7.2 Change of class area across spatial resolutions, results of the SPOT derived land cover in the Niobrara study area.
Figure 7.3  Change of class area across spatial resolutions, results of the SPOT derived land cover in the Niobrara study area.
Figure 7.4 Change of class area across spatial resolutions, results of the 25-class AVHRR derived land cover (grouped from Loveland et al., 1995) of the conterminous U.S.
Figure 7.5 Change of class area across spatial resolutions, results of the 25-class AVHRR derived land cover (grouped from Loveland et al., 1995) of the conterminous U.S.
Figure 7.6 Change of class area across spatial resolutions, results of the 25-class AVHRR derived land cover (grouped from Loveland et al., 1995) of the conterminous U.S.
Figure 7.7 Change of class area across spatial resolutions, results of the 25-class AVHRR derived land cover (grouped from Loveland et al., 1995) of the conterminous U.S.
Figure 7.8 The spatial distributions of six of the 25 land cover classes
Figure 7.9 The spatial distributions of desert shrubland and mixed shrubland/grassland
Figure 7.10 Kappa statistics and consistency rate between full and coarse resolutions and between the two upscaling methods, results of the TM-derived land cover data in the Lincoln/Omaha study area.
Figure 7.11 Change of class area across resolutions, following the aggregation method.
Figure 7.12 Change of class area across resolutions, following the aggregation method.
Figure 7.13 Change of class area across resolutions, following the aggregation method.
Figure 7.14 Change of class area across resolutions, following the aggregation method.
Figure 7.15 Change of class area across resolutions, following the re-classification method.
Figure 7.16 Change of class area across resolutions, following the re-classification method.
Figure 7.17 Change of class area across resolutions, following the re-classification method.
Figure 7.18 Change of class area across resolutions, following the re-classification method.
Figure 7.19 Changes in landscape level indices across resolutions, results of the SPOT-derived land cover in the Niobrara study area.
Figure 7.20 Changes in landscape level indices across resolutions, results of the SPOT-derived land cover in the Niobrara study area.
Figure 7.21 Changes in landscape level indices across resolutions, results of the TM-derived land cover in the Lincoln/Omaha study area. Solid lines indicate results following the aggregation method and dash lines the re-classification method.
Figure 7.22 Changes in landscape level indices across resolutions, results of the TM-derived land cover in the Lincoln/Omaha study area. Solid lines indicate results following the aggregation method and dash lines the re-classification method.
Figure 7.23 Changes in landscape level indices across resolutions, results of the 159-class AVHRR-derived land cover of the conterminous U.S.
Figure 7.24 Changes in landscape level indices across resolutions, results of the 159-class AVHRR-derived land cover of the conterminous U.S.
Figure 7.25 Changes in landscape level indices across resolutions, results of the 25-class AVHRR-derived land cover (grouped from Loveland et al., 1995) of the conterminous U.S.
Figure 7.26 Changes in landscape level indices across resolutions, results of the 25-class AVHRR-derived land cover (grouped from Loveland et al., 1991) of the conterminous U.S.
CHAPTER 8
SUMMARY AND RECOMMENDATIONS FOR FUTURE RESEARCH

8.1. Uniqueness of the study

This dissertation has addressed problems in scaling, problems that are among the main challenges in remote sensing (Malingreau and Belward, 1992; Raffy, 1993). The principal objective of the research was to investigate the effects of changing spatial scale on the representation of land cover. A second objective was to determine the relationship between such effects, characteristics of landscape structure and scaling procedures. The study examined land cover/spatial resolution problems using data obtained from three different sensor systems in three ecological areas. The spatial resolutions involved ranged from 20m to 64km. The study investigated not only the scaling of land cover data but also the scaling of NDVI (the remote sensing parameter most frequently used in coarse resolution land cover mapping). Four fundamental research issues related to spatial scaling were examined: 1) the upscaling of NDVI, an index used widely in regional land cover characterization; 2) the effects of spatial scale on indices of landscape structure; 3) the representation of land cover databases at different spatial scales; and 4) the relationships between landscape indices and land cover area estimations.

Spatial scaling of NDVI has been previously examined by a number of investigators (e.g., Justice et al., 1989; Aman et al., 1992; Friedl et al., 1995; De Cola, 1997). However, no previous study explored relationships between the scaling non-
linearity of NDVI and land cover heterogeneity, and no direct comparison has heretofore been made between NDVI derived from sensors having different resolutions. Effects of spatial scale on landscape structure indices have also been investigated in other studies (e.g., Turner et al., 1989a, 1989b; Cullinan and Thomas, 1992; O’Neill et al., 1992; Qi and Wu, 1996), but prior studies have not shown how such indices react to the types of basic patch units and percentage foreground area. Upscaling of satellite-derived land cover data to GCM cell sizes has been conducted (e.g., Oleson et al., 1996), but no published work systematically examined the agreement between upscaled data and original land cover data in relation to the degree of dominance and heterogeneity of upscaled coarse pixels. Similarly, the influences of landscape structure on coarse resolution land cover mapping and the estimation of areal error due to spatial upscaling were investigated in earlier works (e.g., Mayaux and Lambin, 1995; Moody and Woodcock, 1995; Moody, 1996), but no previous studies have used as many data types, study areas, ranges of spatial resolutions, and upscaling methods as used here.

8.2. Spatial upscaling of NDVI

NDVI data derived from a TM image (30m resolution) and two SPOT images (20m resolution) were spatially upscaled to 1km using two methods: (1) averaging NDVI directly ($M_{NDVI}$) and (2) averaging reflectance prior to calculating NDVI ($S_{NDVI}$). The two upscaling methods were also used in upscaling AVHRR NDVI data from 1km to 64km. The results indicated that overall differences between NDVIs upscaled using the two
methods were small, i.e., less than 0.02 in magnitude, or less than 6% of the NDVI ranges under consideration. The magnitudes were equivalent to, or smaller than, the changes in NDVI caused by other factors such as water vapor and aerosols (Holben, 1986; Goward et al., 1991; Justice et al., 1991a). The results were in agreement with those observed by other investigators (Aman et al., 1992, Friedl et al., 1995). However, my results showed that the differences between $M_{\text{NDVI}}$ and $S_{\text{NDVI}}$ were dependent on both landscape heterogeneity (i.e., number of land cover types in a coarse pixel) and surface spectral properties. In highly heterogeneous areas, the differences could exceed 0.05 in magnitude, about 15% of the NDVI range in predominantly cropland and grassland study areas in Nebraska. The results suggest that the non-linearity of NDVI in relation to spatial resolution, although not significant overall, may introduce substantial bias, depending on the characteristics of the study areas.

There was substantial information loss when the SPOT and TM NDVIs were spatially upscaled from full resolution (20m and 30m, respectively) to 1km. The decreases in standard deviation (an important parameter used in image classification) ranged from about 25% to 40%. A 17% decrease in the standard deviation of the AVHRR MVC NDVI was observed. These decreases far exceeded the between-class variations among land cover types in the respective study areas and, thus, would greatly affect land cover classification.

In previous research, differences among NDVIs derived from different sensors have been theoretically discussed (e.g., Price, 1987) and investigations using simulated data have been conducted (e.g., Gallo and Daughtry, 1987). Direct comparisons using
actual satellite data were recommended (Gallo and Daughtry, 1987) but no studies on the
relationship between AVHRR MVC NDVI and TM and SPOT NDVIs have been
previously reported.

In this research, spatially upscaled NDVI derived from SPOT and TM data were
compared with AVHRR MVC NDVI at the 1km resolution level. The information
content of the AVHRR NDVI and the SPOT and TM NDVI was similar, with differences
in standard deviation ranging from 0 to 8%. Spatial patterns observed on the AVHRR
images and those seen on the upscaled TM and SPOT images were similar, although
AVHRR NDVI exhibited consistently lower values than TM and SPOT NDVI values.
Moderate to high r values (0.45, 0.79, and 0.81) were found from pixel to pixel linear
correlation between AVHRR/SPOT and between AVHRR/TM. The r values increased
when correlation was performed for pixels within 20° of nadir, suggesting that viewing
angle was a factor that may affect consistencies between the AVHRR and the upscaled
SPOT and TM NDVIs.

8.3. Effects of spatial resolution on landscape indices

The literature on the effects of scale on landscape structure parameters is rich and
extensive (e. g., Nellis and Briggs, 1989; Turner et al., 1989a, 1989b, 1991; Cullinan and
Thomas, 1992; O’Neill et al., 1992; Qi and Wu, 1996). However, investigations that
explore relationships between such effects and different basic patch types and changing
proportional areas of foreground land cover are rare. In fact, the relationships between
scale and landscape structure parameters vary as other conditions change. Investigations into scale/landscape relationships for various conditions are important because such studies provide the necessary basis for the development of possible generalized models. Changes in five landscape parameters for three types of landscapes across spatial resolutions and at various proportional areas of foreground land cover were examined in this research. It was found that the critical value of land cover proportion, at which a percolation patch formed, changed with different types of basic patch units which constituted a landscape. The percolation phenomenon occurred most prominently when the basic patch units were single pixels. As the size of basic patch units increased, the formation of percolation patches became less obvious. The critical value for a percolation patch to form decreased as resolution became more coarse.

Landscape indices describing the size characteristics of patches (LPI, MPS, PSSD, and MNN) generally increased as resolution coarsened but the magnitude of such increases changed with different proportional areas of foreground landscape and with different basic patch units forming the landscapes. There seemed no definite statistical relationships between these indices and spatial resolution, although near-linear relationships existed for some indices at certain values of foreground land cover area (e.g., mean patch size in the SPLS and CPLS landscapes when foreground areas were between 0.10 to 0.50). The landscape shape index exhibited more regular patterns across resolutions as compared to other indices. In the square-patch landscape (SPLS) and circular-patch landscape (CPLS), there were significant linear relationships between LSI and resolution, and, in the dot-patch landscape (DPLS), LSI's relationships with resolution
were reciprocal. Again, however, such relationships were not extensible from one landscape to another because the regression slopes and intercepts changed among landscapes and different proportional foreground areas. These results did not completely agree with those obtained in some other studies where landscape indices were found stable across resolution (e.g., Wickham and Ritters, 1995) or could be interpolated (e.g., Benson and MacKenzie, 1995). However, those studies were conducted in only one landscape. In fact, no literature was found that documented the same functional (or statistical) relationships between landscape indices and spatial scales in different landscapes. The results obtained in this study contribute to a better understanding of spatial resolution/landscape structure relationships in various landscapes, and suggest the need for additional research on the generalization of such relationships across landscapes.

8.4. Upscaling of land cover databases

Considerable spatial upscaling is needed when remotely sensed data are to be used in regional/global ecological and climatological models such as GCMs (Justice et al., 1991b; Hall et al., 1992). Aggregation methods based on the majority rule have been utilized in most investigations to derive coarse resolution land cover maps (e.g., Turner et al., 1989b; Benson and MacKenzie, 1995; Moody and Woodcock, 1995; Oleson et al., 1996). Few studies have, however, examined the agreement between the aggregated and the original land cover data and, the actual degree of majority a selected majority class represented. In this research, land cover databases derived from images acquired by three
sensor systems were aggregated to ten levels of coarser resolutions. The representation of land cover in the coarser resolution data was examined in relation to the character of the landscapes under investigation. Possible relationships between the degree of agreement and loss of image information content at coarser resolutions were explored. The results indicated good or very good overall agreement, as represented by Kappa statistics, between full and aggregated land cover data when the coarser pixels were up to four times as large as the original pixels. At very coarse resolutions (e.g., 64km), however, even AVHRR-derived land cover datasets exhibited poor representation in some areas, suggesting that it might be problematic to aggregate satellite-derived land cover data to GCM cell sizes.

The degree of agreement was dependent on the nature (i.e., pureness) of the full resolution pixels. The AVHRR-derived land cover had more heterogenous classes than land cover derived from SPOT and TM images, but exhibited consistently higher Kappa values because most of the original 1km pixels in AVHRR were mixed-pixels. The 25-class AVHRR land cover, grouped from the original 159 classes, showed the highest consistency rate due to its generalized class definitions. Similar results were observed when comparisons were made between the SPOT- and TM-derived datasets. The original TM dataset had more coarse resolution, covered a larger area, and had fewer classes (considering that the classes in the rural and urban areas were separately labeled) compared to the SPOT dataset. Nevertheless, the aggregated SPOT datasets displayed a higher degree of agreement because the cover types in the SPOT classification were more generally defined. The degree of agreement was related to the regional land surface
characteristics. Poor land cover representation was more likely in regions with complex land surface forms, because landscapes in such regions tended to be more heterogeneous.

Examination of the degree of dominance of majority classes in aggregated coarse pixels revealed that the term *majority* was misleading because the majority class often comprised a minority proportion of a coarse pixel. Two other parameters proposed in this research, pixel heterogeneity and degree of dominance, should be used together with the majority class to accurately portray the characteristics in aggregated coarse pixels. This is especially important when the aggregated land cover is used to parameterize ecoclimatological models (see, for example, Oleson *et al.*, 1996).

The degree of agreement between the original and aggregated land cover datasets could be significantly regressed against changes in image information content measured with NDVI standard deviation, regardless of the locations of study areas, types of sensor systems, and definitions of land cover classes. The results suggest that the ability to upscale land cover to coarser resolutions is closely related to image information content, although the exact form of the relationship between the accuracy of land cover classification and image information content might differ from those observed in this study.

8.5. Relationships between land cover upscaling and landscape structure

Recent research has shown that errors in areal estimation of land cover from coarse resolution remotely sensed data are significantly related to the spatial
characteristics of the landscape under investigation (e.g., Moody and Woodcock, 1994). Regression models were used to establish relationships between land cover areas measured from coarse resolution remotely sensed data and landscape structure indices (e.g., Mayaux and Lambin, 1995; Moody and Woodcock, 1995). Such models, if invertible, would enable the correction of areal errors inherent in coarse resolution land cover maps (e.g., Moody, 1996). Operational use of such models requires that they be generalizable over different geographical/ecological regions. Current models, however, have not been tested using data from validation sites (e.g., Mayaux and Lambin, 1995; Moody and Woodcock, 1995) or have only been tested in a site that had the same land cover characteristics as the area where the model was developed (e.g., Moody, 1996). In this study, relationships between land cover class area and landscape structure indices were investigated, both qualitatively and statistically, in three ecological regions with three types of images, i.e., SPOT, TM and AVHRR. For the AVHRR-derived land cover, an original 159-class U.S. dataset was generalized into a 25-class map, and comparisons between the statistical parameters obtained from the two datasets were conducted. For the TM-derived land cover, two upscaling approaches were utilized to generate coarse resolution data and the effects of landscape structure on coarse resolution land cover datasets were examined.

The absolute areas of a land cover class at coarse resolution was found to be predominantly determined by the initial area in the full resolution, regardless of the cover types, areas, and upscaling approaches used. The effects of landscape structure, however, were noticeable in all datasets. Areas of classes having larger mean patch sizes and
agglomerated distributions tended to increase, while those of scattered classes having small patch sizes tended to decrease. In all aggregated datasets, regression models using class area at a coarse resolution as the dependent variable consistently included the number of patches, which reflected the mean patch size, as a variable having significant contribution to the models. When proportional error was used as the dependent variable, either the mean shape index or area weighted mean shape index was the most important predictor variable, suggesting that landscape structure was the dominant factor controlling proportional error. The signs of regression coefficients of the three landscape structure variables (i.e., NP, MSI, AWMSI) that contributed most to regression models, both using absolute area or using proportional error as the dependent variable, were consistent for different study areas. This indicated that these three variables were indeed indicative of areal changes in land cover databases spatially aggregated to coarser resolutions.

The number of variables and the absolute magnitudes of coefficients of the variables, on the other hand, were different among all datasets, suggesting that regression models derived from different data and areas were not quantitatively interchangeable. Therefore, models developed in one area could not be used to calibrate areal errors in another area. The results, nevertheless, should not be considered discouraging because the regression models used in this study were simple linear models and they still identified consistently the most important landscape indices. Additional studies using more sophisticated models should be tested.

The results obtained from the two series of coarse resolution TM land cover databases indicated that upscaling techniques had a significant influence on the magnitude
and direction of land cover area changes across resolutions. Landscape structure played a more important role in the aggregation method than in the re-classification method. Land cover areas were more dependent on spectral properties than on spatial properties when the land cover was classified from degraded spectral data. Regression models showed that variables selected from the re-classified dataset were different from those selected from the aggregated dataset. In the regression with absolute area as the dependent variable, the number of patches, which was selected in all aggregated datasets, was not included. In the regression with proportional error as the dependent variable, AWMSI had a negative coefficient, which was contrary to those models developed from the aggregation method. These results suggest that there are different rules governing land cover area change across resolutions in the two upscaling methods. Models developed from the aggregation datasets did not seem applicable to datasets generated by the re-classification method. This has important implications to both validation and calibration of regional/global land cover datasets, because both upscaling methods can be used to produced such datasets. Impacts of upscaling methods should first be understood before any calibration models are applied.

8.6. Conclusions and recommendations for future research

This dissertation represents a comprehensive investigation into relationships between spatial resolution and land cover characteristics over a broad range of spatial resolutions and geographical/ecological areas. The results of this study contribute to a
better understanding of how land cover data derived from remotely-sensed data are influenced by spatial resolution and how such influences interact with various factors including landscape structure and rescaling techniques.

The major conclusions of this research are summarized as follows:

1) The overall bias resulting from the non-linearity of NDVI in relation to spatial resolution is generally insignificant as compared to other factors such as the influence of aerosols and water vapor. The bias however, is related to land surface characteristics. Significant errors (e.g., 15% of the NDVI range for the study area) may be introduced in heterogeneous areas where different land cover types exhibit strong spectral contrast.

2) Spatially upscaled SPOT and TM NDVIs have information content comparable with the AVHRR-derived NDVI. Pixel to pixel relationships, however, are dependent on other factors such as satellite viewing geometry.

3) Indices of landscape structure and spatial resolution are generally related, but the exact forms of the relationships are subject to changes in other factors including the basic patch unit constituting a landscape and the proportional area of foreground land cover under consideration.

4) The extent of agreement between a full resolution land cover dataset and its spatially aggregated products changes with the properties of the original dataset, including the pixel size and class definition. The class selected by the majority rule often represents a minority proportion of a coarse pixel. Accurate descriptions of the nature of a spatially aggregated land cover dataset should include information regarding pixel heterogeneity and the degree of dominance.
5) There are close relations between landscape structure and class areas estimated from spatially aggregated land cover databases. The relationships, however, do not permit extension from one area (and/or dataset) to another. Inversion calibration across different geographic/ecological areas is, therefore, not feasible. Different rules govern the land cover area changes across resolutions when different upscaling methods are used. Special attention should be given to the comparison between land cover maps derived using different methods.

Results and conclusions from this study point to the need for additional research in several areas:

1) Scaling of the NDVI in relation to atmospheric conditions, viewing/illumination geometry, bandwidth/locations, and surface bidirectional reflectance distribution functions. There are still many uncertainties regarding the scaling of NDVI. These must be resolved if relationships between NDVI and plant biophysical parameters observed at leaf/canopy/plot levels are to be transferable to satellite measurements and further to coarse grids such as those used in GCMs. Such transformations are subject to influences from the atmosphere, the sensor system, and surface characteristics. Research accounting for all these factors is needed to develop solid and robust NDVI scaling models.

2) Examination of the nature of the relationships among landscape structure indices, spatial resolutions, types of basic patch units, and foreground land cover areas. General relationships between landscape structure indices and spatial resolution have been observed. The relationships, however, are influenced by many factors including the types of basic patches comprising the landscape and the proportional area
of the land cover under consideration. Although this study documented qualitatively and statistically (for some indices) such interrelationships, additional experiments and quantitative analyses are needed to explore the exact nature of such relationships.

3) Development of indices to better characterize the quality of spatially upscaled land cover databases and to relate such indices to ecoclimatological models.

Results of this study showed that the land cover class selected by aggregation using the majority rule might represent only a minority proportion of a coarse pixel. Significant errors may result from using such aggregated coarse pixels to parameterize ecoclimatological models. These errors can be better predicted if the heterogeneity and degree of dominance of the coarse pixels are known. Therefore, indices that can better characterize the quality of aggregated land cover need to be developed and used to parameterize ecoclimatological models.

4) Investigation of the stability and extensibility of relationships between land cover area estimated from coarse resolution and landscape structure indices.

Although the effects of landscape structure on land cover area estimation were documented in this and previous studies, no relationships between the two which were extensible across different regions have been found at this point. More studies are needed to test a) if generalizable forward models can be found, and b) if forward models are, should they exist, invertible.

5) Further investigation of the impact of upscaling techniques on land cover representation.

This research indicated that scaling methods have significant influences on the
representation of land cover at coarse resolutions. The factors identified in controlling land cover area change using the majority rule method were very different from those using the re-classification method. This had significant implication on the development of calibration models, as application of such models assumes that the controlling factors are the same. More experimentation is needed to investigate if relationships between differently upscaled land cover databases can be found, and if such relationship can be used to link regression models developed from those databases.
8.7. References


Appendix A

Following are brief descriptions of the class level landscape structure indices used in this study. Class level indices are measured from the land cover class under consideration, but not from an entire landscape. All indices are derived from Fragstats (McGarigal and Marks, 1993).

*Largest patch index* (LPI): the percentage of the landscape that the largest patch in land cover class $i$ comprises.

$$LPI = \frac{1}{A} \frac{n_i}{\max(a_{ij})}$$

*Number of patches* (NP): the total number of patches into which a land cover class $i$ is fragmented.

$$NP = n_i$$

*Mean patch size* (MPS): the ratio of total area of a land cover class $i$ to the number of patches in the class.

$$MPS = \frac{1}{n_i} \sum_{j=1}^{n_i} a_{ij}$$

*Patch size standard deviation* (PSSD): the standard deviation of patch sizes to mean patch size in land cover class $i$, indicating the degree of variation in patch sizes.

$$PSSD = \sqrt{\frac{1}{n_i} \sum_{j=1}^{n_i} (a_{ij} - \frac{1}{n_i} \sum_{j=1}^{n_i} a_{ij})^2}$$

*Mean shape index* (MSI): the average patch perimeter to area ratio of land cover class $i$, representing the complexity of patch shape.

$$MSI = \frac{1}{n_i} \sum_{j=1}^{n_i} \left( \frac{p_{ij}}{\sqrt{a_{ij}}} \right)$$
Area-weighted mean shape index (AWMSI): similar to MSI, but takes patch area into account so that larger patches are weighted more than smaller patches.

\[
AWMSI = \sum_{j=1}^{n_i} \left( \frac{p_{ij}}{a_{ij}} \right) \left( \frac{a_{ij}}{\sum_{j=1}^{n_i} a_{ij}} \right)
\]

Double log fractal (DLF): a measure of patch shape complexity of land cover class \(i\), approaching 1 for very simple shapes and 2 for highly convoluted patches.

\[
DLF = \frac{2}{\left[ n_i \sum_{j=1}^{n_i} \ln(p_{ij}) \ln(a_{ij}) \right] \left[ \left( \sum_{j=1}^{n_i} \ln(p_{ij}) \right) \left( \sum_{j=1}^{n_i} \ln(a_{ij}) \right) \right]^{1/2}}
\]

Interspersion and juxtaposition index (IJI): indicating the degree to which patches of land cover class \(i\) mixed with patches of other land cover types within a landscape.

\[
IJI = \frac{1}{\ln(m-1)} \left[ -\sum_{k=1}^{m} \left( \frac{e_{ik}}{\sum_{k=1}^{m} e_{ik}} \right) \ln\left( \frac{e_{ik}}{\sum_{k=1}^{m} e_{ik}} \right) \right]
\]

where

- \(i\) and \(k\) = 1, 2, ..., \(m\) indicate patch types (land cover classes),
- \(j=1, 2, ..., n_i\) indicates number of patches in class \(i\),
- \(A\) is the total landscape area,
- \(m\) is the number of classes,
- \(n_i\) is the total number of patches in class \(i\),
- \(a_{ij}\) is the area of patch \(j\) in class \(i\),
- \(p_{ij}\) is the perimeter of patch \(j\) in class \(i\), and
- \(e_{ik}\) is the total length of edge in landscape between class types \(i\) and \(k\).
Appendix B

Following are brief descriptions of the landscape level landscape structure indices used in this study. Most landscape level indices are similar to the class level indices, except that the landscape level indices are measured from the entire landscape. All indices are derived from Fragstats (McGarigal and Marks, 1993).

*Largest patch index* (LPI): the percentage of the landscape that the largest patch in the landscape comprises.

\[
LPI = \frac{1}{A} \sum_{j=1}^{n} \max(a_j)
\]

*Number of patches* (NP): the total number of patches in the landscape.

\[
NP = \sum_{i=1}^{m} n_i
\]

*Mean patch size* (MPS): the ratio of total area of the landscape to the number of patches in the landscape.

\[
MPS = \frac{A}{\sum_{i=1}^{m} n_i}
\]

*Patch size standard deviation* (PSSD): the standard deviation of patch sizes to mean patch size in the landscape, indicating the degree of variation in patch sizes.

\[
PSSD = \sqrt{\frac{1}{m} \sum_{i=1}^{m} \sum_{j=1}^{n_i} \left( a_{ij} - \frac{A}{m} \right)^2}
\]

*Mean shape index* (MSI): the average patch perimeter to area ratio in the landscape, representing the complexity of patch shape.
Area-weighted mean shape index (AWMSI): similar to MSI, but takes patch area into account so that larger patches are weighted more than smaller patches.

\[
AWMSI = \sum_{i=1}^{m} \sum_{j=1}^{n_i} \left[ \frac{p_{ij}}{\sqrt{a_{ij} A}} \right]
\]

Double log fractal (DLF): a measure of patch shape complexity of a landscape, approaching 1 for very simple shapes and 2 for highly convoluted patches.

\[
DLF = \frac{2}{\left[ \sum_{i=1}^{m} \sum_{j=1}^{n_i} (\ln p_{ij} \ln a_{ij}) \right] - \left[ (\sum_{i=1}^{m} \sum_{j=1}^{n_i} \ln p_{ij}) (\sum_{i=1}^{m} \sum_{j=1}^{n_i} \ln a_{ij}) \right]}
\]

Interspersion and juxtaposition index (JI): indicating the degree to which patches of different land cover types are mixed together within a landscape.

\[
JI = (\ln \left[ \frac{1}{2} m(m-1) \right])^{-1} \left[ -\sum_{i=1}^{m} \sum_{k=i+1}^{m} \left( \frac{e^k}{E} \right) \ln \left( \frac{e^k}{E} \right) \right]
\]

Mean nearest-neighbor distance (MNN): The sum of the distance to the nearest neighboring patch of the same type based on nearest edge-to-edge distance, divided by the number of patches with a neighbor.

\[
MNN = \frac{1}{m} \sum_{i=1}^{m} \sum_{j=1}^{n_i} \frac{h_{ij}}{\sum_{i=1}^{m} \sum_{j=1}^{n_i} h_{ij}}
\]

Contagion index (CONTAG): the observed contagion over the maximum possible contagion for the given number of patch types.
\[
CONTAG = 1 + \frac{1}{2 \ln(m)} \sum_{i=1}^{m} \sum_{k=1}^{m} \left[ P_i \left( \frac{g_{ik}}{m} \right) \ln P_i \left( \frac{g_{ik}}{m} \right) \right] \sum_{k=1}^{m} g_{ik} \sum_{k=1}^{m} g_{ik}
\]

where

- \( i \) and \( k \) = 1, 2, ..., \( m \) indicate patch types (land cover classes),
- \( j \) = 1, 2, ..., \( n \) indicates number of patches in class \( i \),
- \( A \) is the total landscape area,
- \( m \) is the number of classes,
- \( n_i \) is the total number of patches in class \( i \),
- \( n \) is the total number of patches in the landscape (\( n = \sum n_i \)),
- \( a_{ij} \) is the area of patch \( j \) in class \( i \),
- \( p_{ij} \) is the perimeter of patch \( j \) in class \( i \), and
- \( e_{ik} \) is the total length of edge in landscape between class types \( i \) and \( k \).
- \( h_{ij} \) is the edge-to-edge distance between patch \( i \) and \( j \),
- \( E \) is the total length in the landscape,
- \( P_i \) is the proportion of land cover class \( i \) in the landscape, and
- \( g_{ik} \) is the number of adjacencies between pixels of patch type \( i \) and \( k \).