HERO OBJECT-BASED LAWN MAPPING EXPLORATION OF SUBURBIA:

RATIONALE, METHODS AND RESULTS FOR THE NSF PLUM ISLAND ECOSYSTEMS LONG-TERM ECOLOGICAL RESEARCH SITE

WORKING PAPER NO. 2012-24
DECEMBER 2012

By Colin Polsky, Robert Gilmore Pontius, Jr., Nicholas M. Giner, Albert Decatur, Daniel Miller Runfola, and Rahul Rakshit

HUMAN

CONSTRUCTED

1887
CLARK UNIVERSITY
George Perkins Marsh Institute

HERO Object-based Lawn Mapping Exploration of Suburbia: Rationale, Methods and Results for the NSF Plum Island Ecosystems Long-Term Ecological Research Site

Colin Polsky[1]*, Robert Gilmore Pontius, Jr. [1], Nicholas M. Giner[1], Albert Decatur[1], Daniel Miller Runfola[1], and Rahul Rakshit[1]

[1] Human-Environment Regional Observatory (HERO) and Graduate School of Geography; Clark University; 950 Main Street; Worcester, MA 01610

* Corresponding Author: Graduate School of Geography; Clark University; 950 Main Street; Worcester, MA 01610; 508-421-3828 tel.; cpolsky@clarku.edu

Date of completion: 6 December 2012

Abstract: This document outlines the efforts of the *holmes* team, part of the Human-Environment Regional Observatory (HERO; http://hero.clarku.edu) program at Clark University in Worcester, Massachusetts (USA). The HERO lab dates from 1999, and is composed of several teams, each assigned to answer different questions about the human-environment relationship in a multi-disciplinary environment (Polsky et al. 2007). For most of the >10 years now that the HERO lab has been in operation, the program has been run as a formal National Science Foundation Research Experiences for Undergraduates Site (NSF REU Site) program. This prestigious affiliation means that the student team is, in the summers, composed of some of the country's leading undergraduate students interested in research and action on environmental questions involving land use and land cover. During the academic years, the student teams are composed almost entirely of undergraduates from Clark University.

Since 2007, one of HERO's teams has been called *holmes*, which stands for "HERO Object-based Lawn Mapping Exploration of Suburbia". The *holmes* team focuses on collecting and analyzing geographic data to represent contemporary suburban land cover. Even though this focus generates intrinsic scholarly and applied merits, it also has an instrumental value: to assist scholarly endeavors by other HERO teams and by potential other users not in the HERO lab.

Introduction

This Marsh Institute White Paper reports results from the primary product of the *holmes* team: A collection of fine-resolution (0.5 meter) land-cover maps of 26 suburban Boston towns, for the year 2005. These towns are selected to represent the National Science Foundation's Plum Island Ecosystems Long-Term Ecological Research (NSF PIE LTER) study site . This collection of maps constitutes, to our knowledge, only the second "synoptic" (i.e., full spatial coverage as opposed to a sample of places) fine-resolution land-cover map that has *lawn* as a land-cover category, and the most spatially extensive to date. (Our colleagues at the NSF Baltimore Ecosystem Study LTER site, who instructed us in the methodology (see Zhou and Troy, 2008 and Zhou et al. 2008 for details), produced the first such map, albeit for a slightly smaller study area, and with a less precise definition of *lawn* than in the *holmes* product.) These characteristics of the *holmes* maps reported here will accelerate a new generation of precise studies of, among other topics, homeowner lawncare practices. The reason is that the fine spatial resolution of the *holmes* products permits us to make statements about household parcel-level land cover. Thus, when coupled with other relevant data sources (e.g., homeowner surveys/focus groups, Census data, biogeochemical/water transport modeling), the *holmes* product promises to permit valid and reliable tests of homeowner-level hypotheses.

The *holmes* land-cover categories include bare soils, coniferous tree canopy, deciduous tree canopy, impervious surfaces like roads and rooftops, open water, wetlands, and, of course, fine green grass vegetation (a subset of which is the lawns category, defined simply as fine green grass located in a residential property). The landscape was mapped from imagery (MassGIS, 2011) containing pixels that were 0.5 m on a side, meaning that landscape features such as individual shrubs, sidewalks, and patches of lawn grass were distinguished by our mapping process. The fine spatial resolution of the imagery data is necessary to distinguish the amount of lawn cover present on a household parcel.

The *holmes* project finished this synoptic mapping task in the summer of 2011. The mapping involved a team of up to seven people (depending on the time of year) over four years. Much of this time the team was only a few members, and that our methods have become much faster as our understanding of managing a large, multi-member research project has grown. The methods outlined in this document do not cover the very early attempts by *holmes* to make an effective workflow, but only the later methods with which the great majority of the maps were produced. The important difference with the techniques used later in the project is that they were far faster than the earlier methods, and allowed much more area to be mapped with the same number of person hours expended.

The remainder of this document presents the primary results of the *holmes* work, followed by an Appendix that includes additional information on the *holmes* project, namely: background of the project, data acquisition and management procedures, methods, project standards, and overall suggestions for completing a large, multi-member research project.

Acknowledgements

The United States National Science Foundation (NSF) supported this work via the following programs: Long Term Ecological Research via grants OCE-0423565, OCE-1026859 & OCE- 1238212 for Plum Island Ecosystems and OCE-0620959 for Georgia Coastal Ecosystems, Coupled Natural Human Systems (CNH) via grant BCS-0709685, Research Experiences for Undergraduates (REU Site) via grant SES-0849985, Urban Long Term Research Areas (ULTRA-ex) via grant BCS-0948984, and "Maps and Locals (MALS)" via grant DEB-0620579. Additional support for this research was provided by the Mosakowski Institute for Public Enterprise, the O'Connor '78 Endowment for the Environment, and the George Perkins Marsh Institute, all at Clark University. Any opinions, findings, conclusions, or recommendation expressed in this paper are those of the authors and do not necessarily reflect those of the funders. The Massachusetts Geographic Information System (MASSGIS) supplied some of the data for this project. All of the authors of this document save for Polsky and Pontius were Clark University students. Thank you to the following *additional* (mostly undergraduate) students who were instrumental in the production of this dataset: Jenner Alpern (Clark University); Katrina Barney (New Mexico State University); Sarah Geise (Carthage College); Jackie Gushue (Boston University); Thomas Hamill (Clark University); Cory Keeler (Grinnell College); Joe Krahe (Clark University); Cait McCann (University of Texas, Austin); Nicholas Perdue (Metropolitan University of Denver); Nagraj Rao (Clark University); Matthew Salem (Arizona State University); and James Wilson (Clark University). Major mapping assistance was provided by J. Morgan Grove (US Forest Service) and Jarlath O'Neil-Dunne (University of Vermont) through the affiliation with the NSF LTER Baltimore Ecosystem Study (BES) project (http://beslter.org/).

Contents

Maps showing the PIE study area and lawns in the PIE study area 4

Results of the land-cover mapping effort and map accuracy assessment 6

Table of land-cover percentages by town in the PIE study area 7

Bar charts showing percentages of each individual land-cover category, by town 8

Land-cover maps and pie charts for the PIE study area and individual towns 12

Error matrices and Figures of Merit for the PIE study area and individual towns 39

References 66

Appendix 67

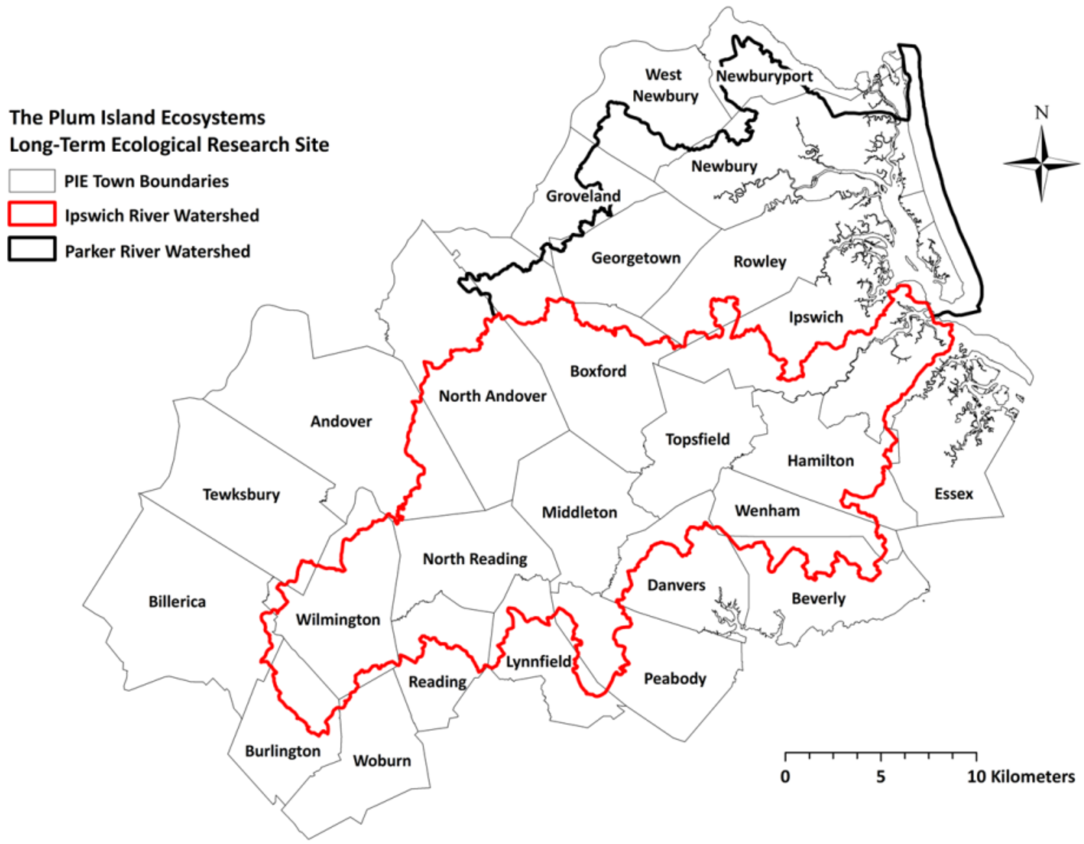


Figure 1. The Plum Island Ecosystems LTER study area. The study area covers 1143 km^2 in northeastern Massachusetts.

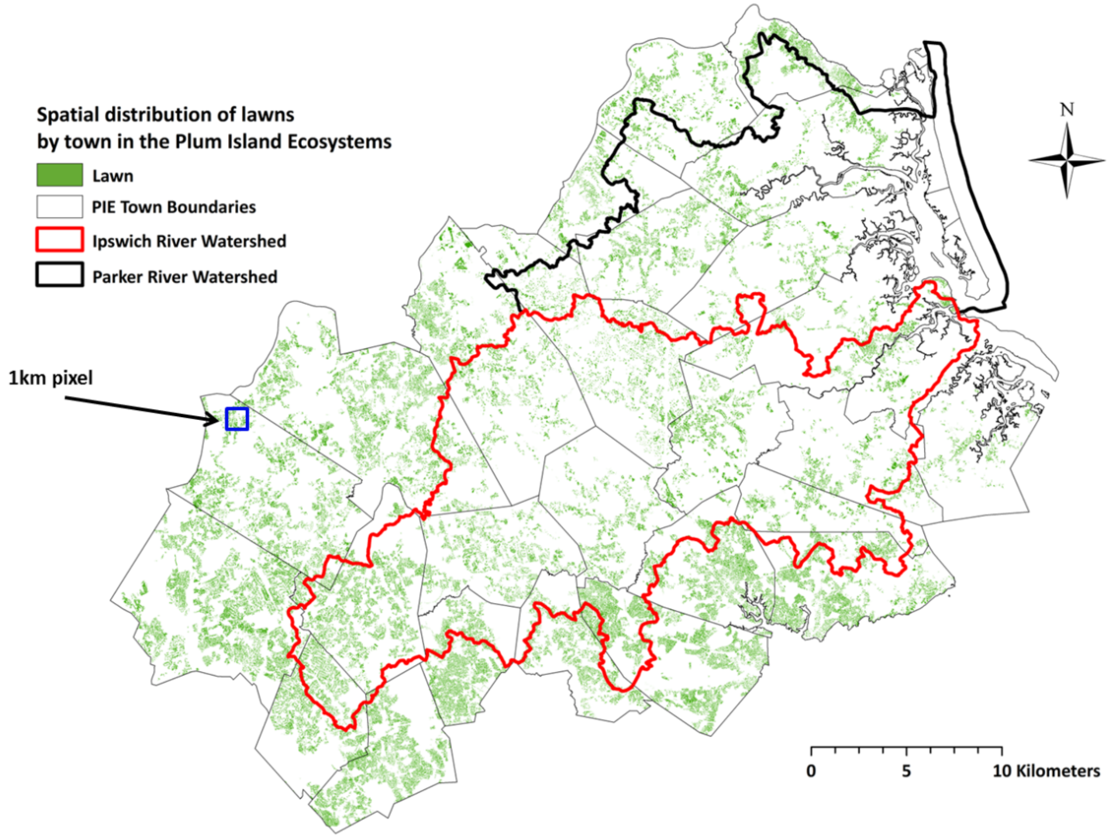


Figure 2. Lawn cover in the Plum Island Ecosystems LTER study area, 0.5-meter spatial resolution. The 1km pixel in blue reflects the spatial resolution of some prior work by other scholars (e.g., Milesi, 2005, 2009). Lawns cover roughly 7% of the total 1143 km^2 study area.

Results

The results of our 1143 km^2 mapping effort are presented in Table 1 and Figures 3-37. Table 1 shows the percentages of each land-cover category by town in the PIE study area. In the entire PIE study area, the dominant land-cover category is deciduous trees (35% of the study area), followed by wetlands (18%), impervious surfaces (14%), coniferous trees and fine green grass (13% each), bare soil (4%), and water (3%). Lawns cover roughly 7% of the PIE study area, and account for over half of all fine green grass. Figures 3-10 show the percentages of each PIE town composed of each individual land-cover category, listed in descending order. The town of Danvers has the highest percentage of bare soil (nearly 13%), while the highest percentage of coniferous trees were in Boxford (21%). The most deciduous trees were observed in West Newbury (45%), while the highest percentage of impervious surfaces occurred in Woburn (35%). West Newbury also had the highest percentage of water (nearly 9%), while the highest percentage of wetlands was observed in Newbury (nearly 40%). The highest percentages of grass and lawn were observed in Peabody, accounting for 19% and nearly 12% of the town, respectively. Figures 11-37 show the fine-resolution land-cover maps for the entire PIE study area and each study area town. Each town land-cover map is accompanied by a pie chart, which shows the land-cover percentages (including lawn) of each town.

The results of the accuracy assessment of the land-cover maps are presented in Tables 2-28 and Figures 38-64. The tables show population-estimated error matrices for the entire PIE study area and each town, with each matrix entry in percent of the study area. The on-diagonal entries (bold) in each matrix show the percentage of each land-cover category on the landscape in which there is agreement between the reference map and the land-cover classification. Figures 38-64 show visual representations of the information in each error matrix, where the three segments of each bar represent omission disagreement, agreement, and commission disagreement for each land-cover category. The colored segments (and their corresponding numerical values) in each bar are analogous to the on-diagonal entries in each error matrix. Across the entire PIE study area, the population-estimated overall agreement between the reference image and land-cover classification is 80%, with the highest overall agreements occurring in Boxford and Rowley (90% agreement) and the lowest agreement occurring in Reading (68% agreement), a range of 22 percentage points.

For a full description of the concepts and methods used in the accuracy assessment, please refer to Giner (2013). Pontius and Millones (2011) outline the methods for the derivation and interpretation of the error matrices shown in Tables 2-28, while Pontius et al. (2008) describes the derivation and interpretation of Figures 38-64.

Table 1. Land-cover percentages by town in the PIE study area.

Town	Bare soil	Coniferous	Deciduous	Grass	Lawn	Impervious	Water	Wetlands	Total
Andover	1.32	12.58	42.85	12.00	7.12	16.68	3.79	10.79	100
Beverly	1.83	14.99	32.50	16.96	10.39	23.75	2.46	7.51	100
Billerica	1.38	14.30	33.96	15.44	10.04	20.13	2.92	11.88	100
Boxford	2.02	20.99	43.61	8.24	4.98	6.64	2.81	15.68	100
Burlington	5.84	13.25	26.77	14.62	9.77	28.34	1.11	10.06	100
Danvers	12.52	6.34	27.19	16.04	10.31	28.61	4.96	4.35	100
Essex	3.5	16.00	34.20	6.25	1.84	4.63	1.52	33.90	100
Georgetown	2.49	17.06	42.88	9.89	6.45	9.80	1.67	16.23	100
Groveland	6.23	13.19	42.15	7.51	5.22	9.77	4.68	16.48	100
Hamilton	4.22	15.59	33.63	13.31	4.07	7.06	2.95	23.23	100
Ipswich	7.81	9.23	34.23	9.39	3.06	6.10	2.42	30.82	100
Lynnfield	2.71	17.69	29.88	11.82	7.04	15.45	4.31	18.14	100
Middleton	2.16	17.07	35.30	11.32	4.29	10.24	5.46	18.45	100
Newbury	6.17	9.61	28.75	9.78	4.28	5.39	1.14	39.15	100
Newburyport	8.95	11.02	26.53	16.99	7.29	20.91	1.29	14.30	100
North Andover	0.79	9.52	43.78	13.22	5.67	13.75	4.76	14.19	100
North Reading	2.81	18.22	30.88	10.75	6.95	16.32	2.86	18.16	100
Peabody	3.21	6.27	29.86	18.78	11.51	31.48	4.19	6.21	100
Reading	4.02	10.55	21.77	16.60	11.31	21.06	0.18	25.82	100
Rowley	6.03	13.07	36.31	6.75	3.37	6.15	0.82	30.88	100
Tewksbury	2.73	13.53	28.90	17.44	10.51	17.31	1.86	18.22	100
Topsfield	1.10	15.03	38.41	16.56	7.53	8.11	2.22	18.57	100
Wenham	0.56	15.20	32.65	13.70	5.55	7.95	5.74	24.20	100
West Newbury	3.86	6.45	44.96	15.68	6.77	5.35	8.40	15.29	100
Wilmington	1.06	14.12	29.45	13.57	10.15	22.23	1.29	18.29	100
Woburn	6.39	7.64	26.49	15.91	9.49	34.98	2.24	6.34	100
PIE study area	3.74	12.83	34.68	12.73	7.11	14.65	2.97	18.40	100

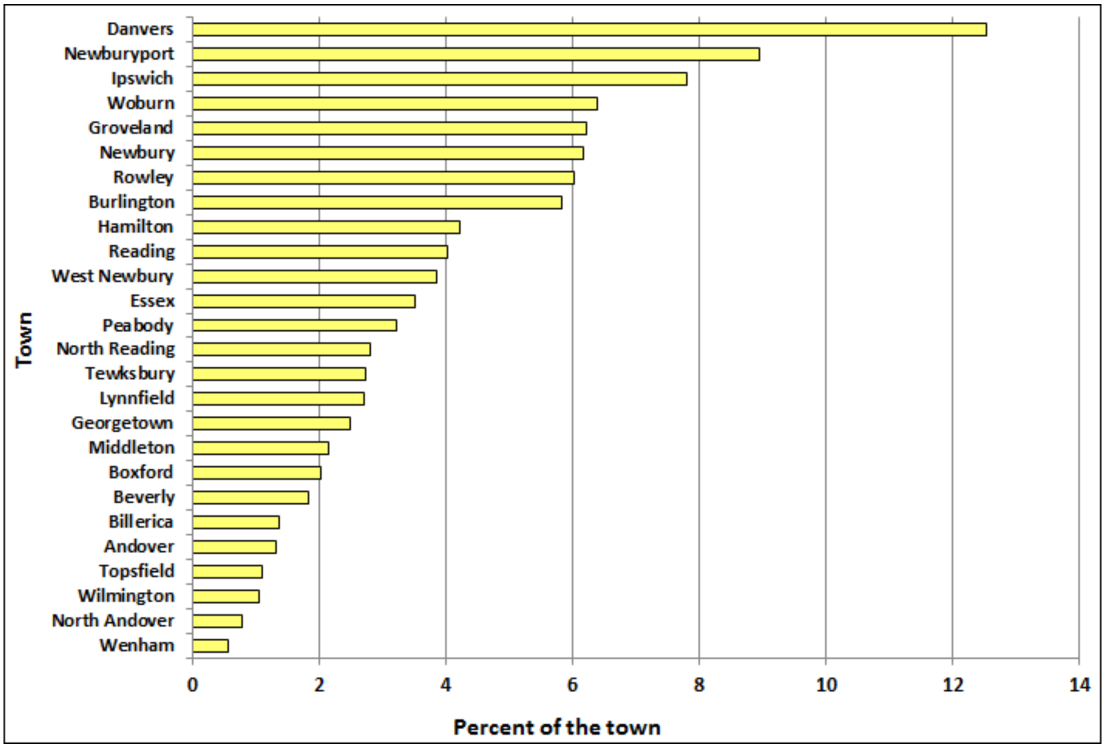


Figure 3. Percentage of each PIE town – bare soil.

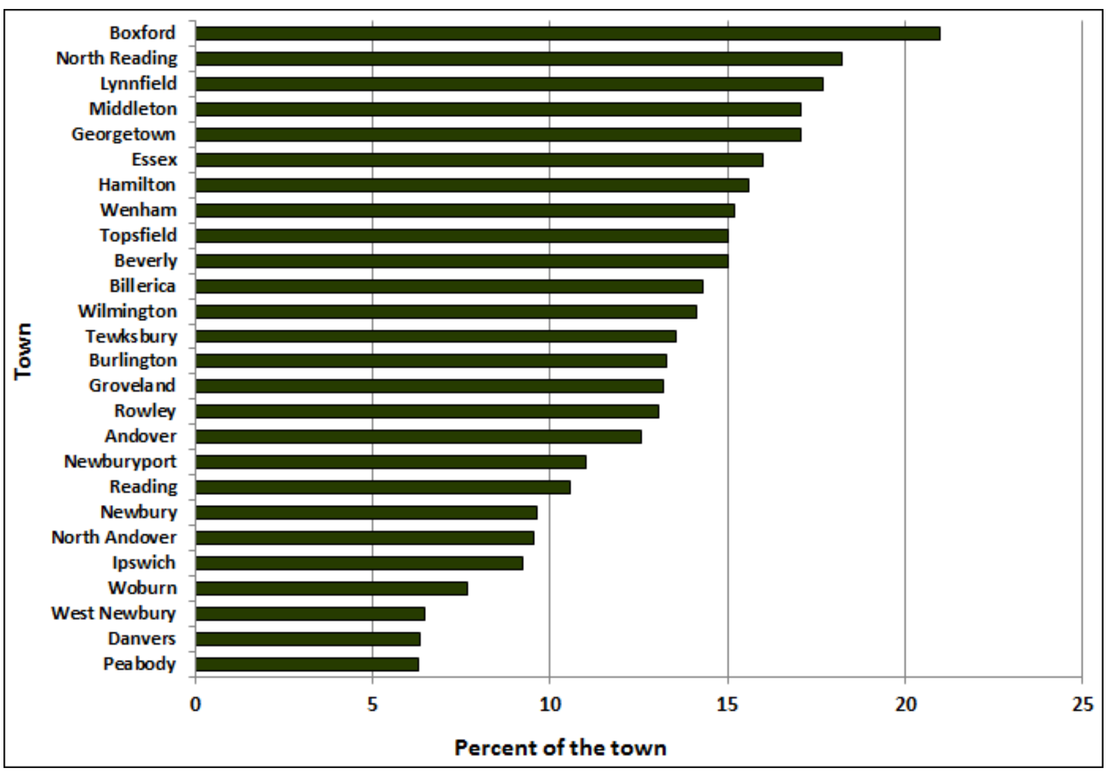


Figure 4. Percentage of each PIE town – coniferous trees.

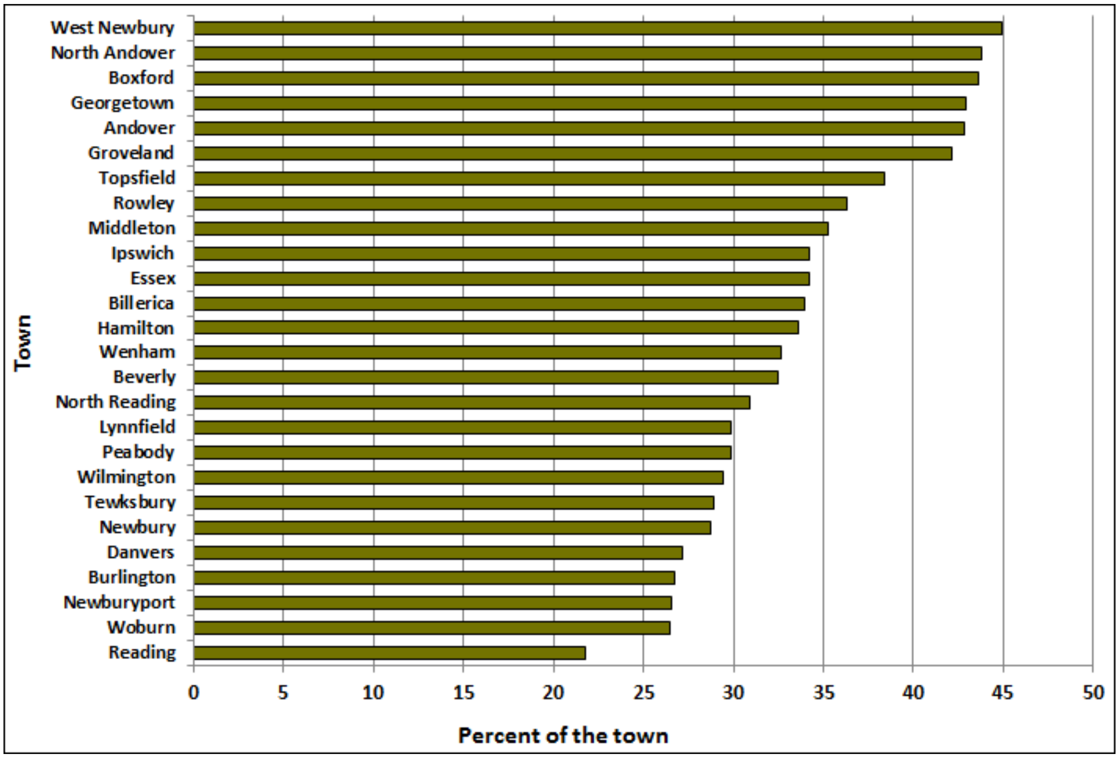


Figure 5. Percentage of each PIE town – deciduous trees.

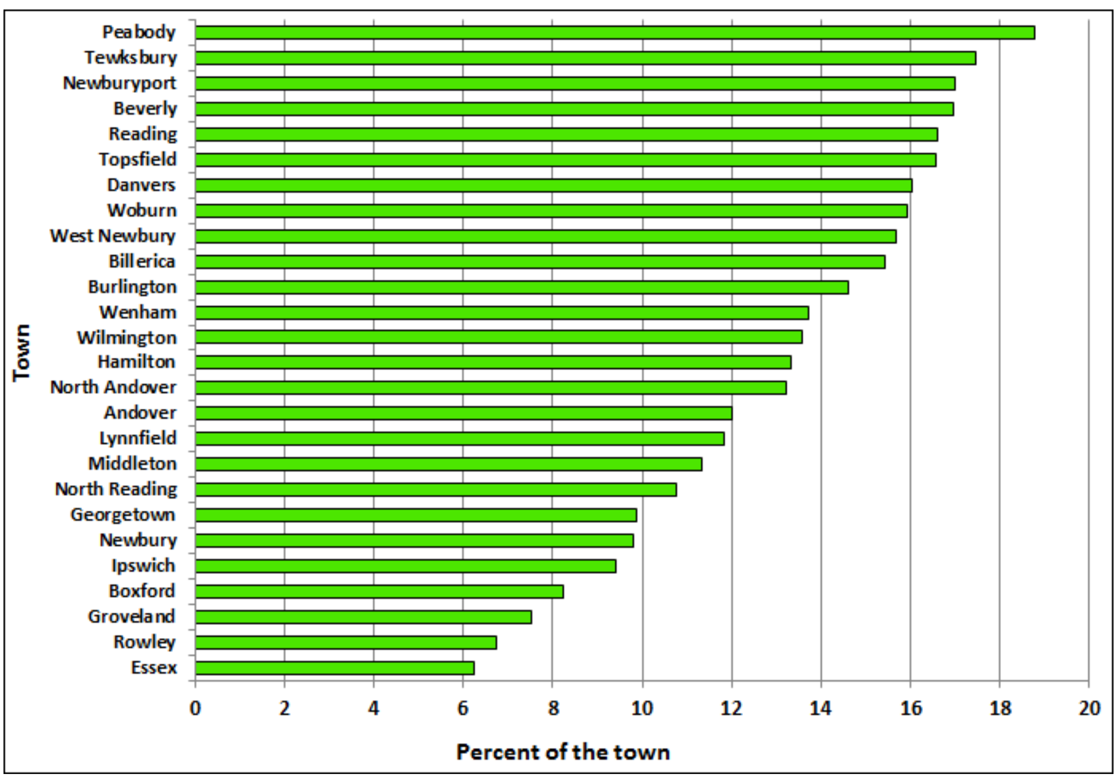


Figure 6. Percentage of each PIE town – fine green grass.

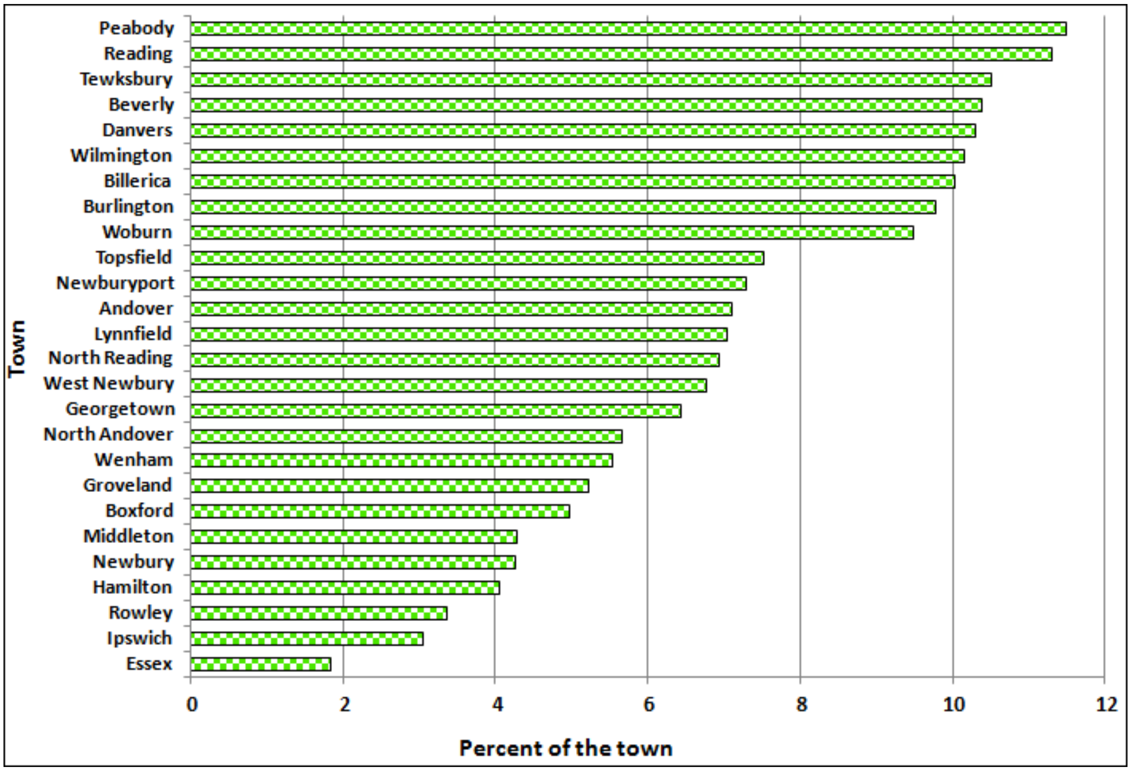


Figure 7. Percentage of each PIE town – lawn.

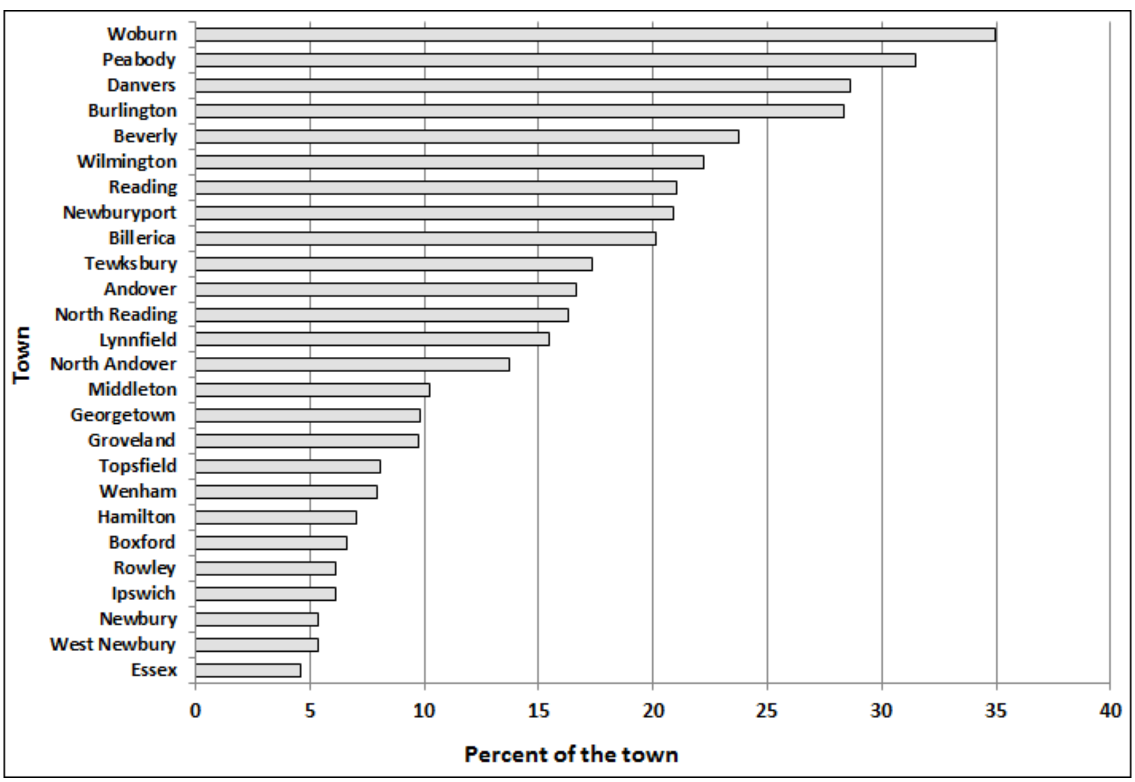


Figure 8. Percentage of each PIE town – impervious surfaces.

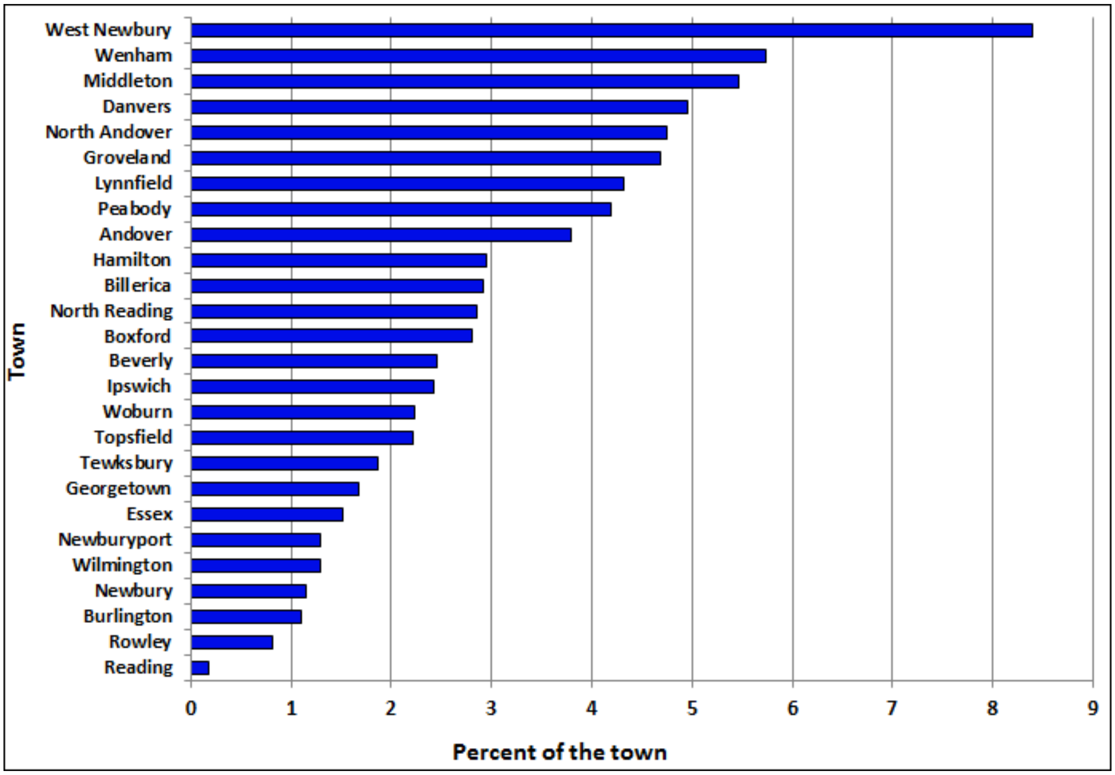


Figure 9. Percentage of each PIE town – water.

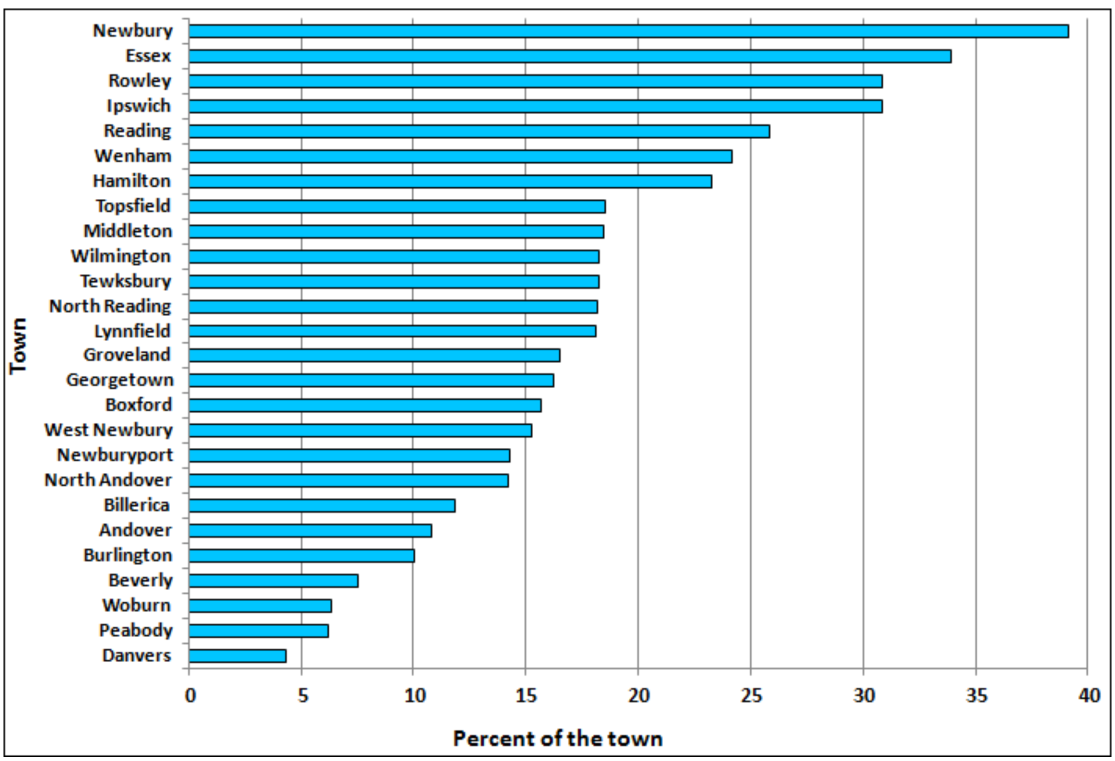
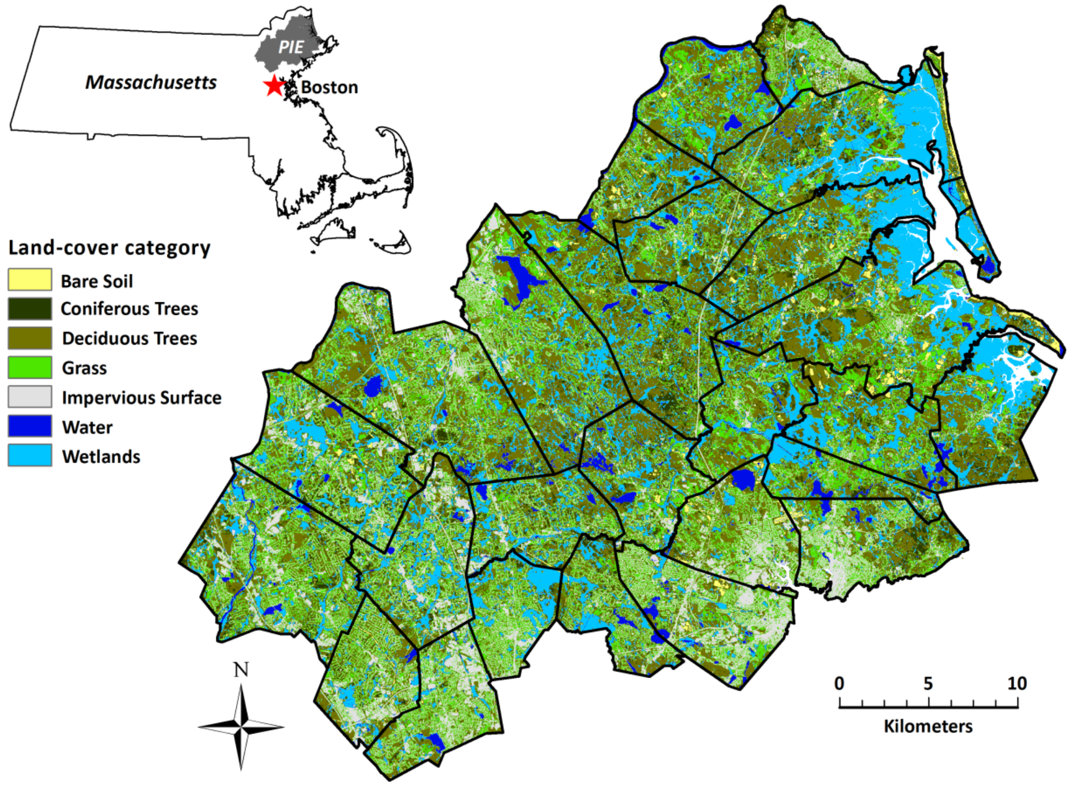


Figure 10. Percentage of each PIE town – wetlands.



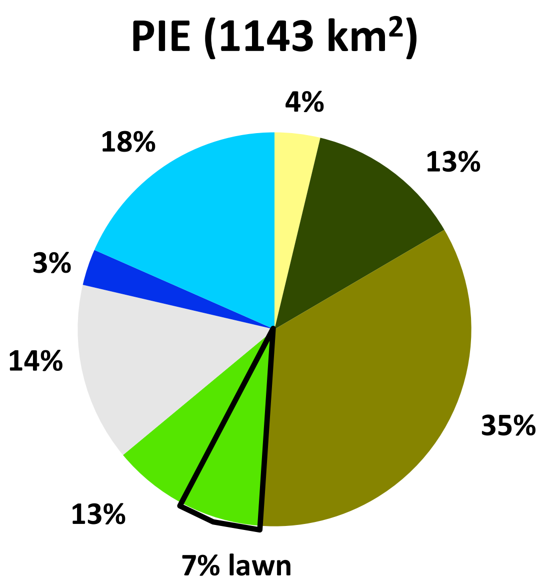
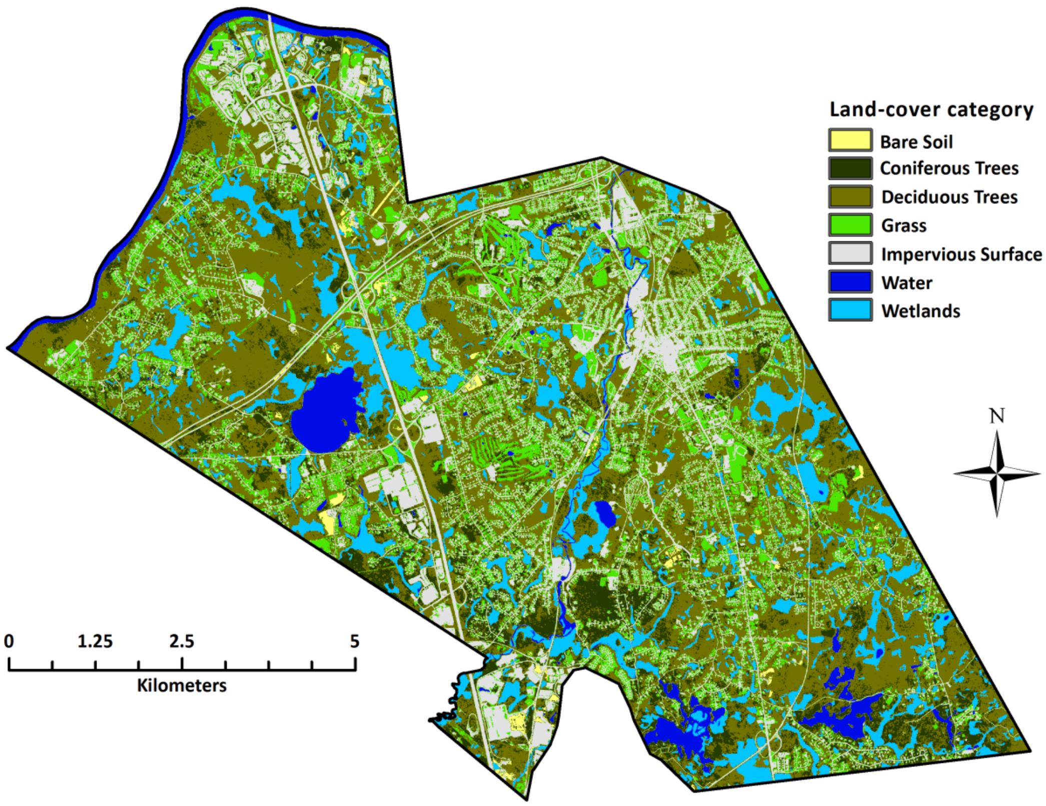


Figure 11. Land-cover map and corresponding land-cover pie chart for the PIE study area.



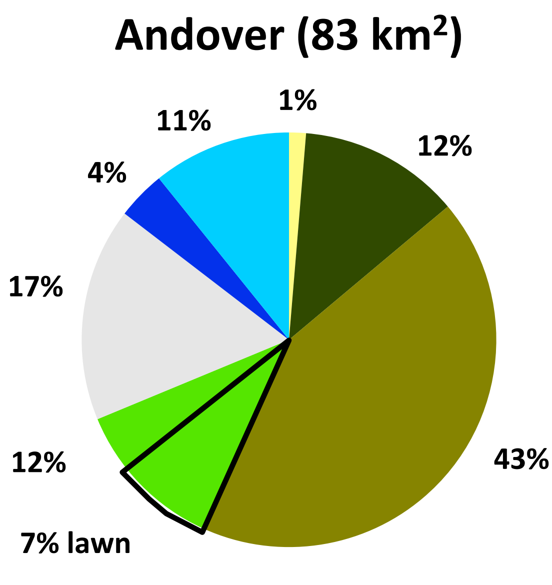
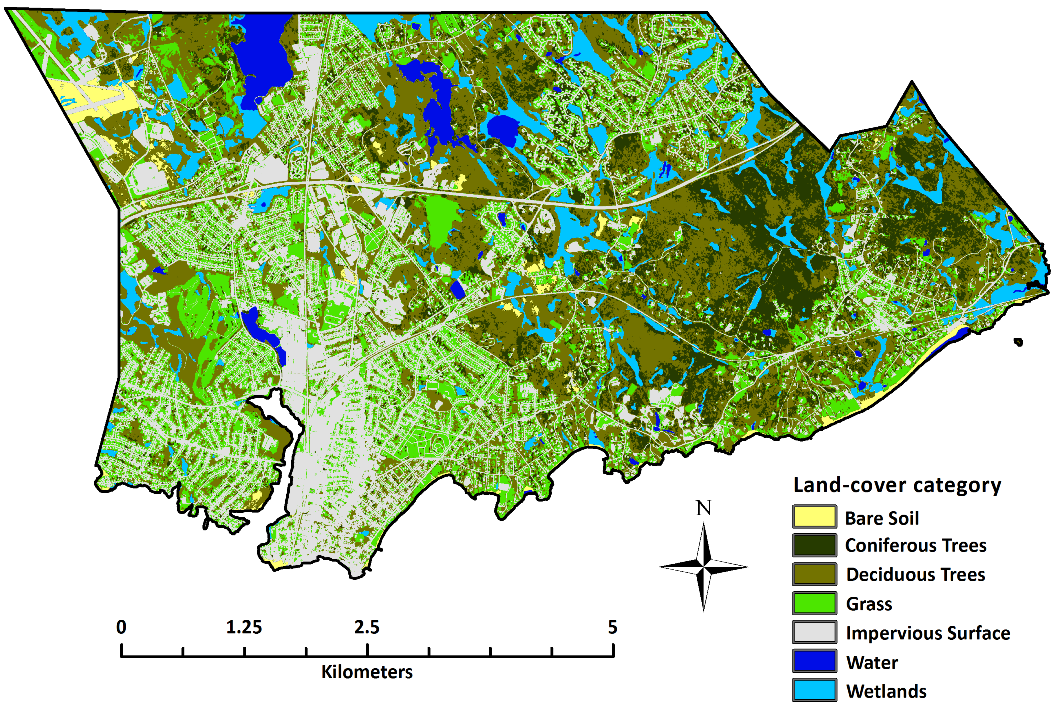


Figure 12. Land-cover map and corresponding land-cover pie chart for the town of Andover.



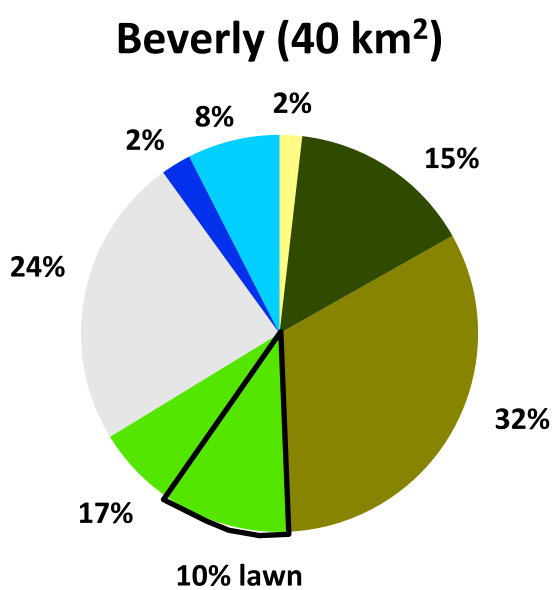
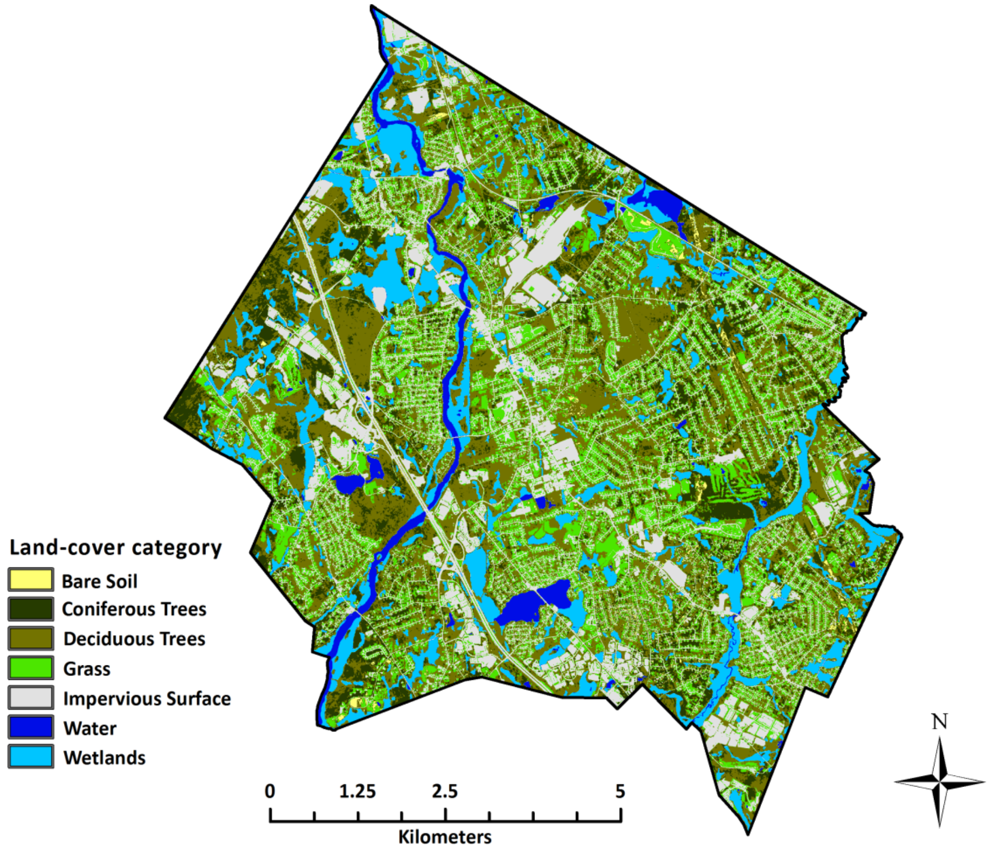


Figure 13. Land-cover map and corresponding land-cover pie chart for the town of Beverly.



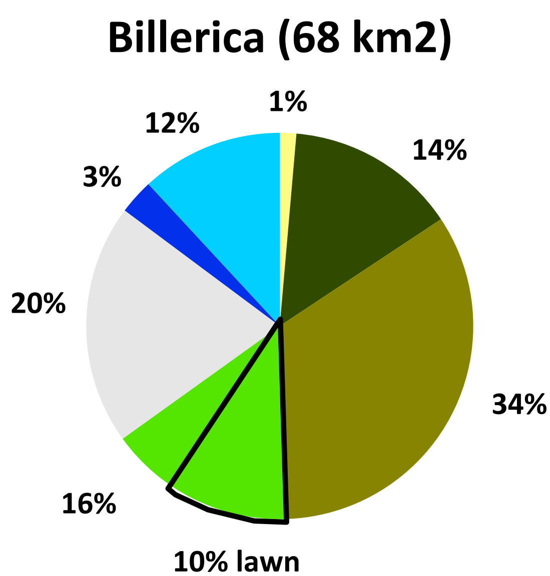
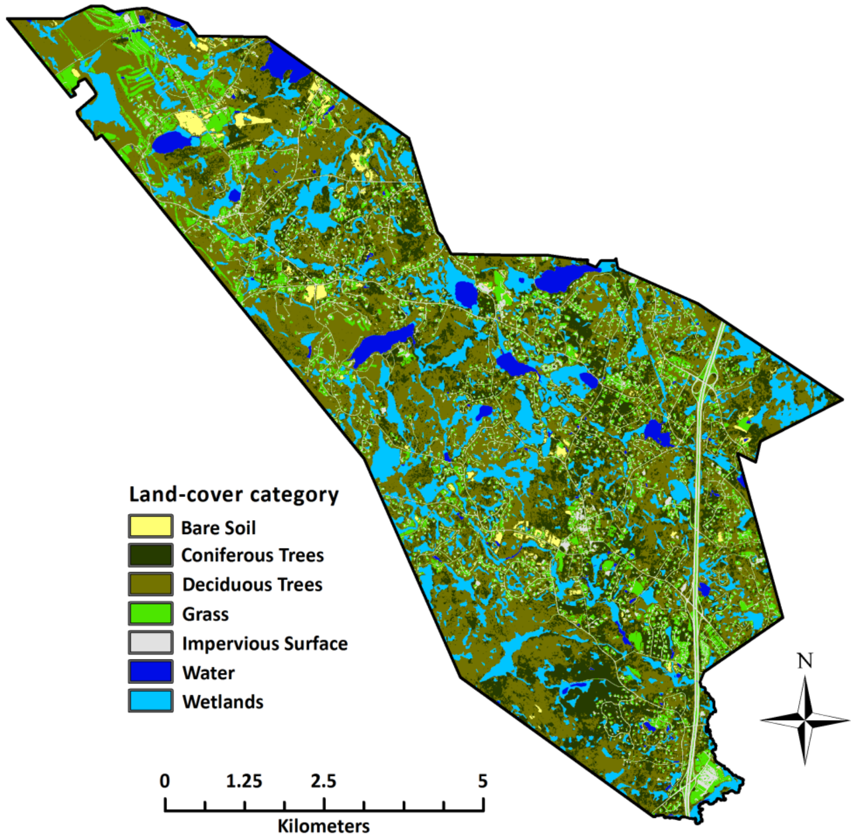


Figure 14. Land-cover map and corresponding land-cover pie chart for the town of Billerica.



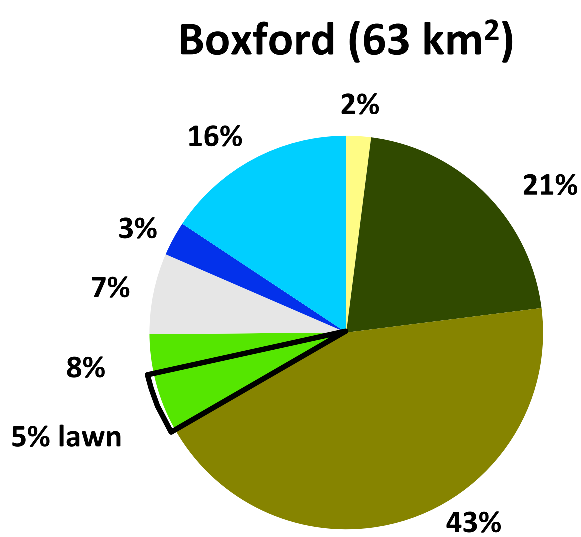


Figure 15. Land-cover map and corresponding land-cover pie chart for the town of Boxford.

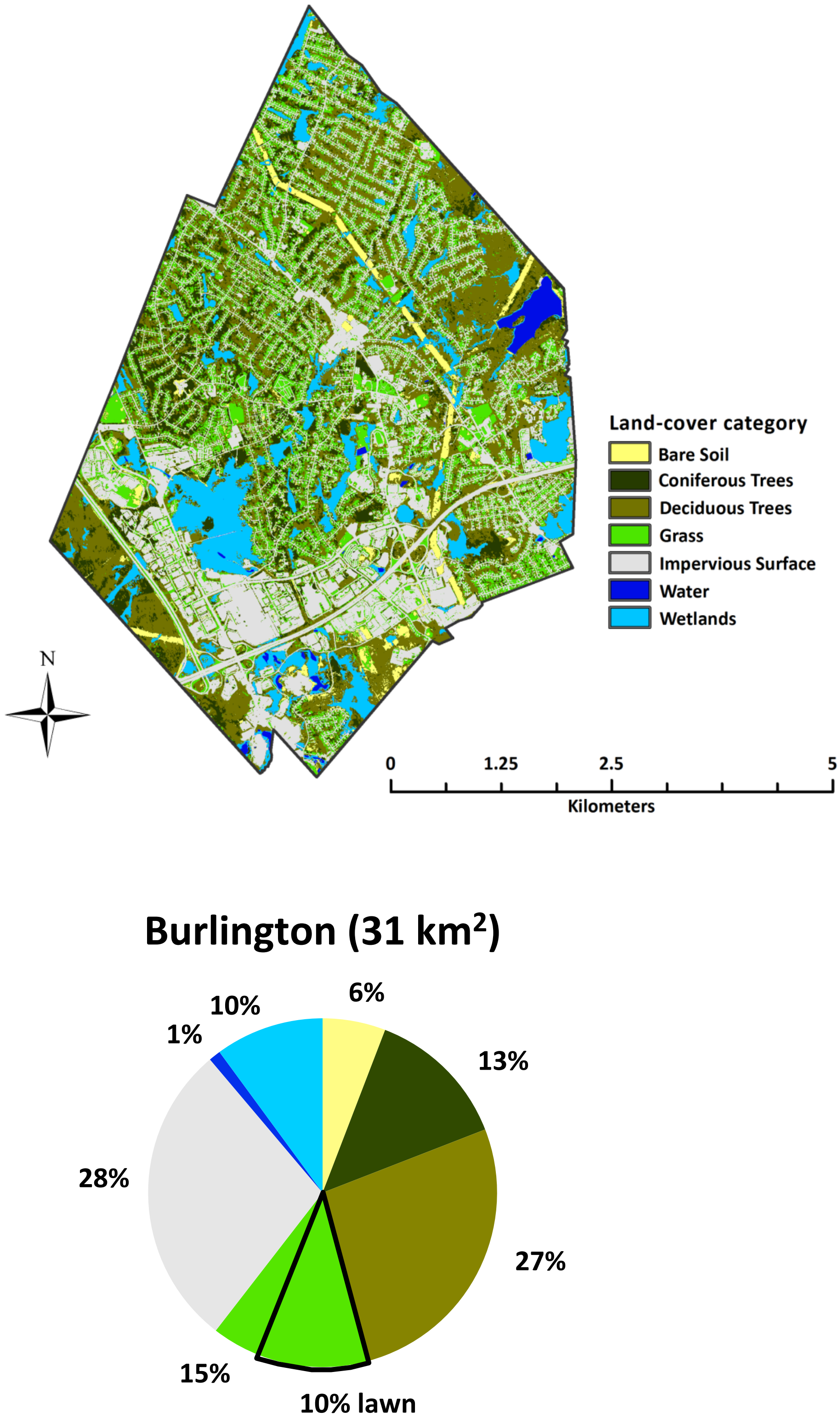
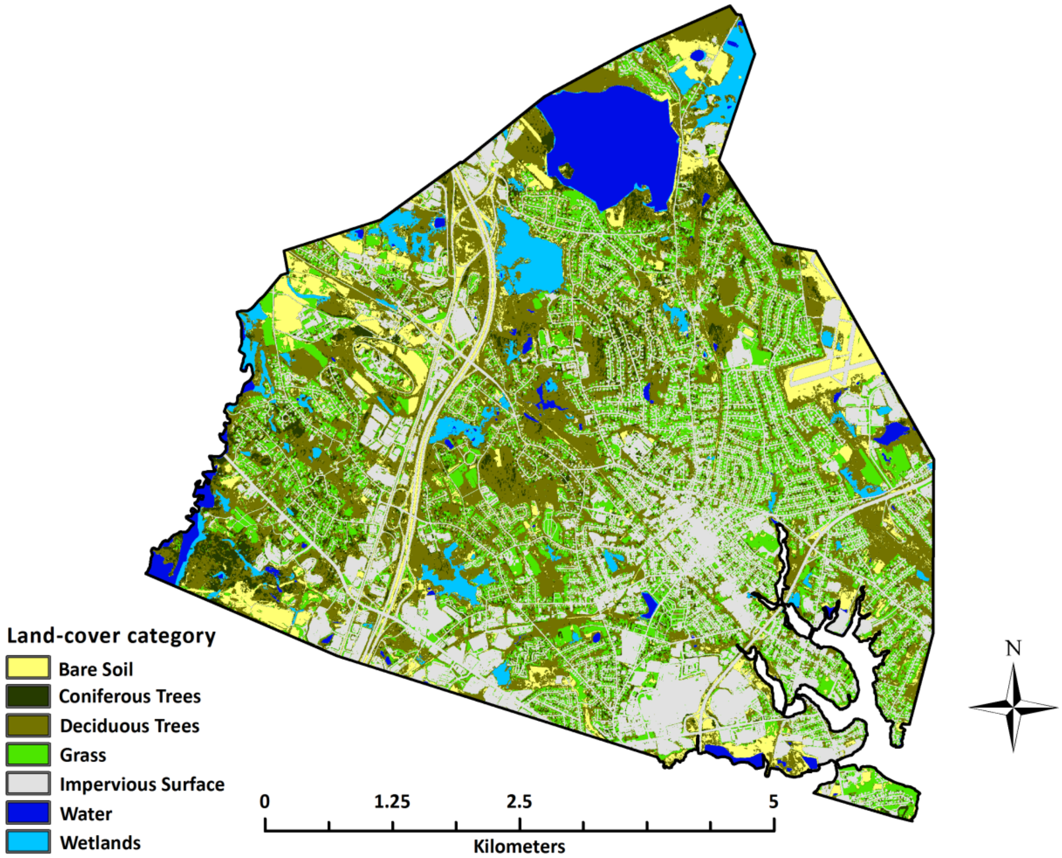


Figure 16. Land-cover map and corresponding land-cover pie chart for the town of Burlington.



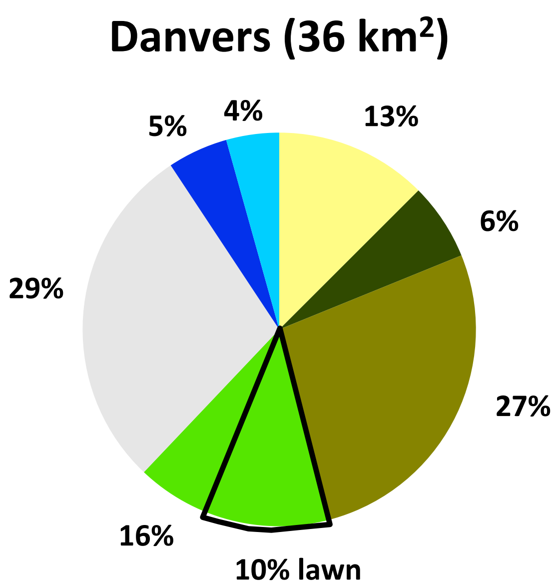
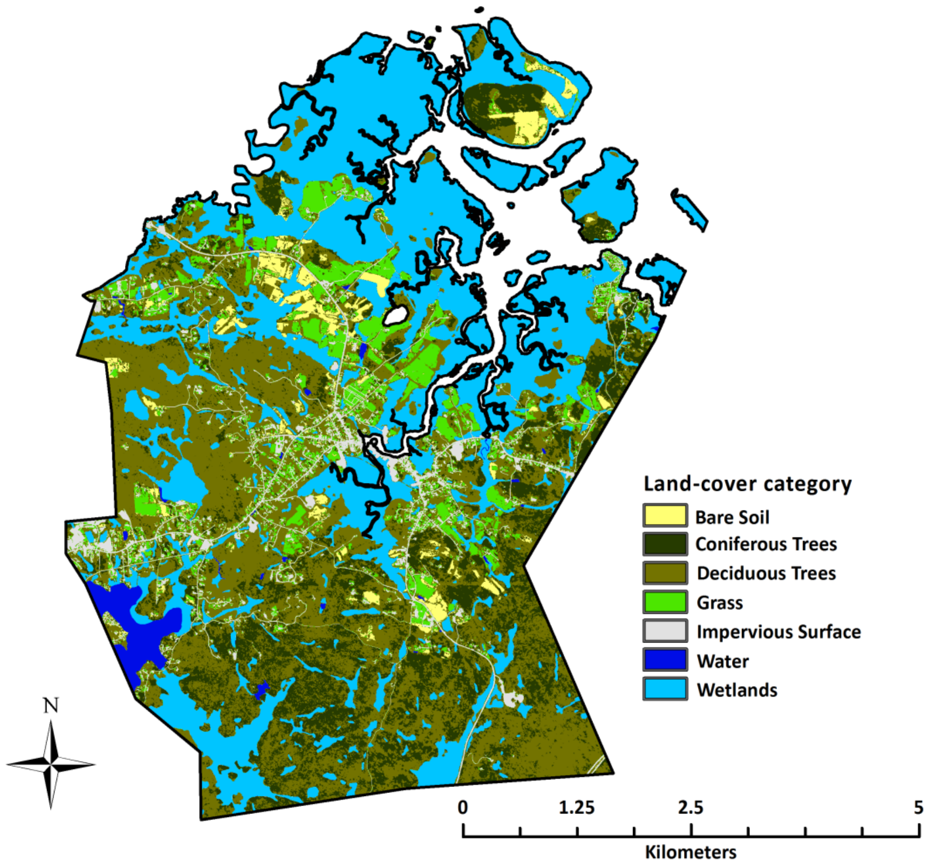


Figure 17. Land-cover map and corresponding land-cover pie chart for the town of Danvers.



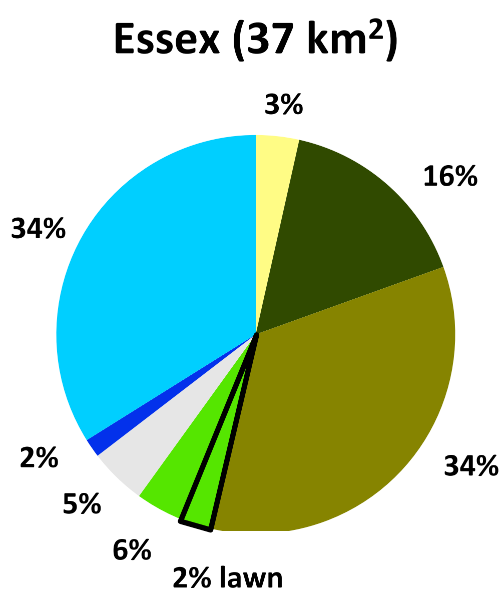
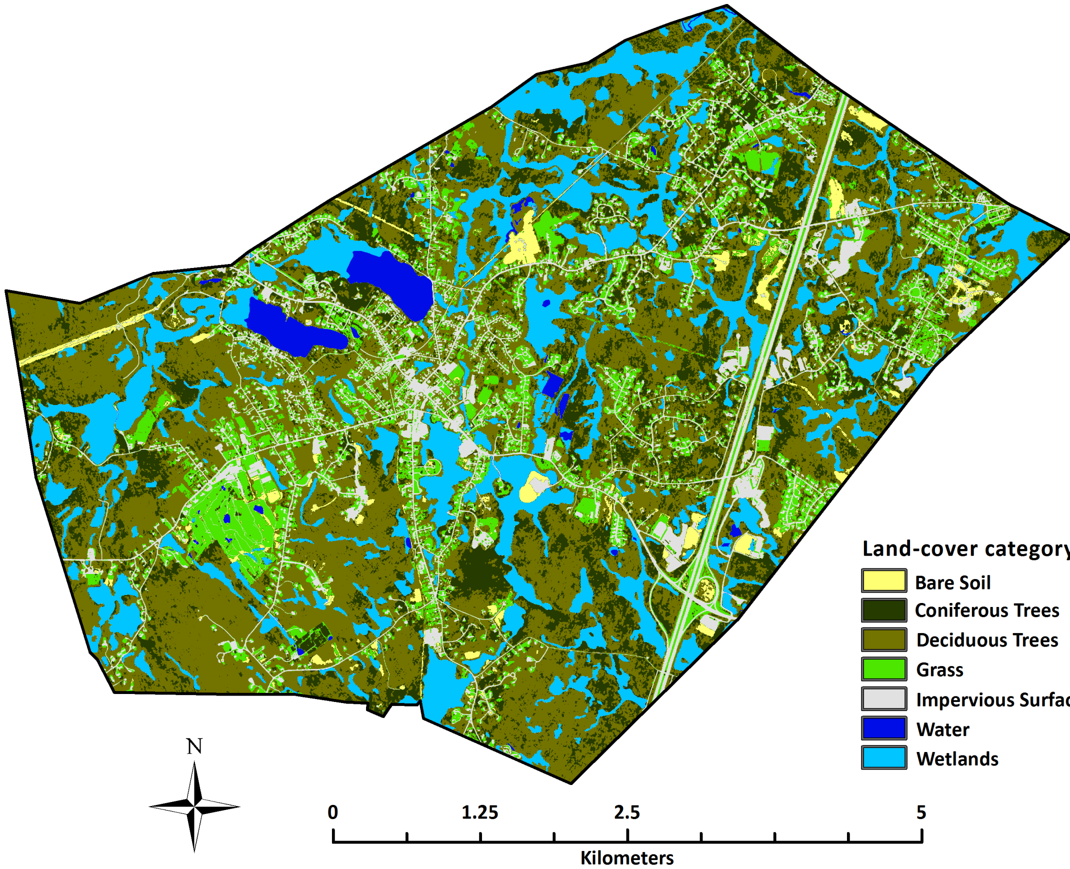


Figure 18. Land-cover map and corresponding land-cover pie chart for the town of Essex.



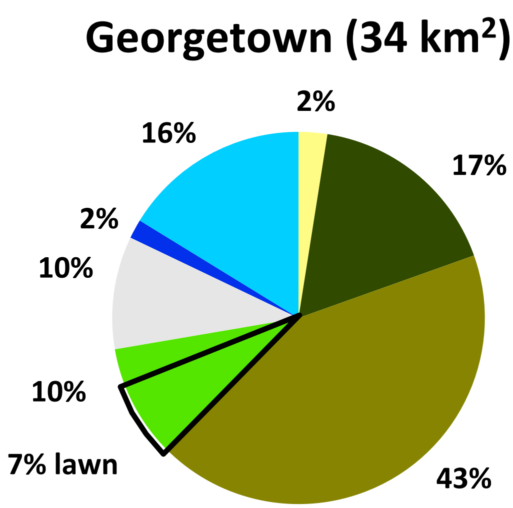
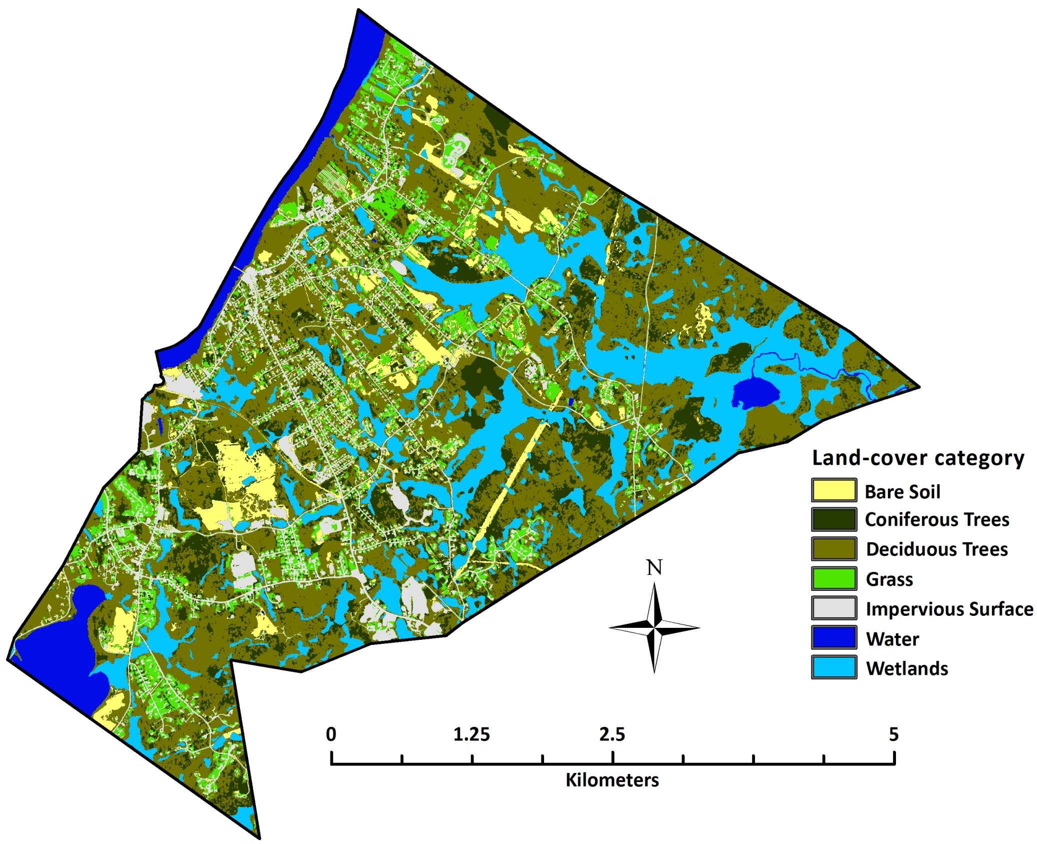


Figure 19. Land-cover map and corresponding land-cover pie chart for the town of Georgetown.



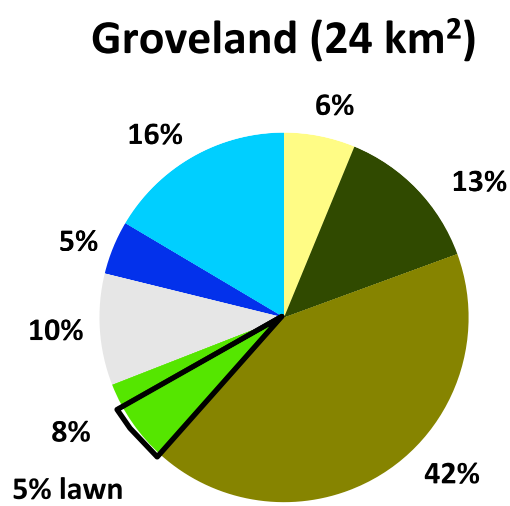
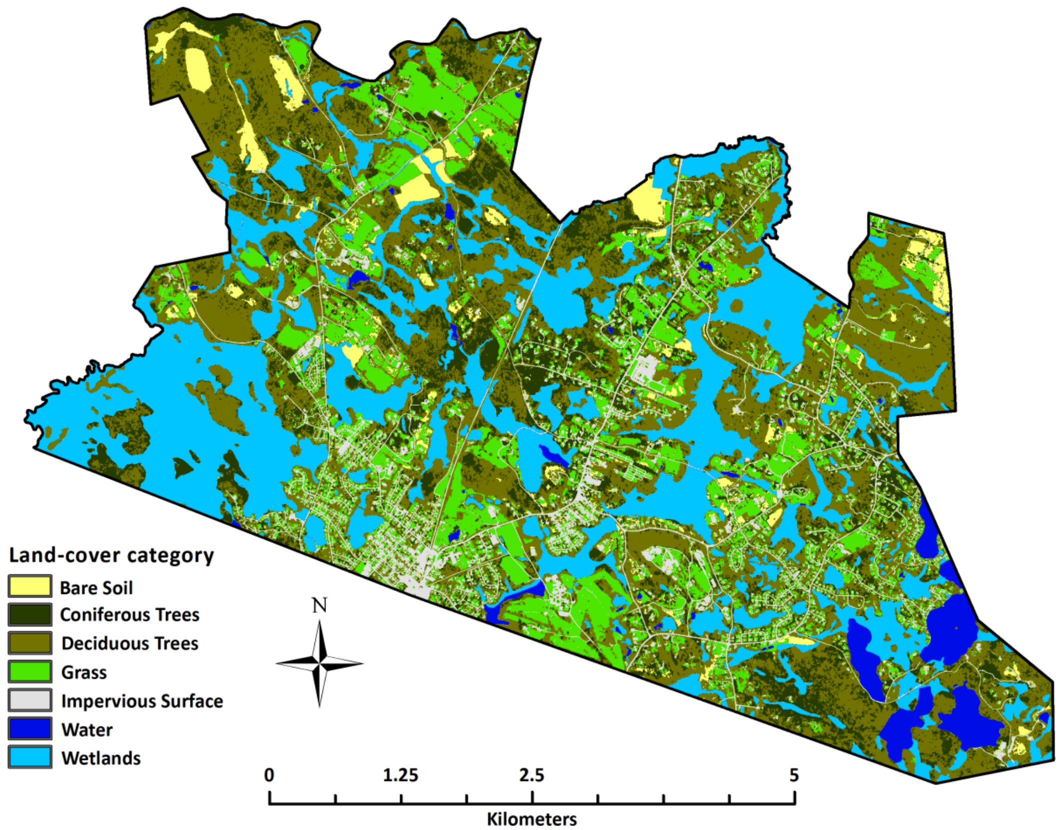


Figure 20. Land-cover map and corresponding land-cover pie chart for the town of Groveland.



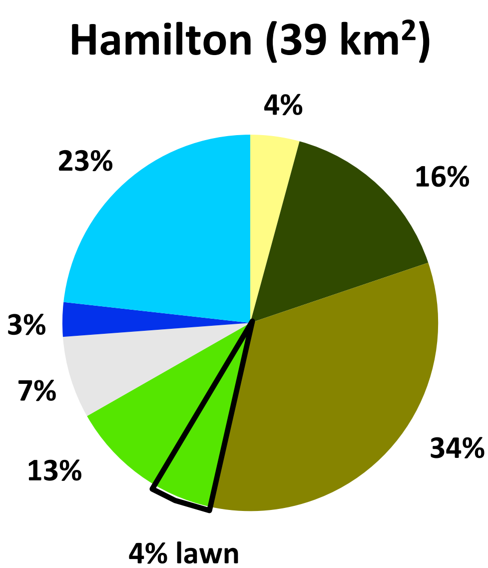
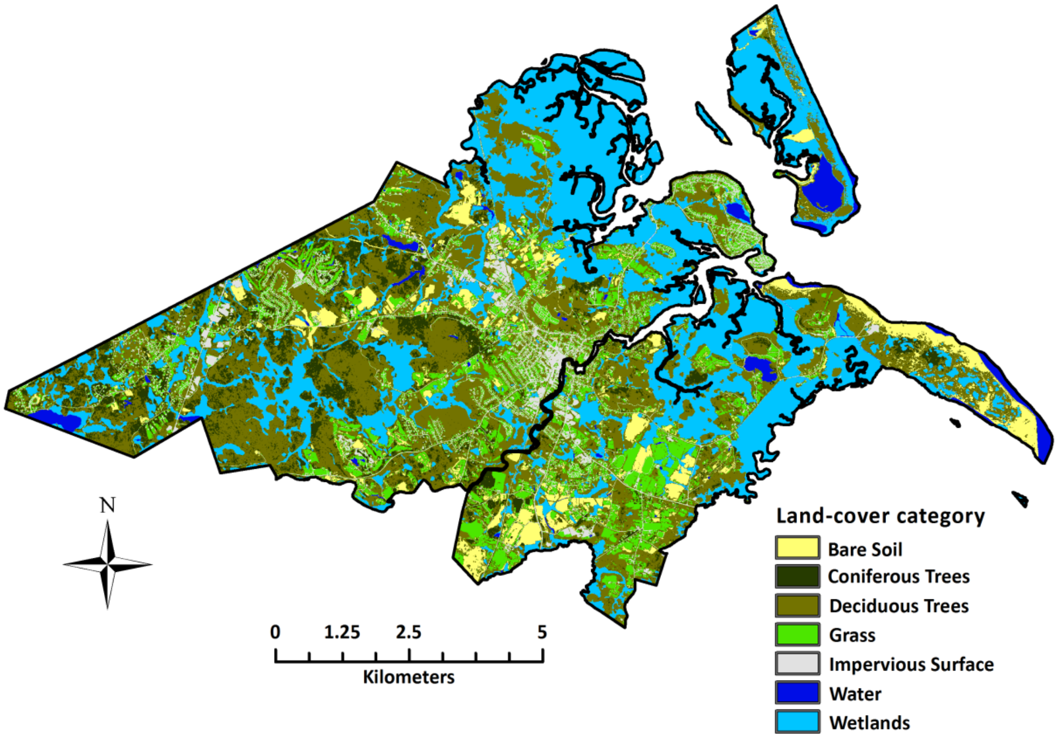


Figure 21. Land-cover map and corresponding land-cover pie chart for the town of Hamilton.



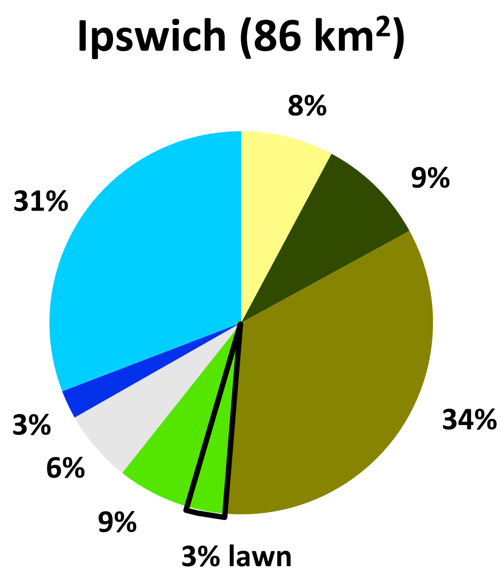
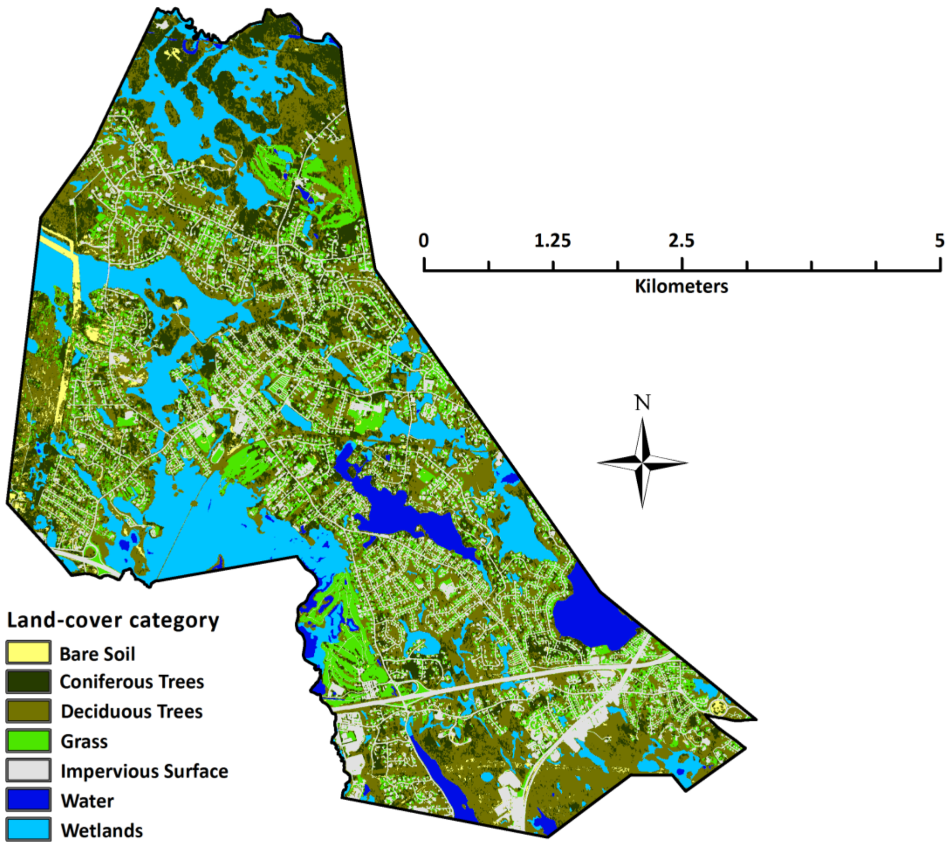


Figure 22. Land-cover map and corresponding land-cover pie chart for the town of Ipswich.



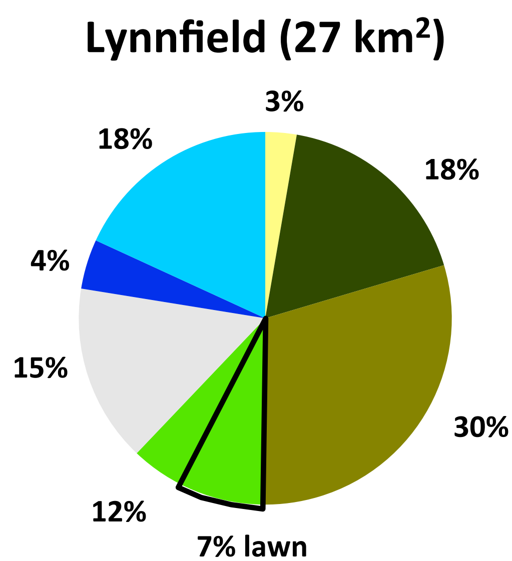
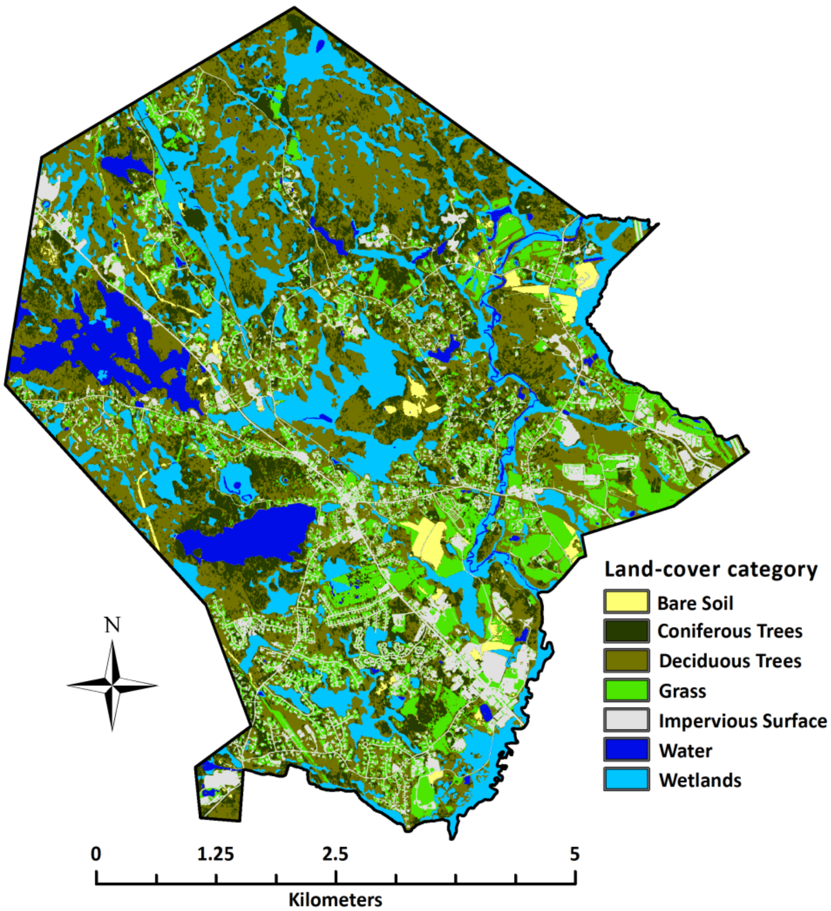


Figure 23. Land-cover map and corresponding land-cover pie chart for the town of Lynnfield.



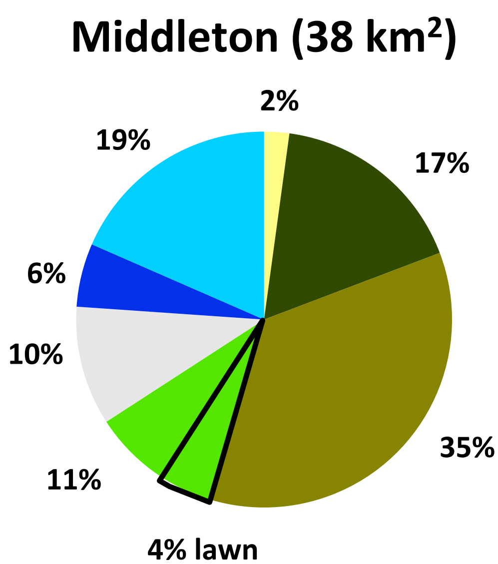
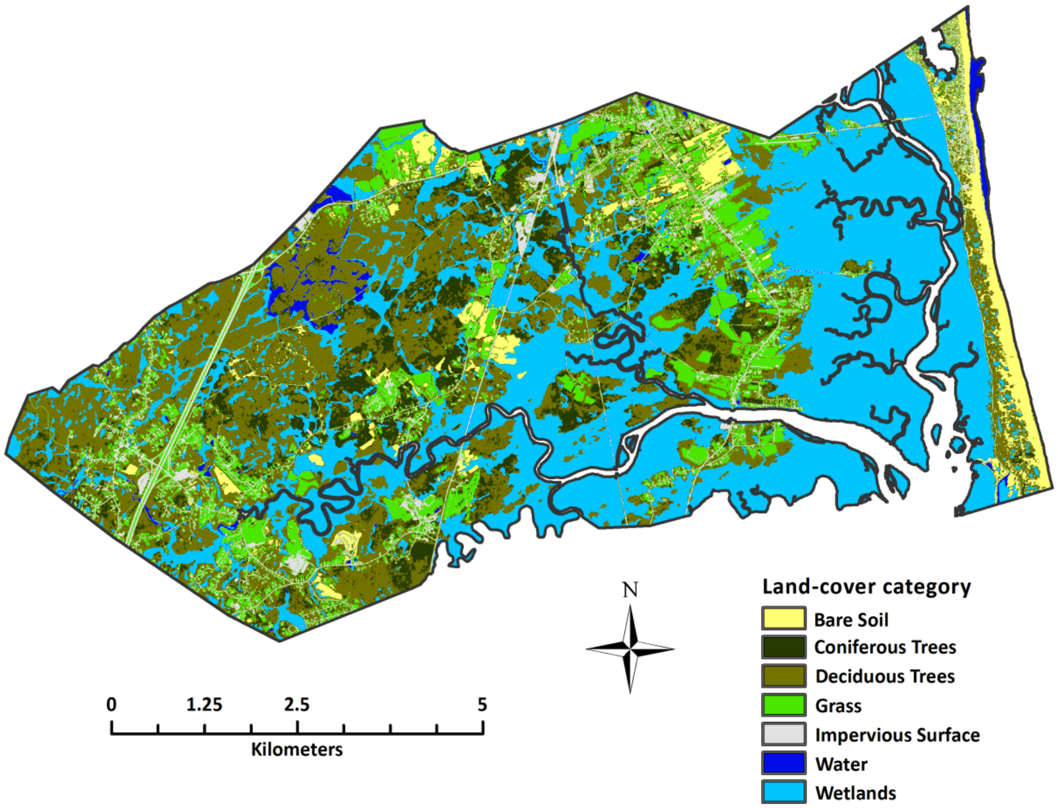


Figure 24. Land-cover map and corresponding land-cover pie chart for the town of Middleton.



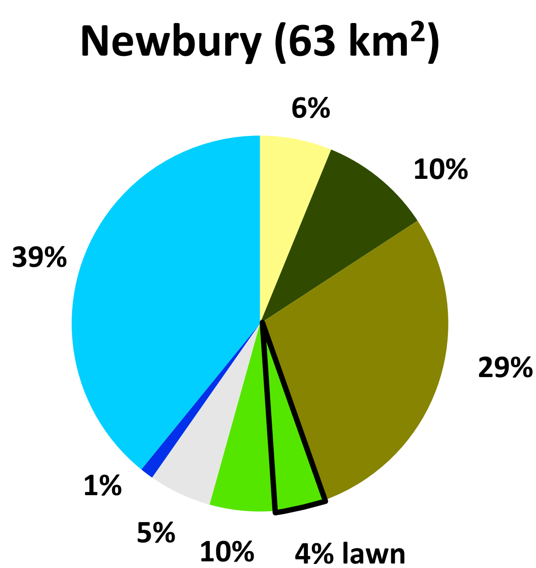
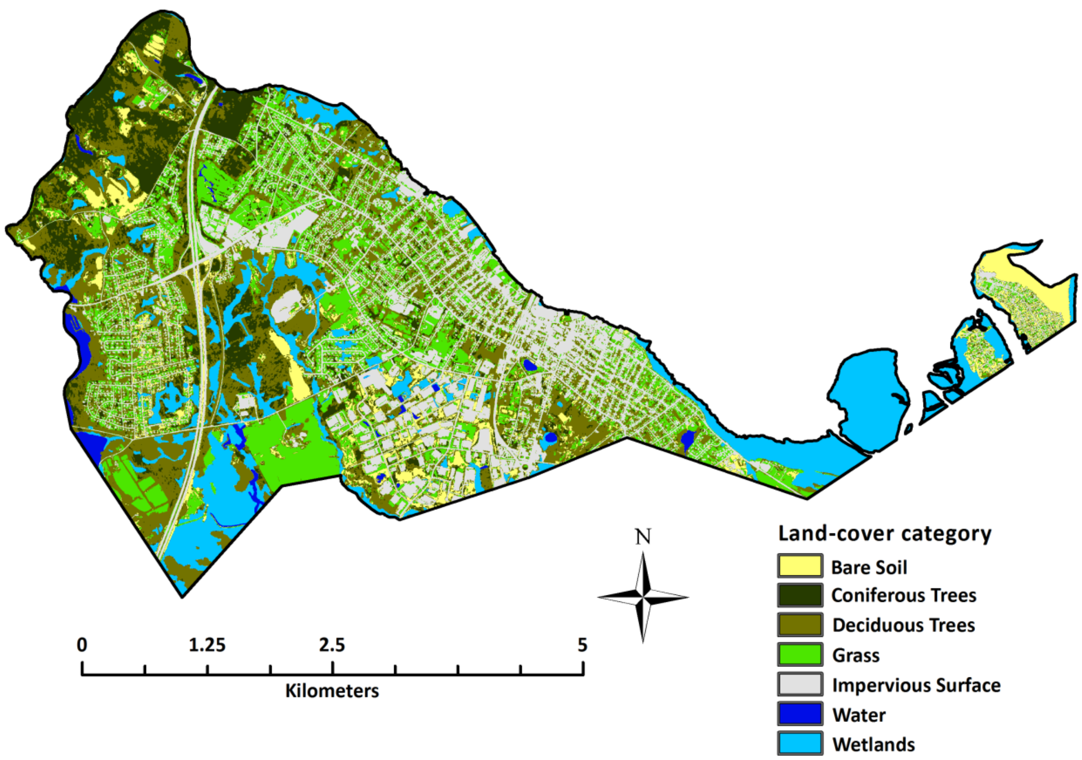


Figure 25. Land-cover map and corresponding land-cover pie chart for the town of Newbury.



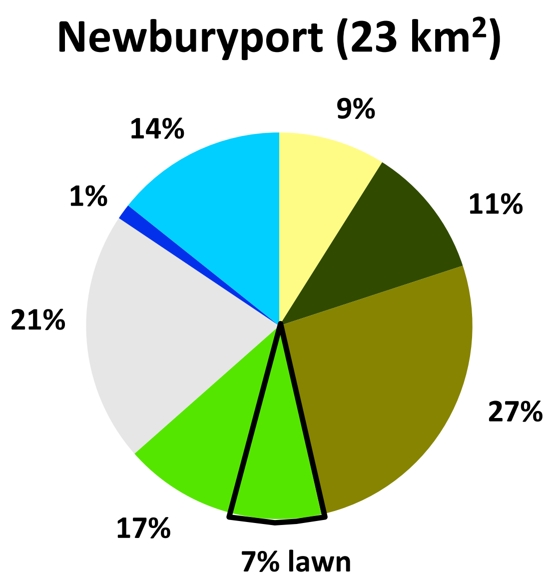
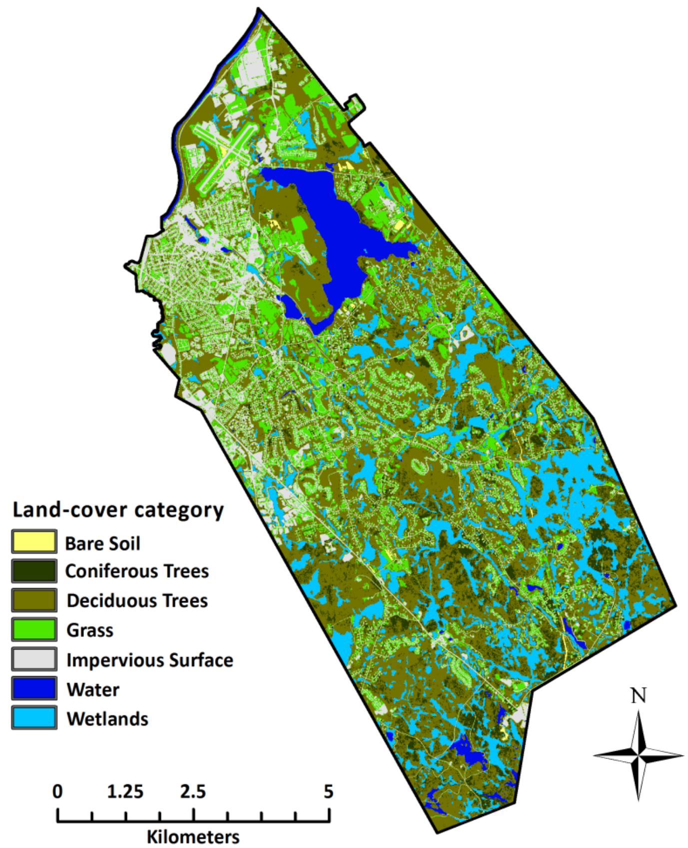


Figure 26. Land-cover map and corresponding land-cover pie chart for the town of Newburyport.



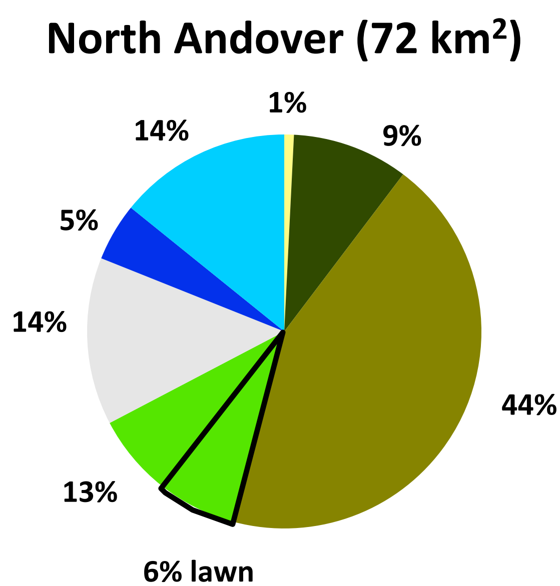
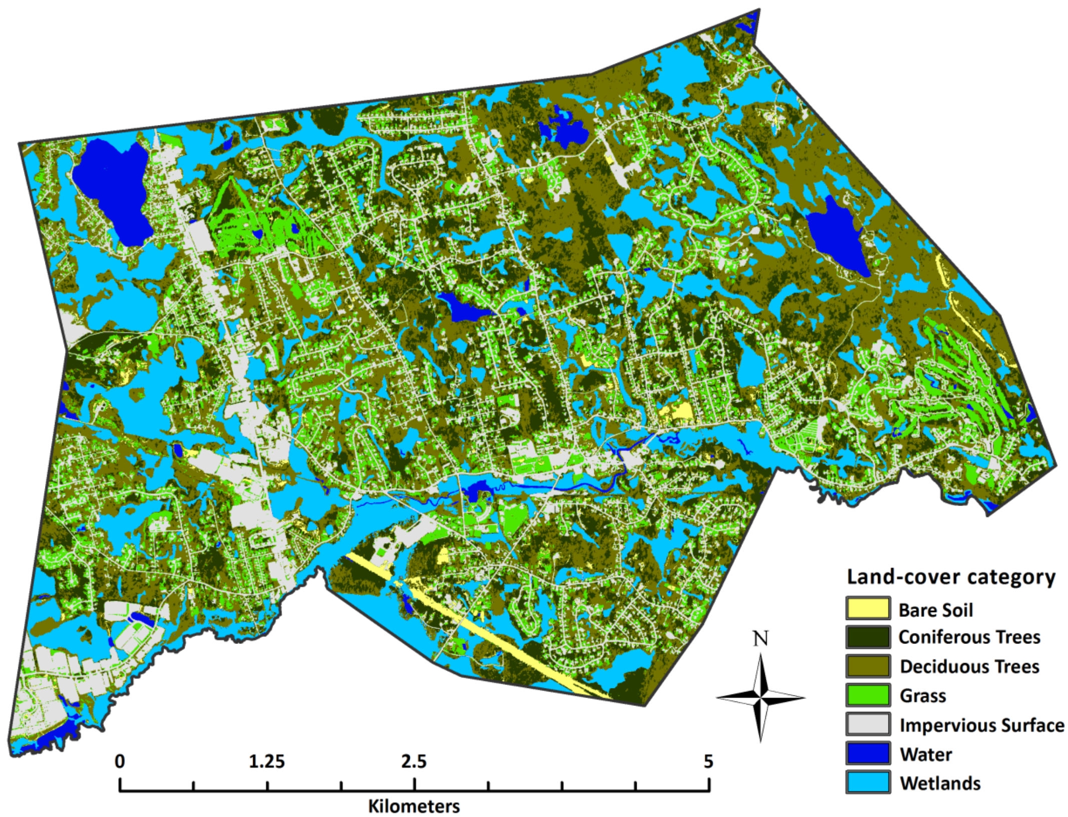


Figure 27. Land-cover map and corresponding land-cover pie chart for the town of North Andover.



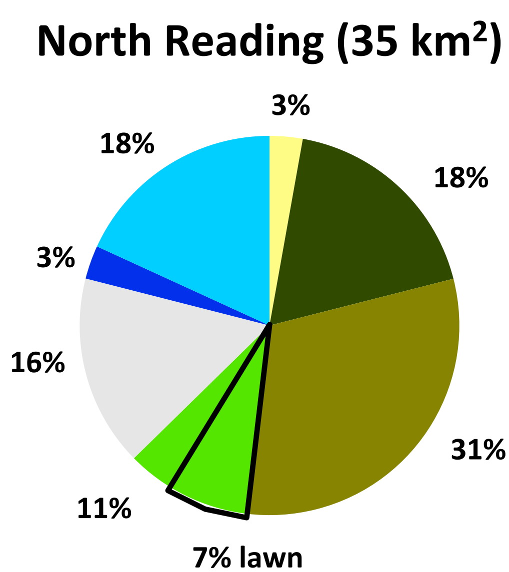
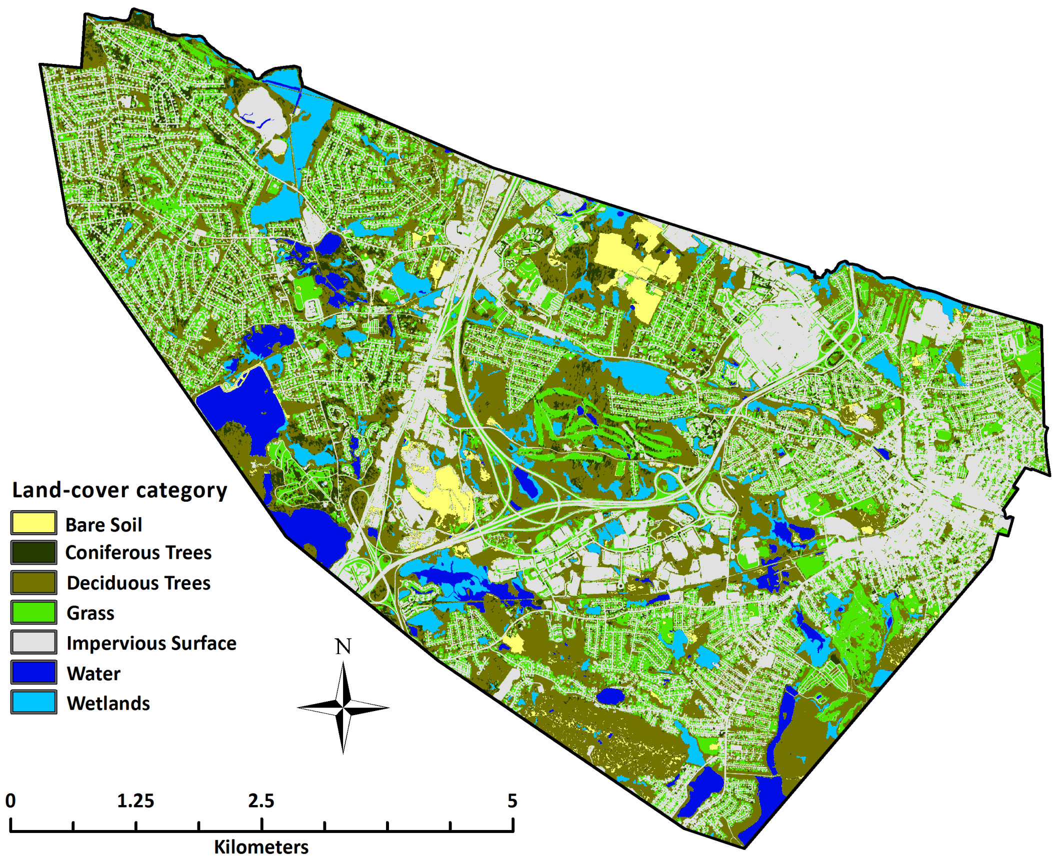


Figure 28. Land-cover map and corresponding land-cover pie chart for the town of North Reading.



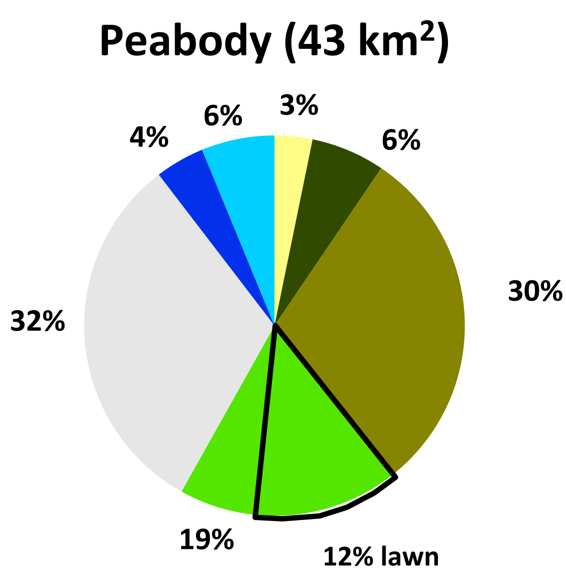
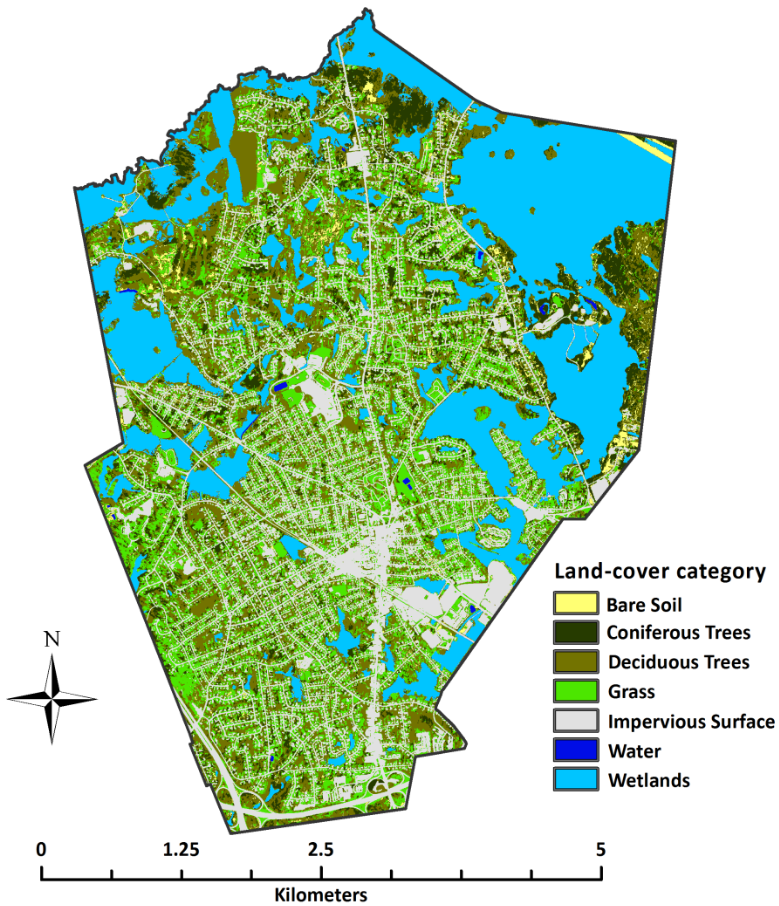


Figure 29. Land-cover map and corresponding land-cover pie chart for the town of Peabody.



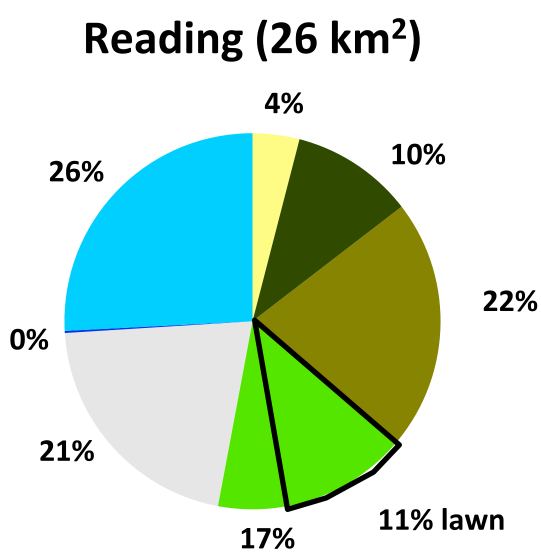
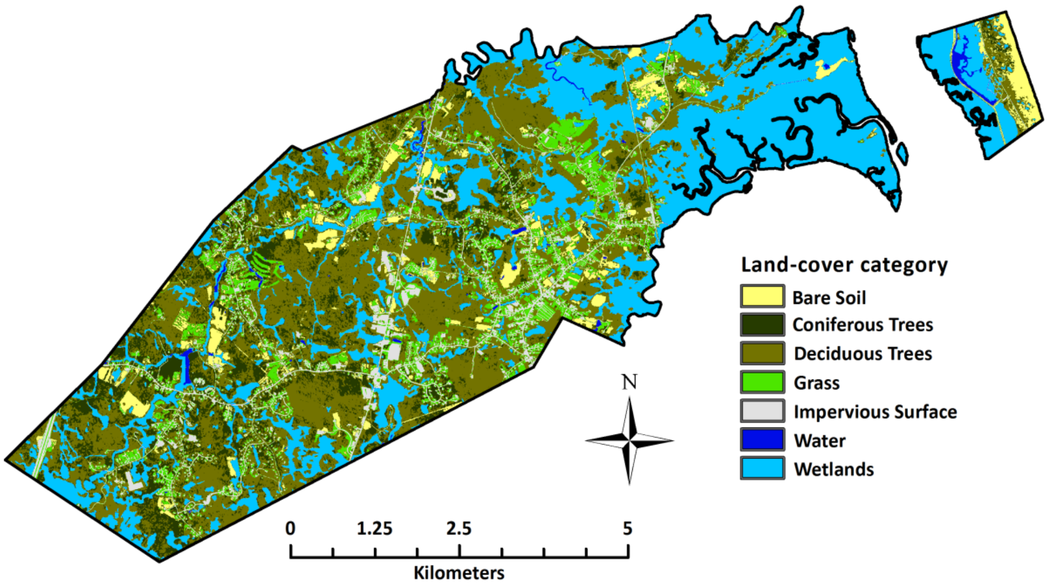


Figure 30. Land-cover map and corresponding land-cover pie chart for the town of Reading.



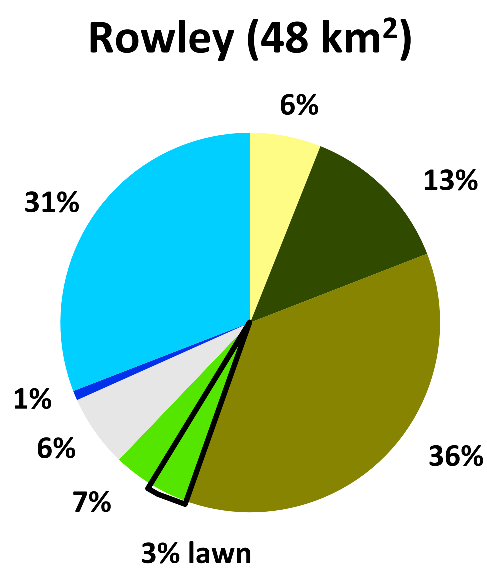
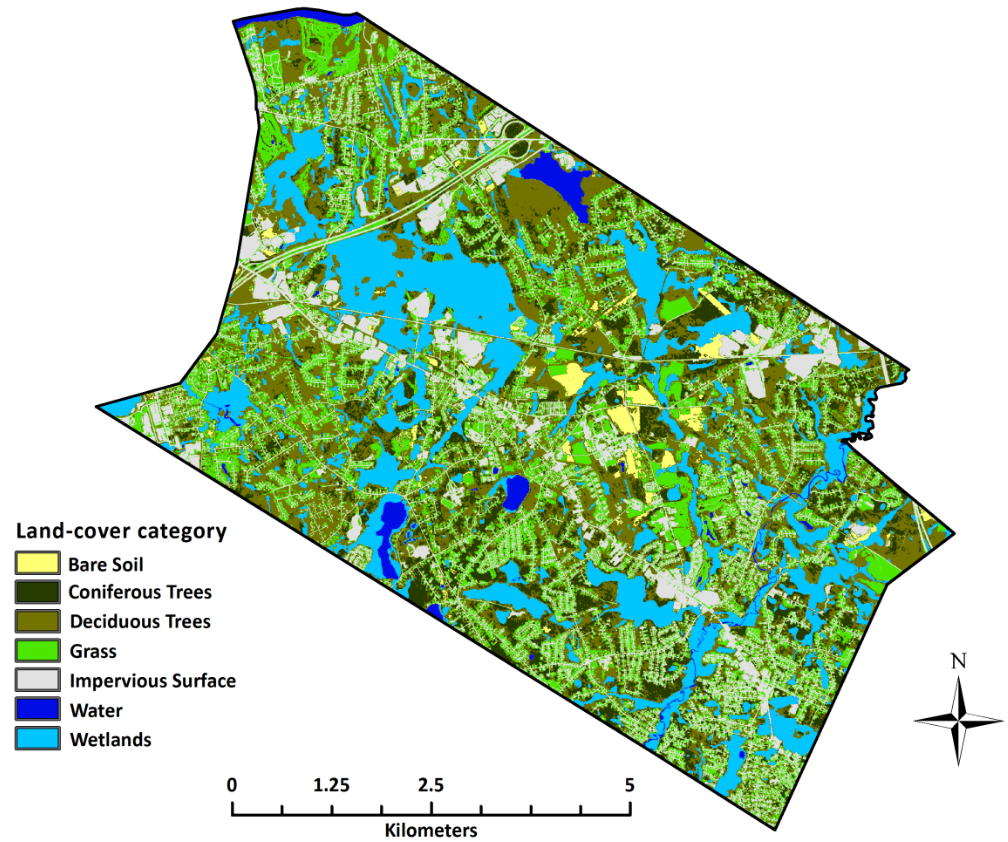


Figure 31. Land-cover map and corresponding land-cover pie chart for the town of Rowley.



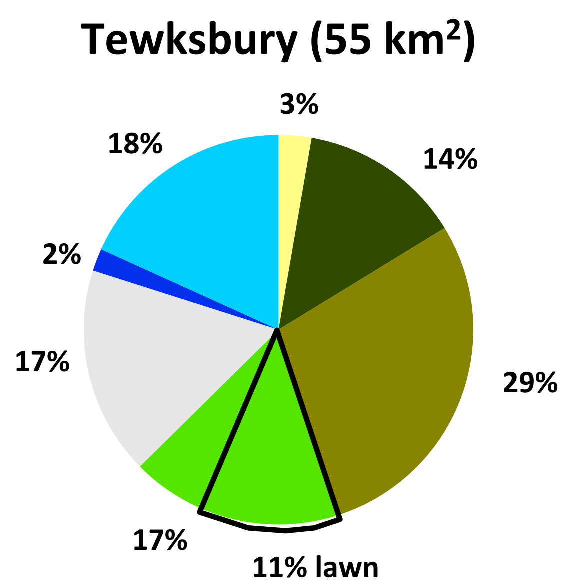
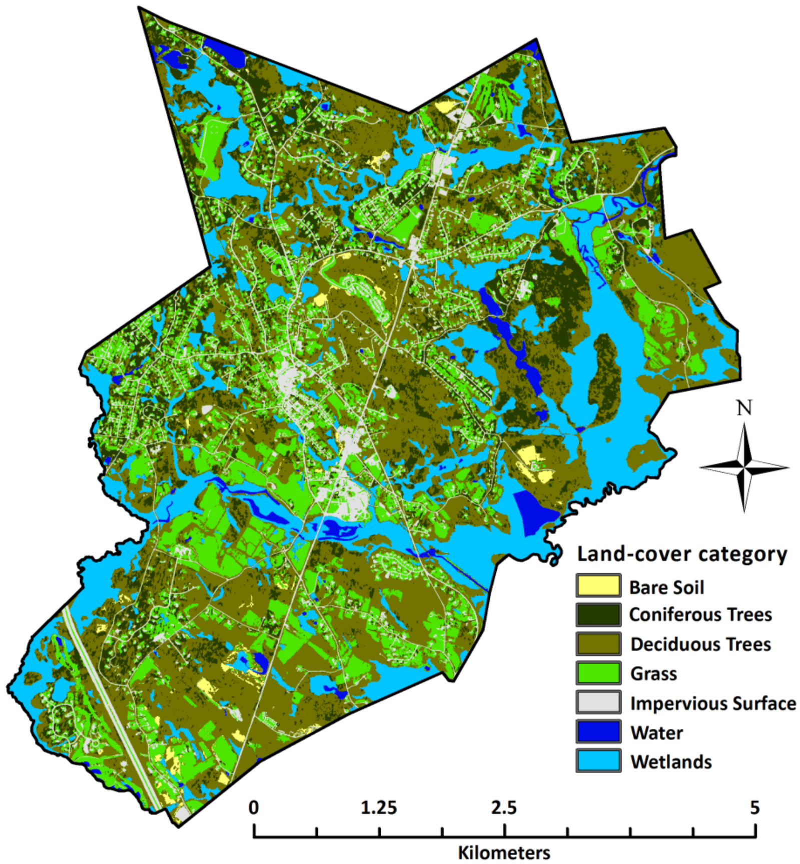


Figure 32. Land-cover map and corresponding land-cover pie chart for the town of Tewksbury.



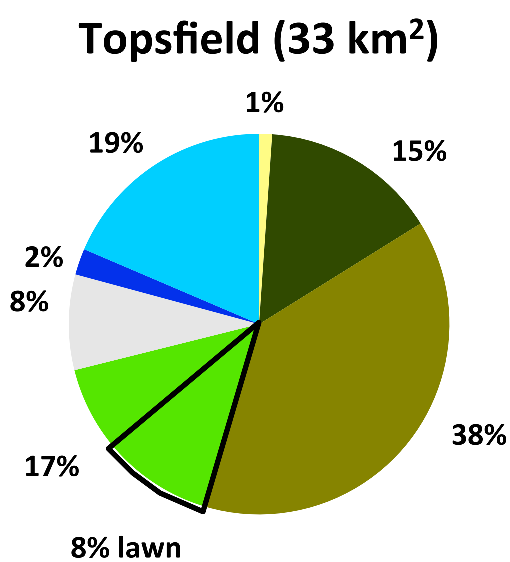
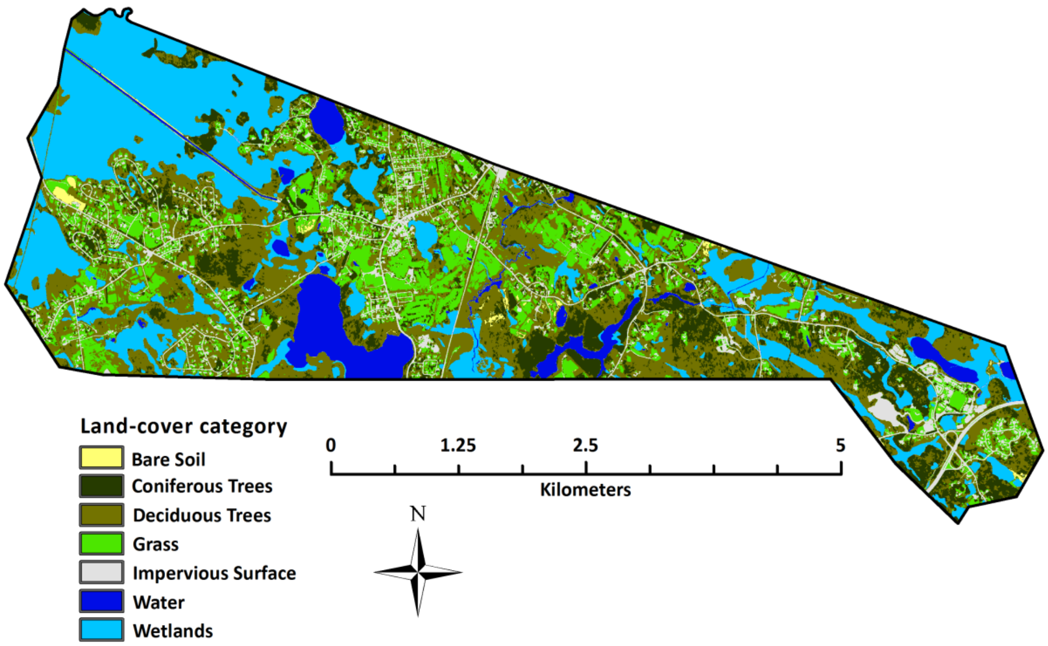


Figure 33. Land-cover map and corresponding land-cover pie chart for the town of Topsfield.



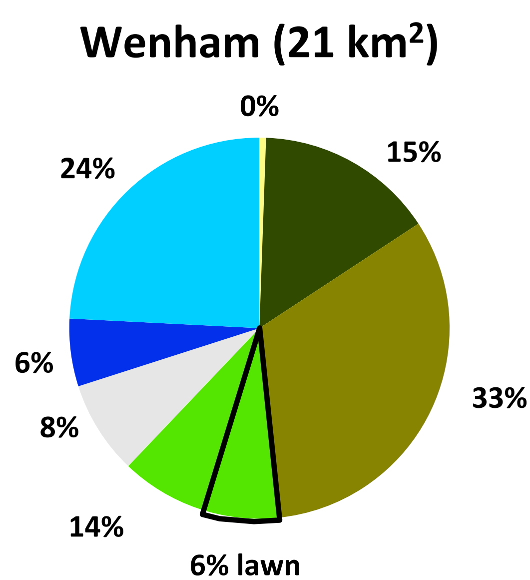
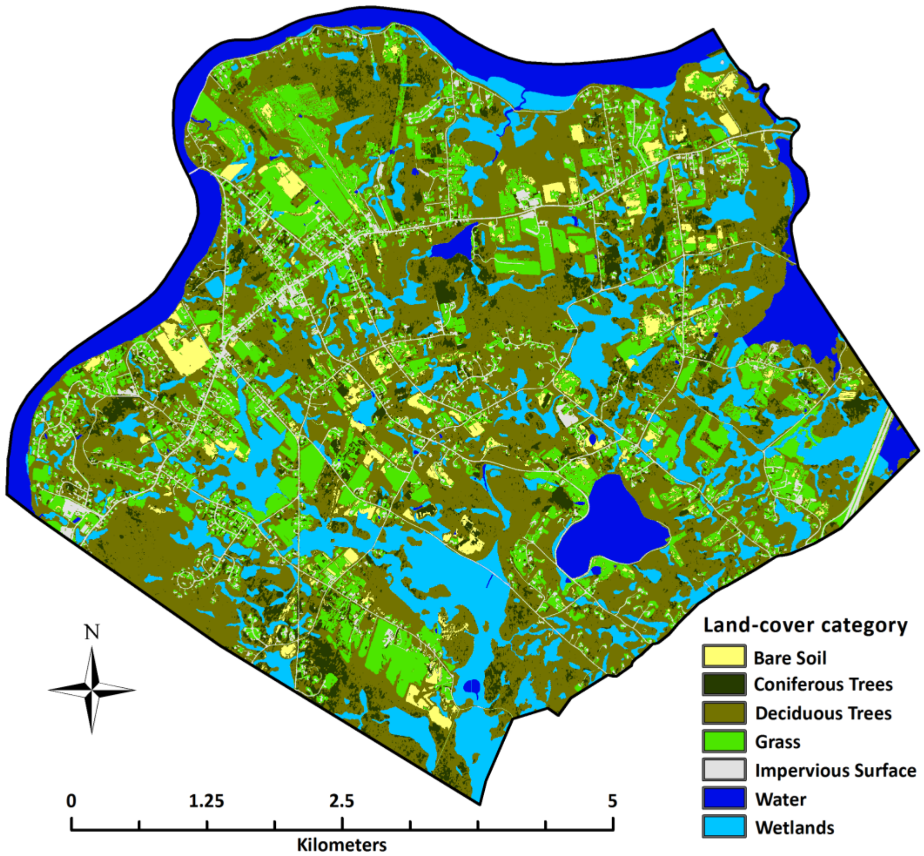


Figure 34. Land-cover map and corresponding land-cover pie chart for the town of Wenham.



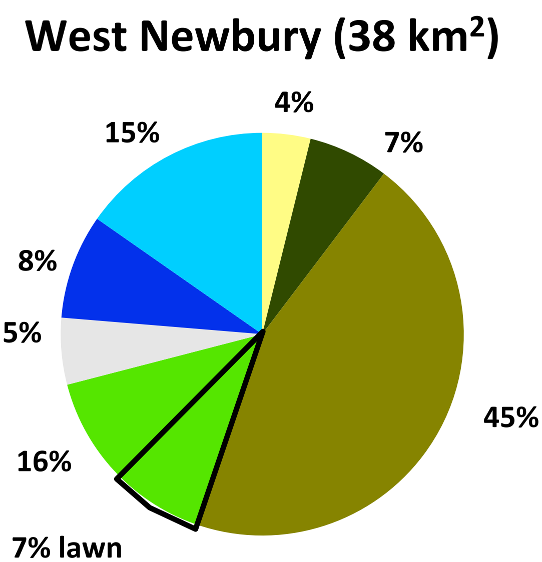
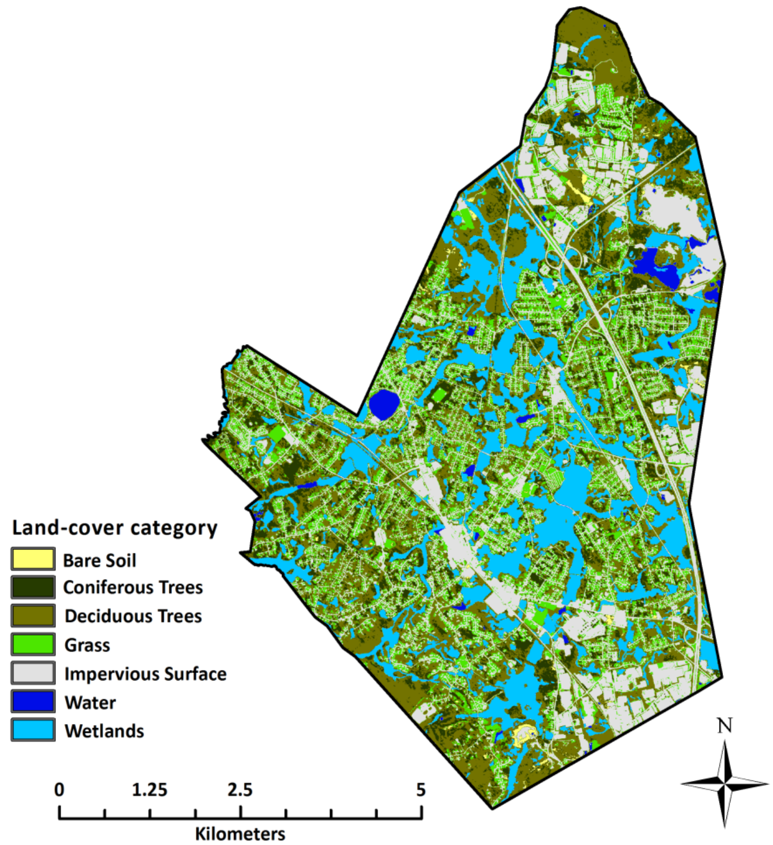


Figure 35. Land-cover map and corresponding land-cover pie chart for the town of West Newbury.



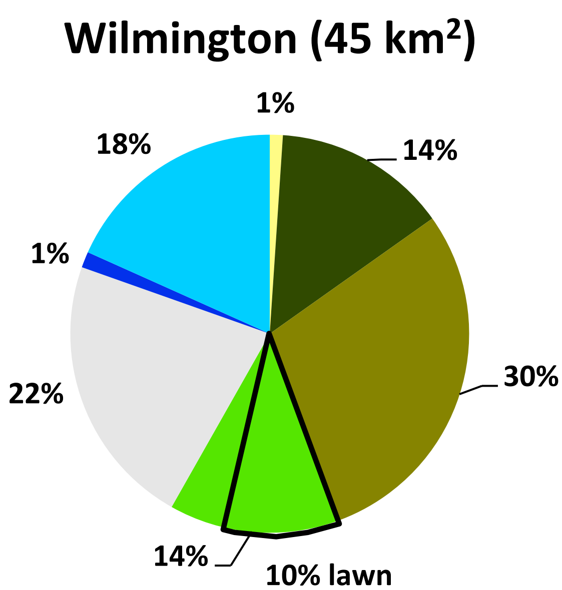
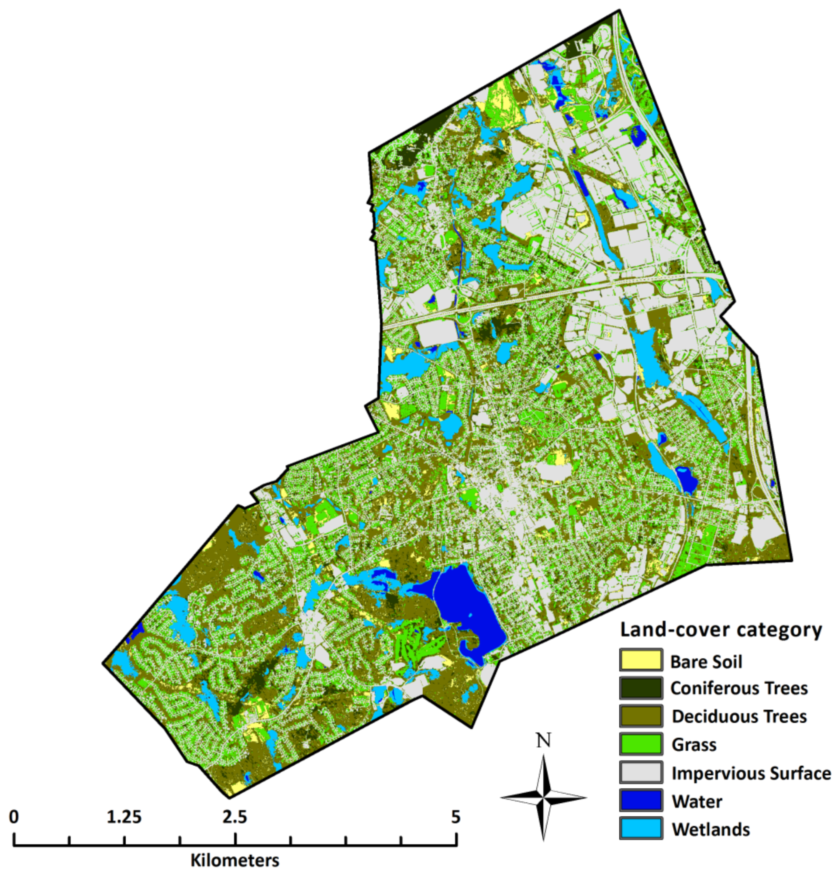


Figure 36. Land-cover map and corresponding land-cover pie chart for the town of Wilmington.



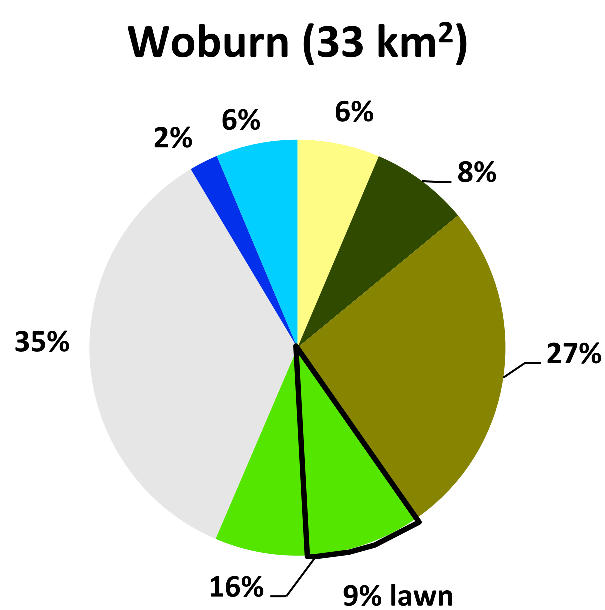


Figure 37. Land-cover map and corresponding land-cover pie chart for the town of Woburn.

Table 2. Estimated population error matrix for the entire PIE study area.

	PIE error matrix (entries in percent of the study area)								
	Reference map								
Land-cover map	Bare soil	Coniferous	Deciduous	Grass	Impervious	Water	Wetlands	Map total	Commission disagreement
Bare soil	**2.32**	0.01	0.34	0.65	0.33	0.00	0.09	3.74	1.42
Coniferous	0.04	**9.77**	2.45	0.37	0.07	0.06	0.07	12.83	3.06
Deciduous	0.55	1.59	**29.62**	1.73	0.27	0.28	0.63	34.68	5.06
Grass	1.29	0.23	0.88	**9.87**	0.41	0.04	0.02	12.73	2.86
Impervious	0.49	0.06	0.30	0.17	**13.61**	0.02	0.00	14.64	1.04
Water	0.02	0.02	0.09	0.01	0.00	**2.65**	0.18	2.97	0.32
Wetlands	0.45	0.62	4.67	0.08	0.04	0.74	**11.79**	18.40	6.61
Reference total	5.16	12.30	38.35	12.88	14.73	3.79	12.78	100.00	20.36
Omission disagreement	2.84	2.53	8.73	3.01	1.12	1.14	0.99	20.36	

Overall agreement = 80%

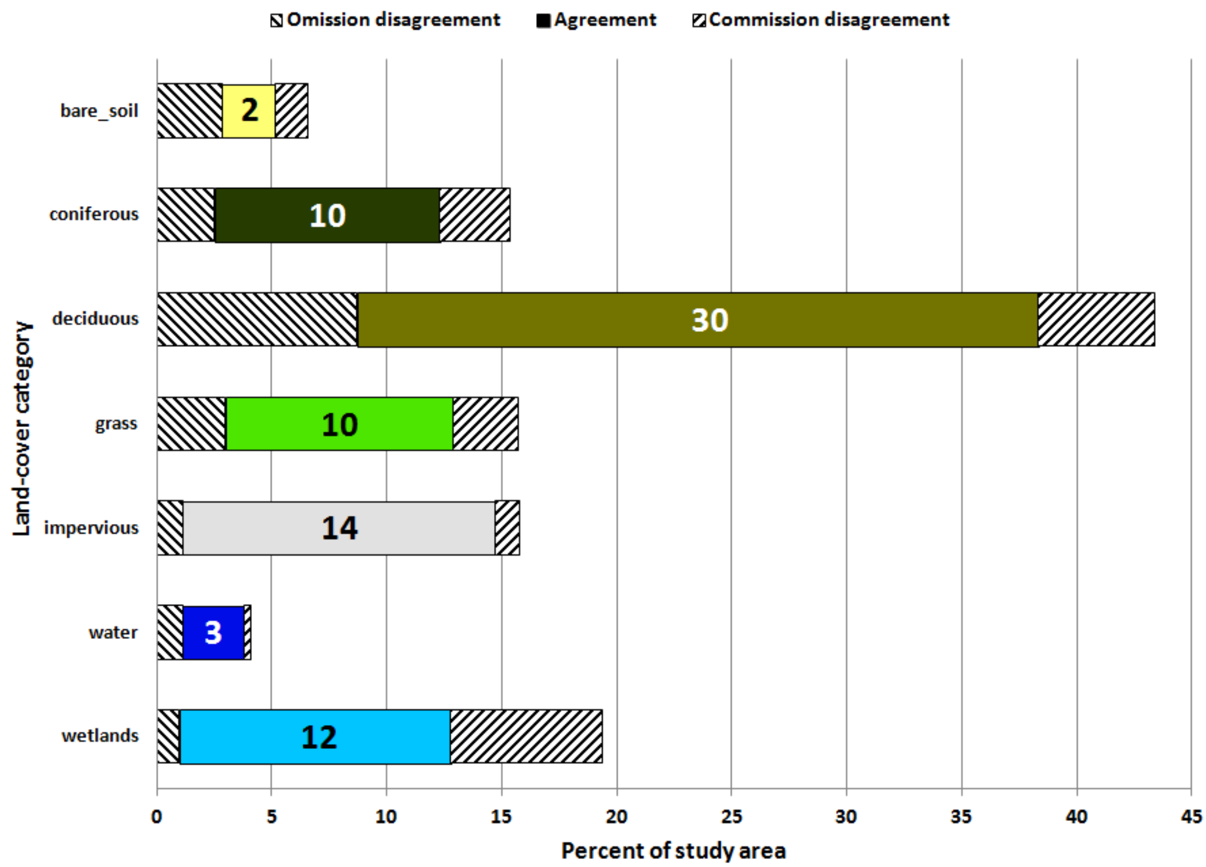


Figure 38. Figure of merit for the entire PIE study area. The colored segments of the bars indicate the percentage of that land-cover category in the study area correctly classified.

Table 3. Estimated population error matrix for the town of Andover.

Andover error matrix (entries in percent of the study area)									
	Reference map								
Land-cover map	Bare soil	Coniferous	Deciduous	Grass	Impervious	Water	Wetlands	Map total	Commission disagreement
Bare soil	**0.76**	0.00	0.31	0.08	0.16	0.00	0.00	1.32	0.56
Coniferous	0.29	**10.86**	1.22	0.17	0.04	0.00	0.00	12.58	1.72
Deciduous	0.12	1.03	**37.68**	1.32	0.26	2.43	0.00	42.85	5.17
Grass	1.35	0.10	0.39	**10.10**	0.07	0.00	0.00	12.00	1.90
Impervious	0.00	0.00	0.00	0.00	**16.68**	0.00	0.00	16.68	0.00
Water	0.00	0.17	0.12	0.01	0.00	**3.40**	0.09	3.79	0.38
Wetlands	0.34	0.67	6.27	0.20	0.00	1.37	**1.94**	10.79	8.86
Reference total	2.86	12.84	45.99	11.87	17.21	7.20	2.02	100.00	18.59
Omission disagreement	2.10	1.98	8.31	1.78	0.54	3.80	0.09	18.59	

Overall agreement = 81%

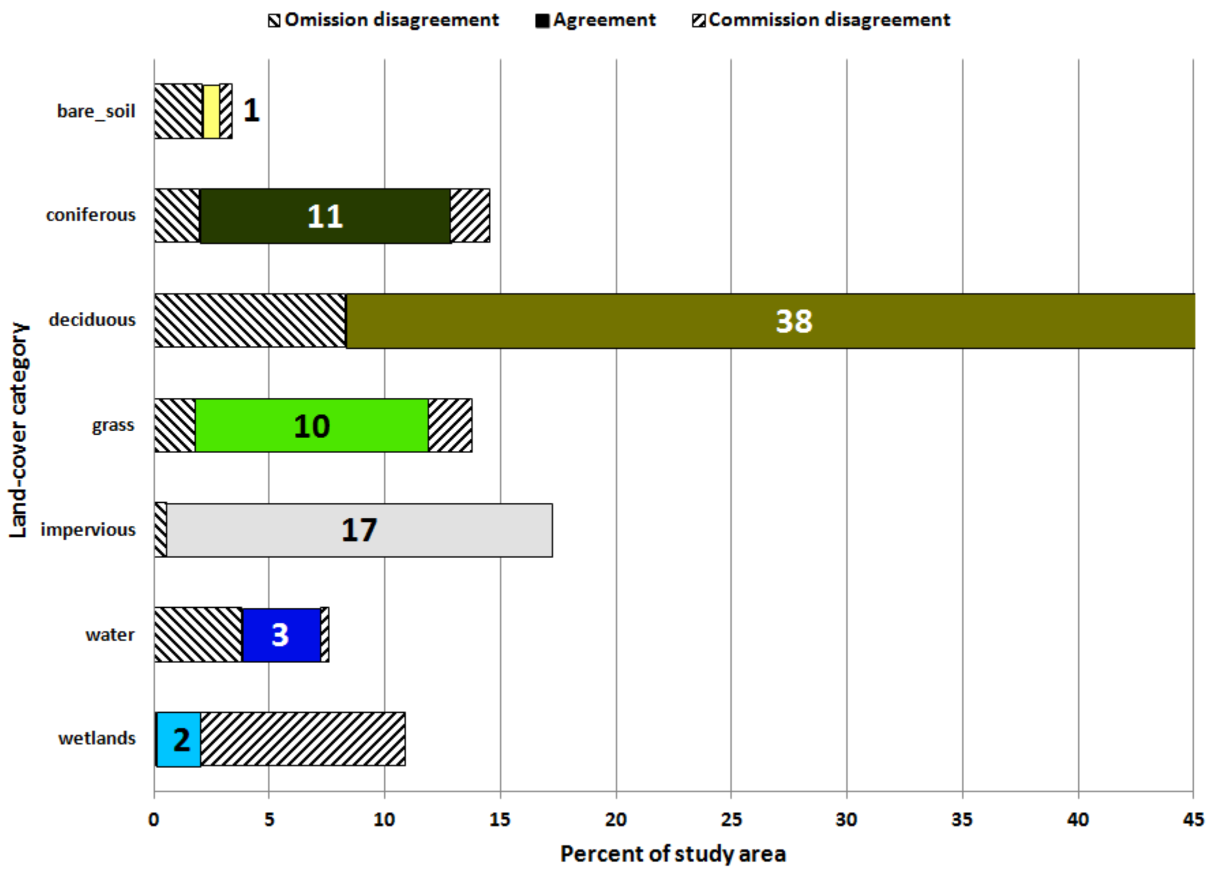


Figure 39. Figure of merit for the town of Andover. The colored segments of the bars indicate the percentage of that land-cover category in the town correctly classified.

Table 4. Estimated population error matrix for the town of Beverly.

				Beverly error matrix (entries in percent of the study area)					
				Reference map					
Land-cover map	Bare soil	Coniferous	Deciduous	Grass	Impervious	Water	Wetlands	Map total	Commission disagreement
Bare soil	**1.39**	0.00	0.05	0.36	0.03	0.00	0.00	1.83	0.44
Coniferous	0.00	**11.32**	3.14	0.33	0.09	0.00	0.12	14.99	3.67
Deciduous	2.18	1.84	**25.53**	2.65	0.31	0.00	0.00	32.50	6.98
Grass	1.22	0.38	1.23	**13.19**	0.92	0.00	0.00	16.96	3.76
Impervious	0.76	0.11	0.25	0.12	**22.51**	0.00	0.00	23.75	1.23
Water	0.04	0.00	0.06	0.01	0.00	**2.27**	0.08	2.46	0.19
Wetlands	0.48	0.14	1.63	0.24	0.00	0.03	**5.00**	7.51	2.51
Reference total	6.06	13.80	31.89	16.89	23.87	2.30	5.20	100.00	18.80
Omission disagreement	4.67	2.48	6.36	3.70	1.36	0.03	0.20	18.80	

Overall agreement = 81%

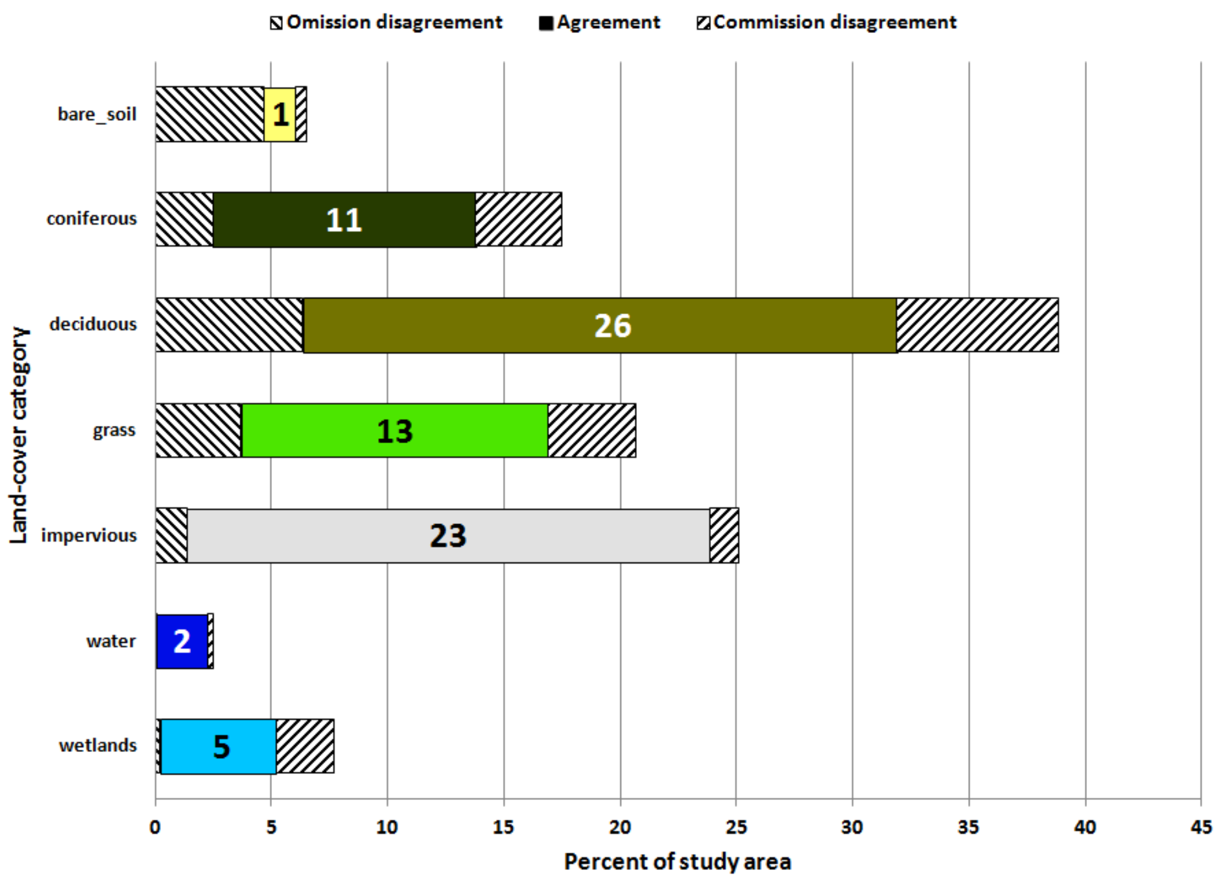


Figure 40. Figure of merit for the town of Beverly. The colored segments of the bars indicate the percentage of that land-cover category in the town correctly classified.

Table 5. Estimated population error matrix for the town of Billerica.

	Billerica error matrix (entries in percent of the study area)								
	Reference map								
Land-cover map	Bare soil	Coniferous	Deciduous	Grass	Impervious	Water	Wetlands	Map total	Commission disagreement
Bare soil	**0.73**	0.00	0.07	0.44	0.14	0.00	0.00	1.38	0.65
Coniferous	0.00	**12.20**	1.20	0.75	0.07	0.07	0.00	14.30	2.10
Deciduous	3.23	3.53	**24.23**	1.95	0.63	0.38	0.00	33.96	9.73
Grass	1.09	0.23	0.75	**12.95**	0.42	0.00	0.00	15.44	2.49
Impervious	0.30	0.00	0.77	0.07	**18.98**	0.00	0.00	20.13	1.15
Water	0.00	0.00	0.05	0.03	0.00	**2.50**	0.34	2.92	0.42
Wetlands	0.32	0.67	5.86	0.23	0.00	0.29	**4.50**	11.88	7.38
Reference total	5.67	16.64	32.93	16.43	20.25	3.25	4.83	100.00	23.90
Omission disagreement	4.94	4.44	8.70	3.47	1.27	0.75	0.34	23.90	

Overall agreement = 76%

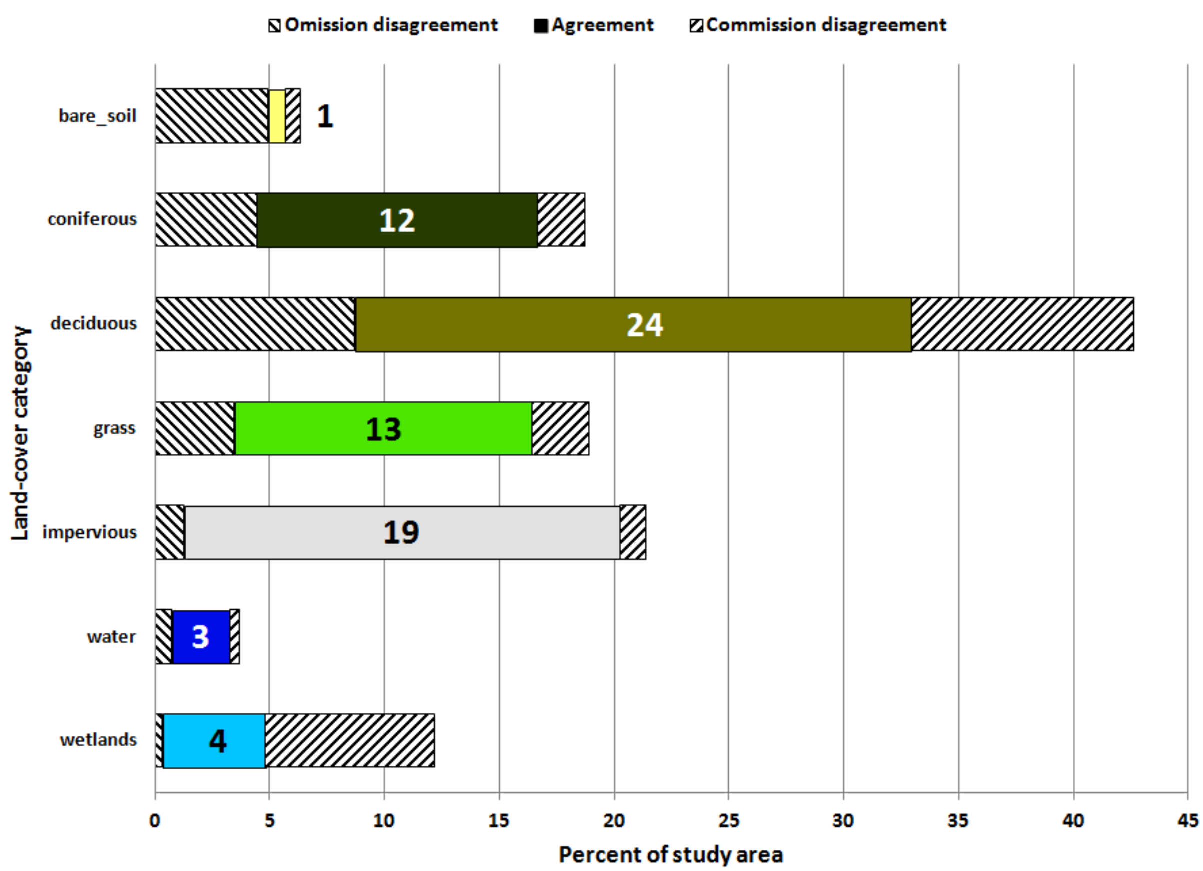


Figure 41. Figure of merit for the town of Billerica. The colored segments of the bars indicate the percentage of that land-cover category in the town correctly classified.

Table 6. Estimated population error matrix for the town of Boxford.

Boxford error matrix (entries in percent of the study area)									
	Reference map								
Land-cover map	Bare soil	Coniferous	Deciduous	Grass	Impervious	Water	Wetlands	Map total	Commission disagreement
Bare soil	**1.79**	0.00	0.04	0.14	0.01	0.00	0.04	2.02	0.23
Coniferous	0.00	**19.01**	1.84	0.00	0.02	0.00	0.12	20.99	1.99
Deciduous	0.33	1.24	**41.24**	0.67	0.02	0.00	0.11	43.61	2.37
Grass	1.61	0.12	0.31	**6.12**	0.09	0.00	0.00	8.24	2.12
Impervious	0.49	0.02	0.07	0.01	**6.06**	0.00	0.00	6.64	0.59
Water	0.02	0.02	0.08	0.00	0.00	**2.61**	0.09	2.81	0.20
Wetlands	0.02	0.93	0.45	0.01	0.00	0.97	**13.31**	15.68	2.37
Reference total	4.26	21.32	44.02	6.94	6.21	3.58	13.67	100.00	9.88
Omission disagreement	2.47	2.32	2.79	0.82	0.15	0.97	0.36	9.88	

Overall agreement = 90%

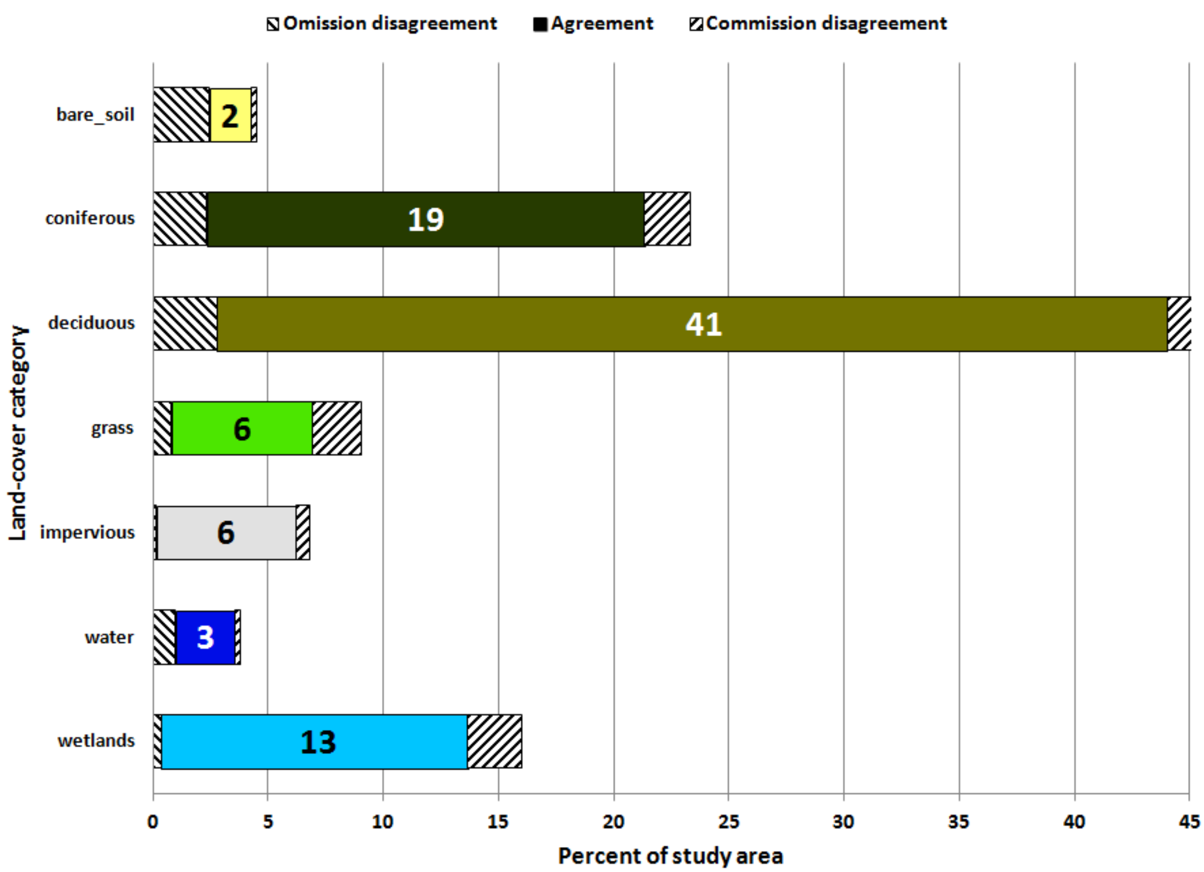


Figure 42. Figure of merit for the town of Boxford. The colored segments of the bars indicate the percentage of that land-cover category in the town correctly classified.

Table 7. Estimated population error matrix for the town of Burlington.

Burlington error matrix (entries in percent of the study area)									
	Reference map								
Land-cover map	Bare soil	Coniferous	Deciduous	Grass	Impervious	Water	Wetlands	Map total	Commission disagreement
Bare soil	**0.71**	0.04	1.61	2.14	1.02	0.00	0.32	5.84	5.13
Coniferous	0.13	**9.24**	2.52	0.71	0.15	0.20	0.31	13.25	4.02
Deciduous	0.00	0.75	**19.72**	0.81	0.19	0.00	5.30	26.77	7.05
Grass	0.23	1.17	1.41	**11.67**	0.14	0.00	0.00	14.62	2.96
Impervious	0.02	0.18	0.25	1.01	**26.76**	0.00	0.12	28.34	1.58
Water	0.01	0.02	0.05	0.01	0.00	**0.94**	0.07	1.11	0.16
Wetlands	0.00	0.16	0.79	0.08	0.01	0.11	**8.91**	10.06	1.16
Reference total	1.11	11.56	26.35	16.42	28.27	1.26	15.03	100.00	22.05
Omission disagreement	0.40	2.33	6.63	4.76	1.51	0.31	6.12	22.05	

Overall agreement = 78%

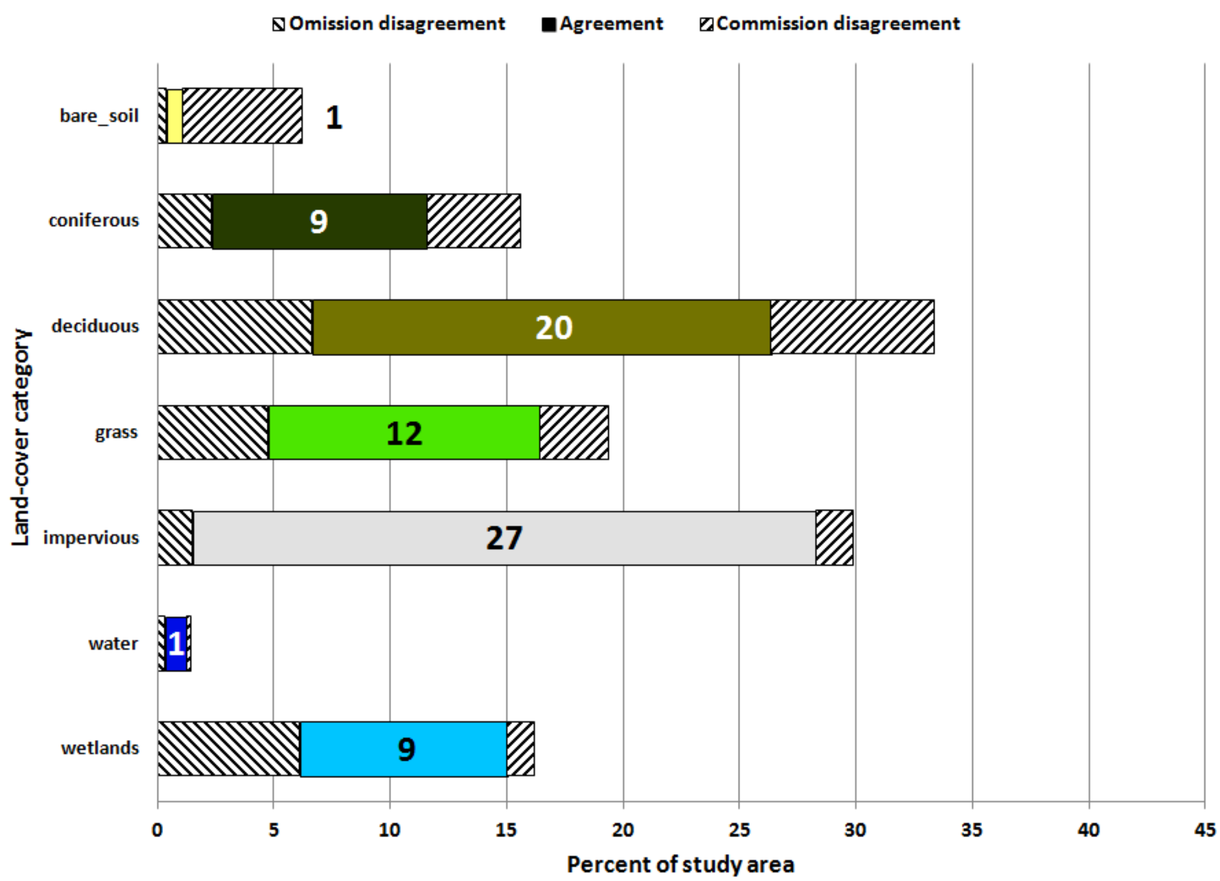


Figure 43. Figure of merit for the town of Burlington. The colored segments of the bars indicate the percentage of that land-cover category in the town correctly classified.

Table 8. Estimated population error matrix for the town of Danvers.

	Danvers error matrix (entries in percent of the study area)								
	Reference map								
Land-cover map	Bare soil	Coniferous	Deciduous	Grass	Impervious	Water	Wetlands	Map total	Commission disagreement
Bare soil	**9.60**	0.05	0.73	1.36	0.76	0.01	0.00	12.52	2.92
Coniferous	0.00	**3.34**	2.44	0.30	0.19	0.00	0.07	6.34	3.00
Deciduous	0.00	0.19	**24.39**	1.12	0.33	0.00	1.16	27.19	2.79
Grass	0.15	0.39	1.03	**13.87**	0.53	0.00	0.07	16.04	2.17
Impervious	1.10	0.15	0.42	0.72	**26.22**	0.00	0.00	28.61	2.38
Water	0.02	0.02	0.19	0.00	0.00	**3.71**	1.02	4.96	1.25
Wetlands	0.13	0.04	1.94	0.00	0.00	0.01	**2.22**	4.35	2.13
Reference total	11.00	4.18	31.14	17.37	28.03	3.73	4.55	100.00	16.64
Omission disagreement	1.39	0.84	6.75	3.50	1.81	0.02	2.32	16.64	

Overall agreement = 83%

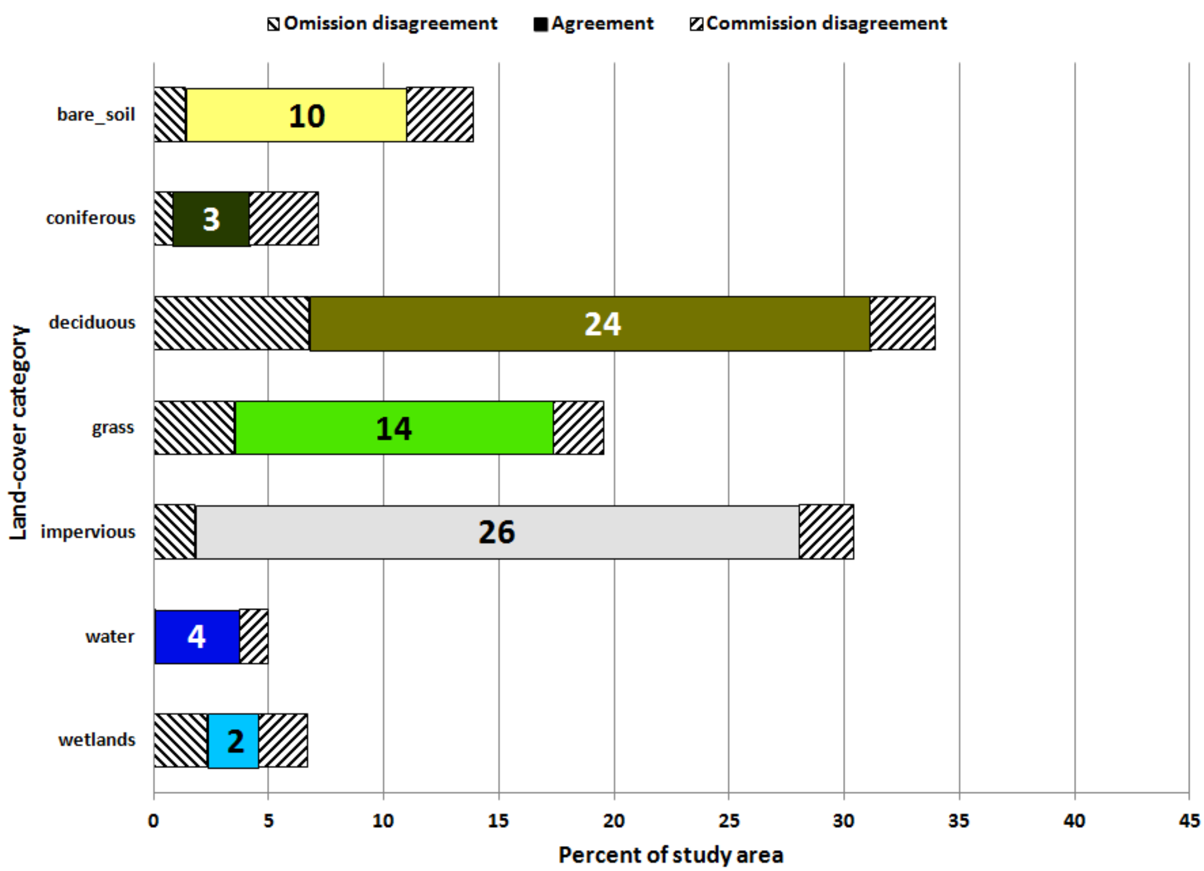


Figure 44. Figure of merit for the town of Danvers. The colored segments of the bars indicate the percentage of that land-cover category in the town correctly classified.

Table 9. Estimated population error matrix for the town of Essex.

	Essex error matrix (entries in percent of the study area)								
	Reference map								
Land-cover map	Bare soil	Coniferous	Deciduous	Grass	Impervious	Water	Wetlands	Map total	Commission disagreement
Bare soil	**2.09**	0.01	0.12	1.20	0.08	0.00	0.00	3.50	1.41
Coniferous	0.01	**12.37**	2.44	1.16	0.01	0.00	0.01	16.00	3.63
Deciduous	1.34	1.09	**31.55**	0.18	0.02	0.00	0.02	34.20	2.65
Grass	0.95	0.06	0.08	**4.90**	0.03	0.00	0.22	6.25	1.35
Impervious	0.18	0.00	0.17	0.04	**3.74**	0.50	0.00	4.63	0.89
Water	0.00	0.02	0.03	0.00	0.00	**1.43**	0.03	1.52	0.09
Wetlands	0.78	1.10	1.12	0.03	0.00	0.73	**30.13**	33.90	3.77
Reference total	5.36	14.66	35.51	7.53	3.88	2.67	30.40	100.00	13.78
Omission disagreement	3.27	2.29	3.96	2.62	0.14	1.23	0.27	13.78	

Overall agreement = 86%

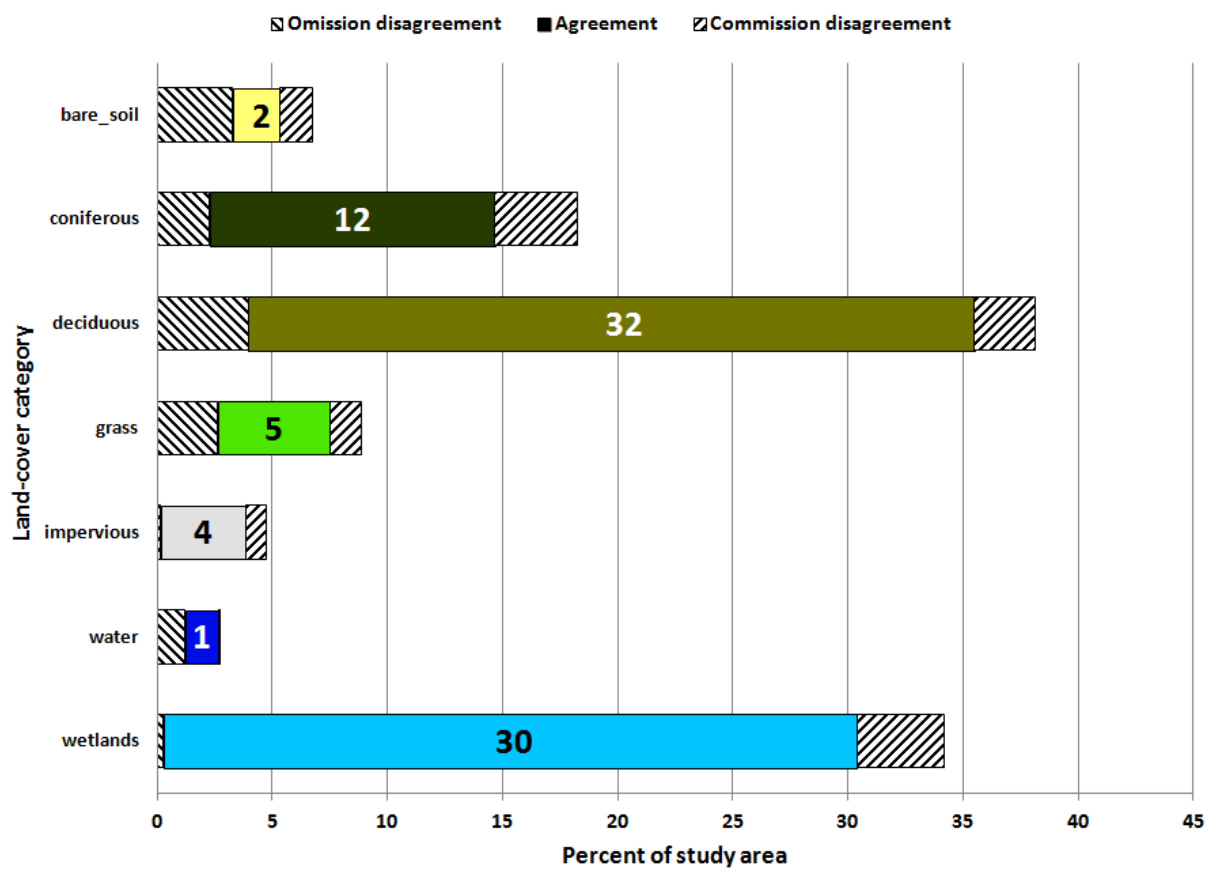


Figure 45. Figure of merit for the town of Essex. The colored segments of the bars indicate the percentage of that land-cover category in the town correctly classified.

Table 10. Estimated population error matrix for the town of Georgetown.

	Georgetown error matrix (entries in percent of the study area)								
	Reference map								
Land-cover map	Bare soil	Coniferous	Deciduous	Grass	Impervious	Water	Wetlands	Map total	Commission disagreement
Bare soil	**1.13**	0.00	0.33	0.42	0.60	0.00	0.00	2.49	1.36
Coniferous	0.00	**13.58**	3.32	0.09	0.07	0.00	0.00	17.06	3.48
Deciduous	0.22	1.89	**39.73**	0.77	0.23	0.03	0.00	42.88	3.15
Grass	0.83	0.04	0.89	**7.98**	0.15	0.00	0.00	9.89	1.91
Impervious	0.43	0.00	0.52	0.07	**8.78**	0.00	0.00	9.80	1.01
Water	0.01	0.01	0.12	0.00	0.00	**1.51**	0.02	1.67	0.16
Wetlands	0.02	1.24	7.21	0.30	0.00	0.89	**6.57**	16.23	9.66
Reference total	2.64	16.77	52.12	9.62	9.84	2.43	6.59	100.00	20.72
Omission disagreement	1.51	3.18	12.39	1.64	1.05	0.92	0.02	20.72	

Overall agreement = 79%

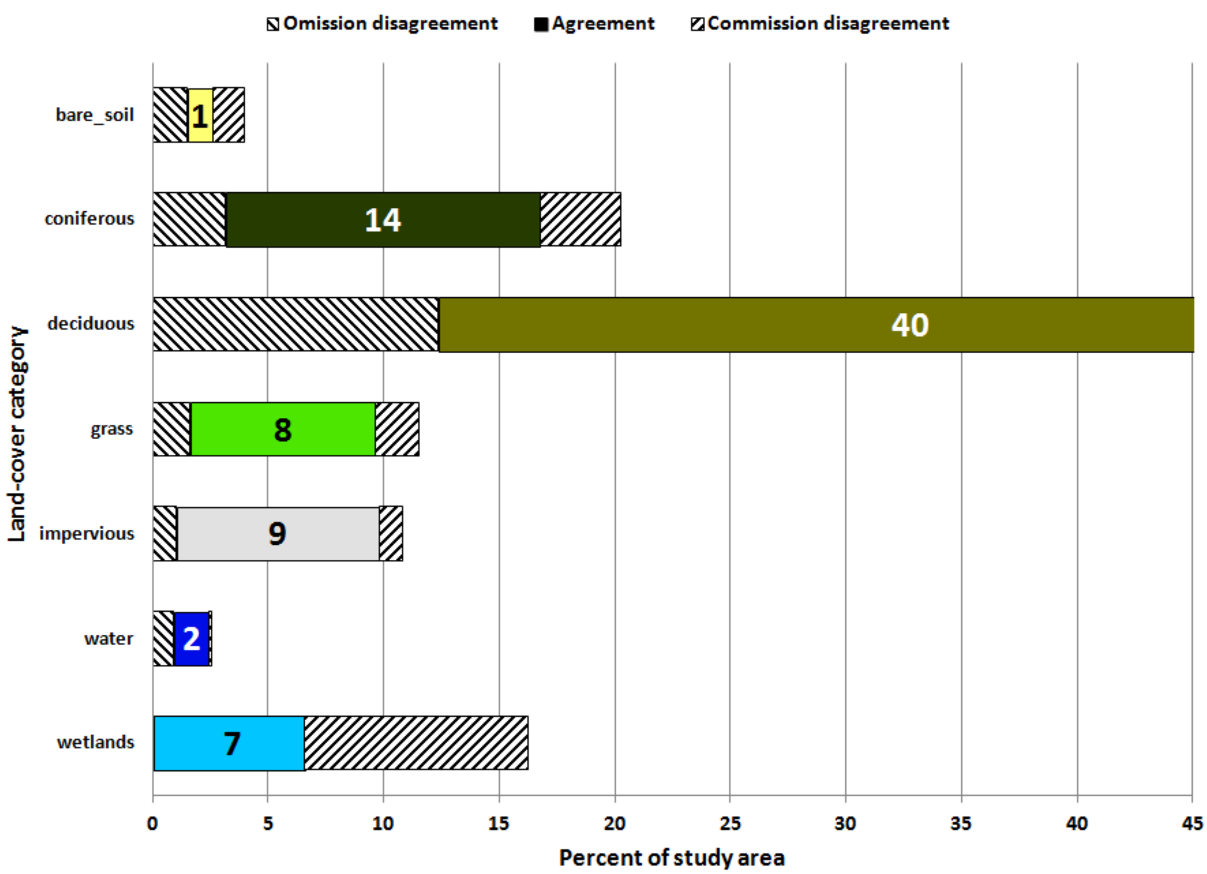


Figure 46. Figure of merit for the town of Georgetown. The colored segments of the bars indicate the percentage of that land-cover category in the town correctly classified.

Table11. Estimated population error matrix for the town of Groveland.

Groveland error matrix (entries in percent of the study area)									
	Reference map								
Land-cover map	Bare soil	Coniferous	Deciduous	Grass	Impervious	Water	Wetlands	Map total	Commission disagreement
Bare soil	**2.14**	0.02	1.35	2.43	0.27	0.00	0.01	6.23	4.09
Coniferous	0.02	**8.54**	2.94	1.25	0.15	0.12	0.17	13.19	4.65
Deciduous	1.43	0.70	**36.71**	2.64	0.61	0.05	0.00	42.15	5.44
Grass	0.65	0.04	0.88	**5.54**	0.41	0.00	0.00	7.51	1.97
Impervious	0.73	0.04	0.52	0.18	**8.30**	0.00	0.00	9.77	1.47
Water	0.00	0.00	0.14	0.00	0.00	**4.53**	0.01	4.68	0.15
Wetlands	0.49	0.18	3.58	0.00	0.00	0.40	**11.82**	16.48	4.66
Reference total	5.45	9.53	46.12	12.04	9.75	5.11	12.01	100.00	22.42
Omission disagreement	3.32	0.99	9.40	6.50	1.44	0.58	0.19	22.42	

Overall agreement = 78%

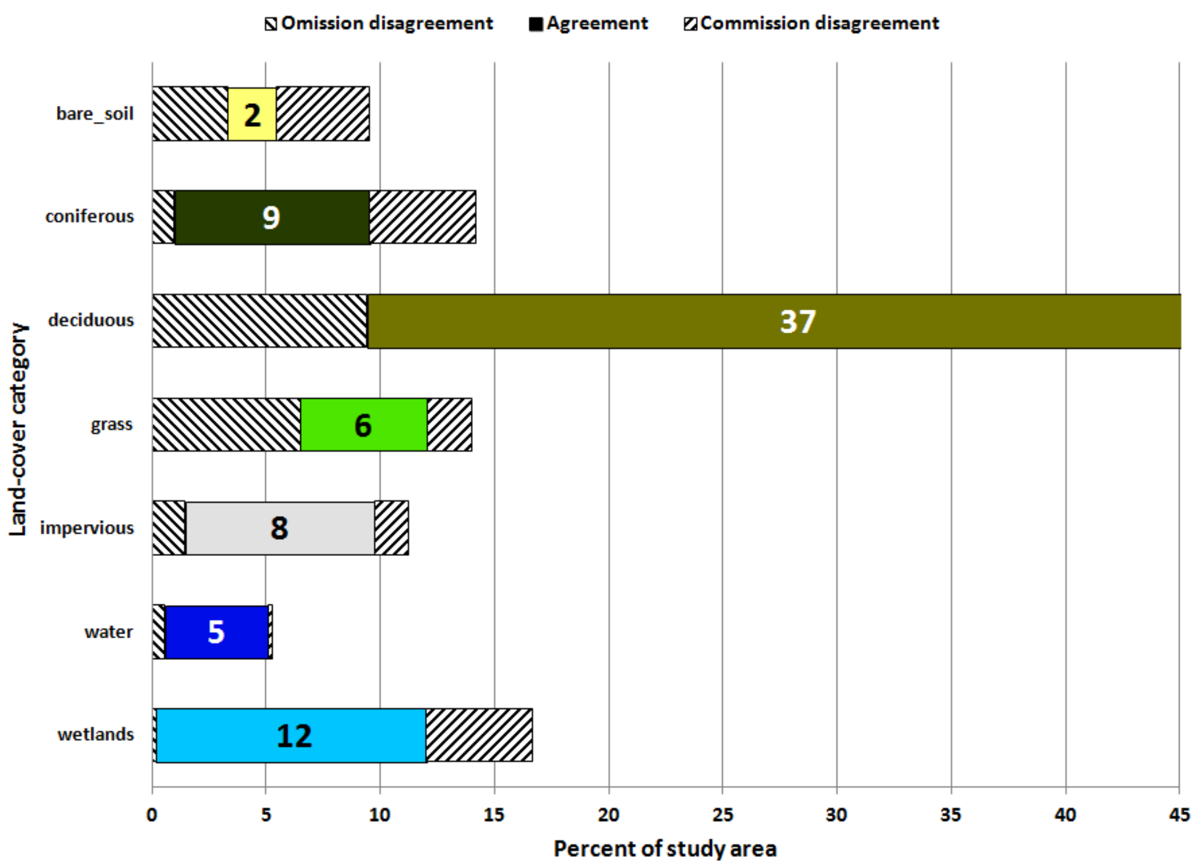


Figure 47. Figure of merit for the town of Groveland. The colored segments of the bars indicate the percentage of that land-cover category in the town correctly classified.

Table 12. Estimated population error matrix for the town of Hamilton.

Hamilton error matrix (entries in percent of the study area)									
	Reference map								
Land-cover map	Bare soil	Coniferous	Deciduous	Grass	Impervious	Water	Wetlands	Map total	Commission disagreement
Bare soil	**2.24**	0.01	0.20	1.50	0.27	0.00	0.00	4.22	1.98
Coniferous	0.18	**10.40**	4.24	0.59	0.06	0.00	0.13	15.59	5.19
Deciduous	0.87	1.23	**27.87**	2.87	0.28	0.00	0.50	33.63	5.76
Grass	1.15	0.19	0.44	**11.26**	0.28	0.00	0.00	13.31	2.06
Impervious	0.03	0.01	0.30	0.06	**6.65**	0.00	0.00	7.06	0.41
Water	0.00	0.02	0.18	0.00	0.01	**2.67**	0.06	2.95	0.28
Wetlands	0.68	0.60	3.57	0.06	0.09	1.13	**17.09**	23.23	6.14
Reference total	5.17	12.45	36.81	16.35	7.64	3.80	17.78	100.00	21.82
Omission disagreement	2.92	2.06	8.94	5.09	0.99	1.13	0.69	21.82	

Overall agreement = 78%

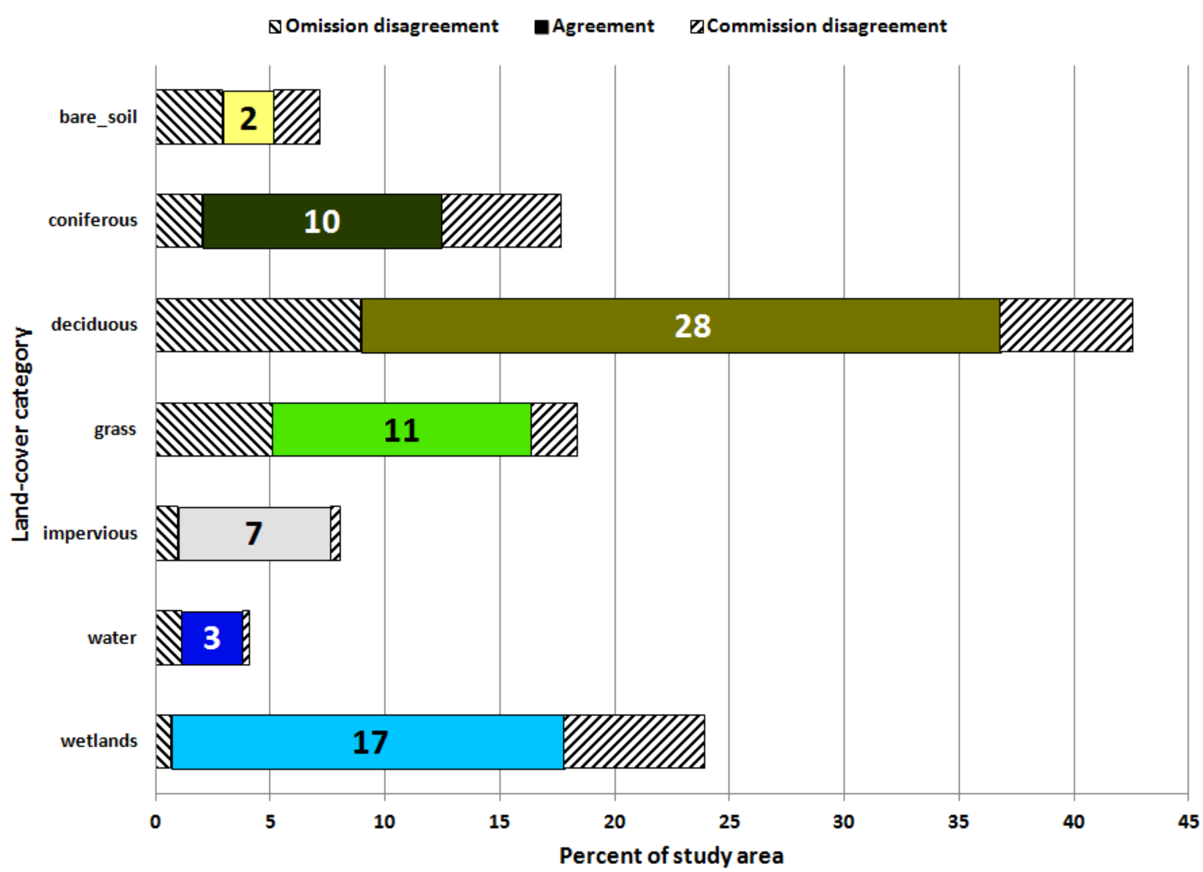


Figure 48. Figure of merit for the town of Hamilton. The colored segments of the bars indicate the percentage of that land-cover category in the town correctly classified.

Table 13. Estimated population error matrix for the town of Ipswich.

Ipswich error matrix (entries in percent of the study area)

Land-cover map	Reference map: Bare soil	Coniferous	Deciduous	Grass	Impervious	Water	Wetlands	Map total	Commission disagreement
Bare soil	**5.78**	0.00	0.55	0.27	0.51	0.03	0.66	7.81	2.02
Coniferous	0.11	**5.33**	3.24	0.26	0.02	0.00	0.26	9.23	3.90
Deciduous	0.35	5.31	**25.52**	1.73	0.10	0.09	1.13	34.23	8.71
Grass	1.97	0.21	0.31	**6.36**	0.54	0.00	0.00	9.39	3.04
Impervious	0.31	0.13	0.14	0.10	**5.42**	0.00	0.00	6.10	0.68
Water	0.01	0.00	0.06	0.00	0.00	**2.19**	0.16	2.42	0.23
Wetlands	1.48	0.00	5.86	0.00	0.23	1.24	**22.01**	30.82	8.81
Reference total	10.00	10.99	35.69	8.72	6.82	3.55	24.24	100.00	27.39
Omission disagreement	4.22	5.66	10.17	2.36	1.40	1.36	2.22	27.39	

Overall agreement = 73%

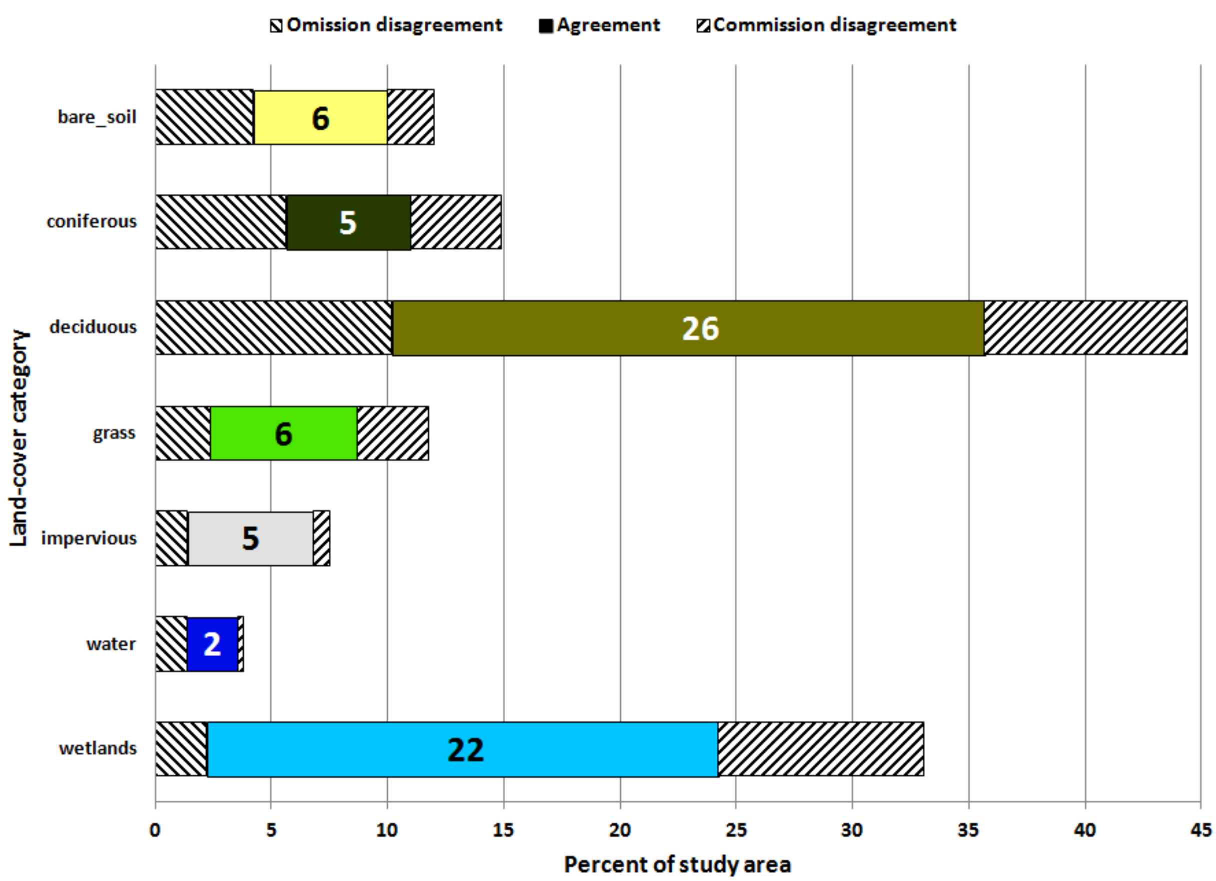


Figure 49. Figure of merit for the town of Ipswich. The colored segments of the bars indicate the percentage of that land-cover category in the town correctly classified.

Table 14. Estimated population error matrix for the town of Lynnfield.

Lynnfield error matrix (entries in percent of the study area)									
	Reference map								
Land-cover map	Bare soil	Coniferous	Deciduous	Grass	Impervious	Water	Wetlands	Map total	Commission disagreement
Bare soil	**0.42**	0.00	1.15	0.80	0.34	0.00	0.00	2.71	2.28
Coniferous	0.00	**11.03**	6.19	0.37	0.10	0.00	0.00	17.69	6.66
Deciduous	0.08	2.29	**25.94**	1.45	0.12	0.00	0.00	29.88	3.94
Grass	0.06	0.18	2.38	**8.99**	0.20	0.00	0.00	11.82	2.83
Impervious	0.02	0.07	0.72	0.20	**14.43**	0.00	0.00	15.45	1.02
Water	0.00	0.02	0.05	0.01	0.00	**4.23**	0.01	4.31	0.08
Wetlands	0.81	0.43	11.67	0.00	0.00	0.05	**5.19**	18.14	12.95
Reference total	1.40	14.02	48.10	11.82	15.19	4.28	5.19	100.00	29.77
Omission disagreement	0.98	2.99	22.16	2.83	0.76	0.05	0.01	29.77	

Overall agreement = 70%

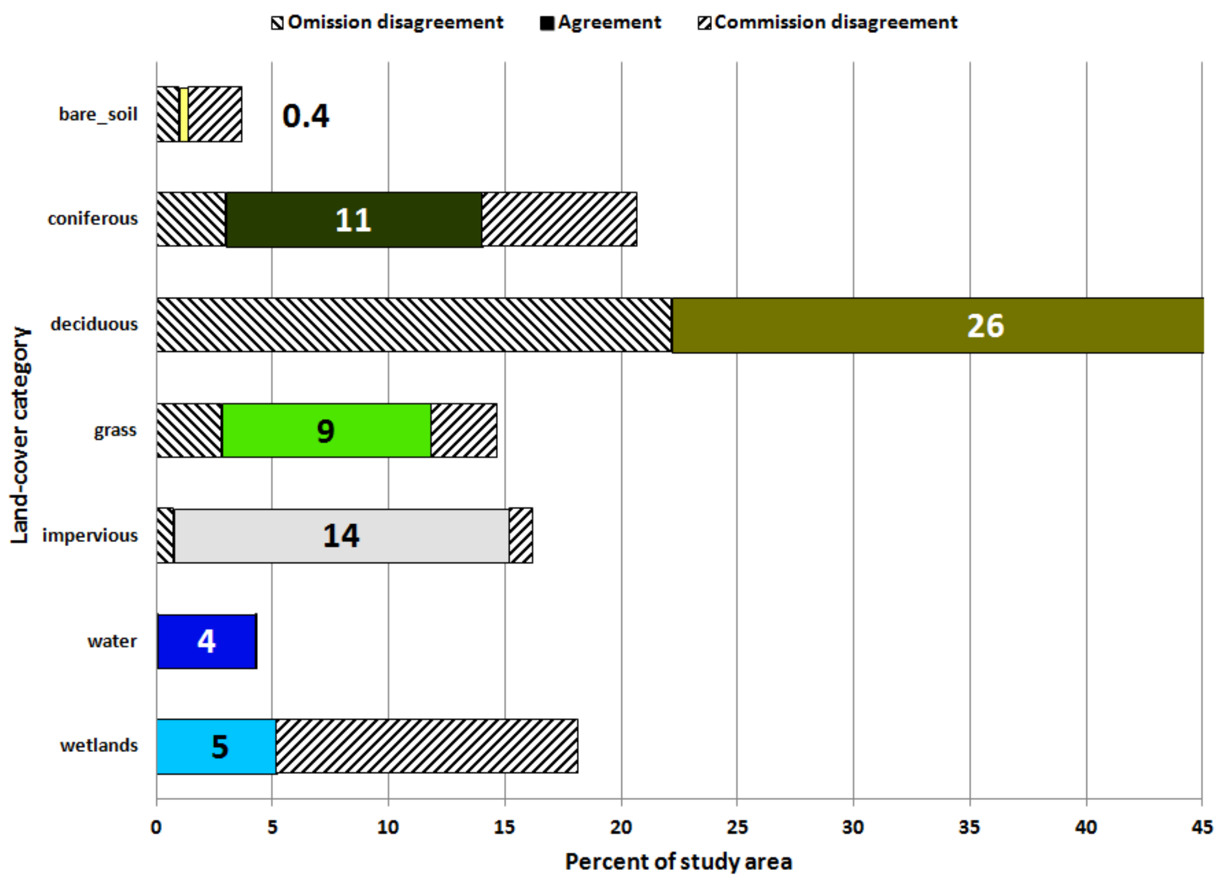


Figure 50. Figure of merit for the town of Lynnfield. The colored segments of the bars indicate the percentage of that land-cover category in the town correctly classified.

Table 15. Estimated population error matrix for the town of Middleton.

				Middleton error matrix (entries in percent of the study area)					
				Reference map					
Land-cover map	Bare soil	Coniferous	Deciduous	Grass	Impervious	Water	Wetlands	Map total	Commission disagreement
Bare soil	**1.42**	0.00	0.12	0.58	0.03	0.00	0.00	2.16	0.73
Coniferous	0.00	**10.68**	6.11	0.02	0.04	0.07	0.16	17.07	6.39
Deciduous	0.00	2.28	**31.81**	0.97	0.13	0.00	0.10	35.30	3.49
Grass	1.05	0.16	0.56	**8.69**	0.61	0.00	0.25	11.32	2.63
Impervious	1.71	0.13	0.62	0.37	**7.41**	0.00	0.00	10.24	2.83
Water	0.00	0.00	0.07	0.03	0.00	**4.86**	0.49	5.46	0.60
Wetlands	0.02	0.24	6.76	0.02	0.00	0.61	**10.80**	18.45	7.65
Reference total	4.20	13.51	46.04	10.68	8.23	5.54	11.80	100.00	24.32
Omission disagreement	2.78	2.83	14.23	1.99	0.82	0.68	0.99	24.32	

Overall agreement = 76%

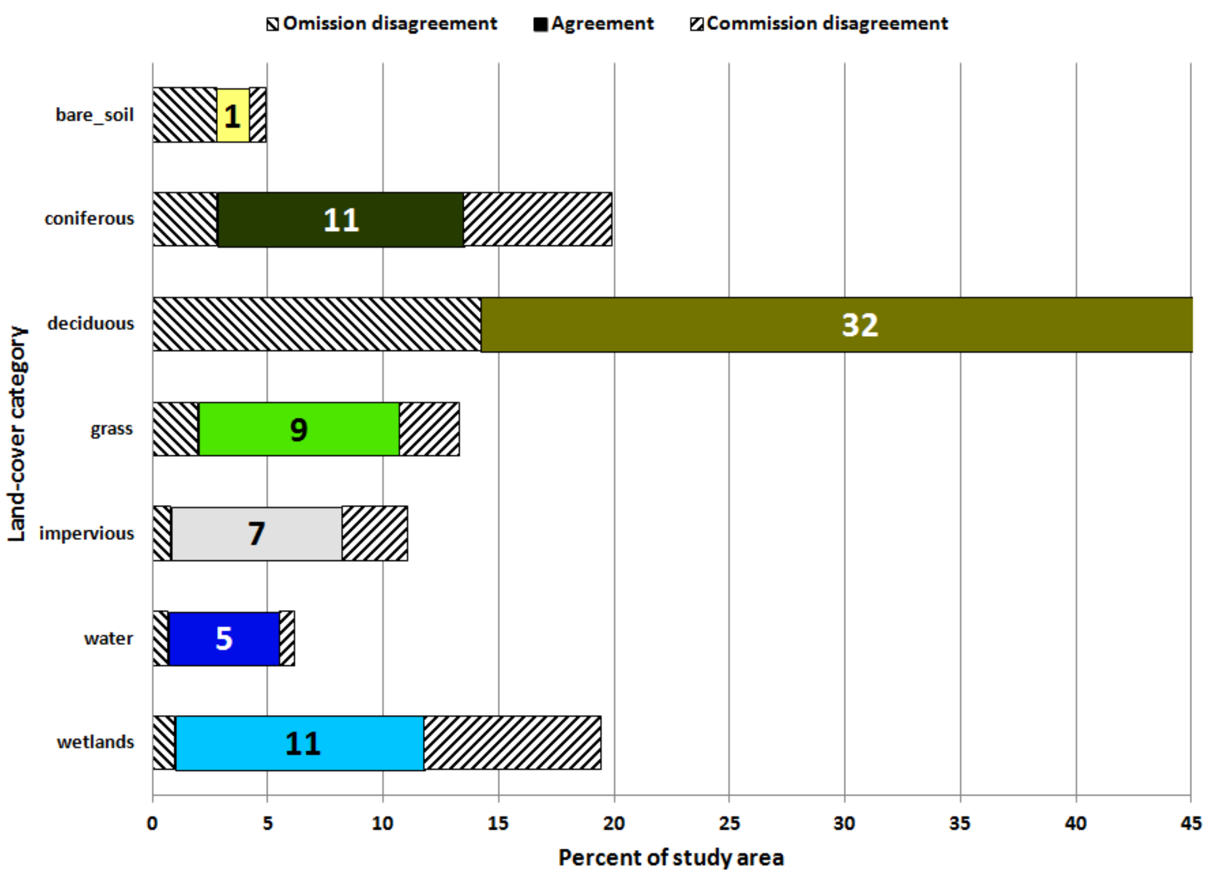


Figure 51. Figure of merit for the town of Middleton. The colored segments of the bars indicate the percentage of that land-cover category in the town correctly classified.

Table 16. Estimated population error matrix for the town of Newbury.

Newbury error matrix (entries in percent of the study area)									
	Reference map								
Land-cover map	Bare soil	Coniferous	Deciduous	Grass	Impervious	Water	Wetlands	Map total	Commission disagreement
Bare soil	**5.02**	0.00	0.07	0.87	0.15	0.00	0.06	6.17	1.15
Coniferous	0.00	**5.26**	4.03	0.02	0.03	0.27	0.00	9.61	4.35
Deciduous	0.15	0.18	**26.12**	1.04	0.10	0.02	1.14	28.75	2.63
Grass	1.14	0.12	0.68	**7.79**	0.06	0.00	0.00	9.78	1.99
Impervious	0.21	0.00	0.25	0.21	**4.72**	0.00	0.00	5.39	0.67
Water	0.20	0.00	0.13	0.01	0.00	**0.60**	0.20	1.14	0.54
Wetlands	1.61	0.00	4.73	0.00	0.26	1.56	**30.98**	39.15	8.17
Reference total	8.32	5.56	36.01	9.95	5.32	2.46	32.38	100.00	19.51
Omission disagreement	3.30	0.31	9.89	2.16	0.60	1.86	1.40	19.51	

Overall agreement = 80%

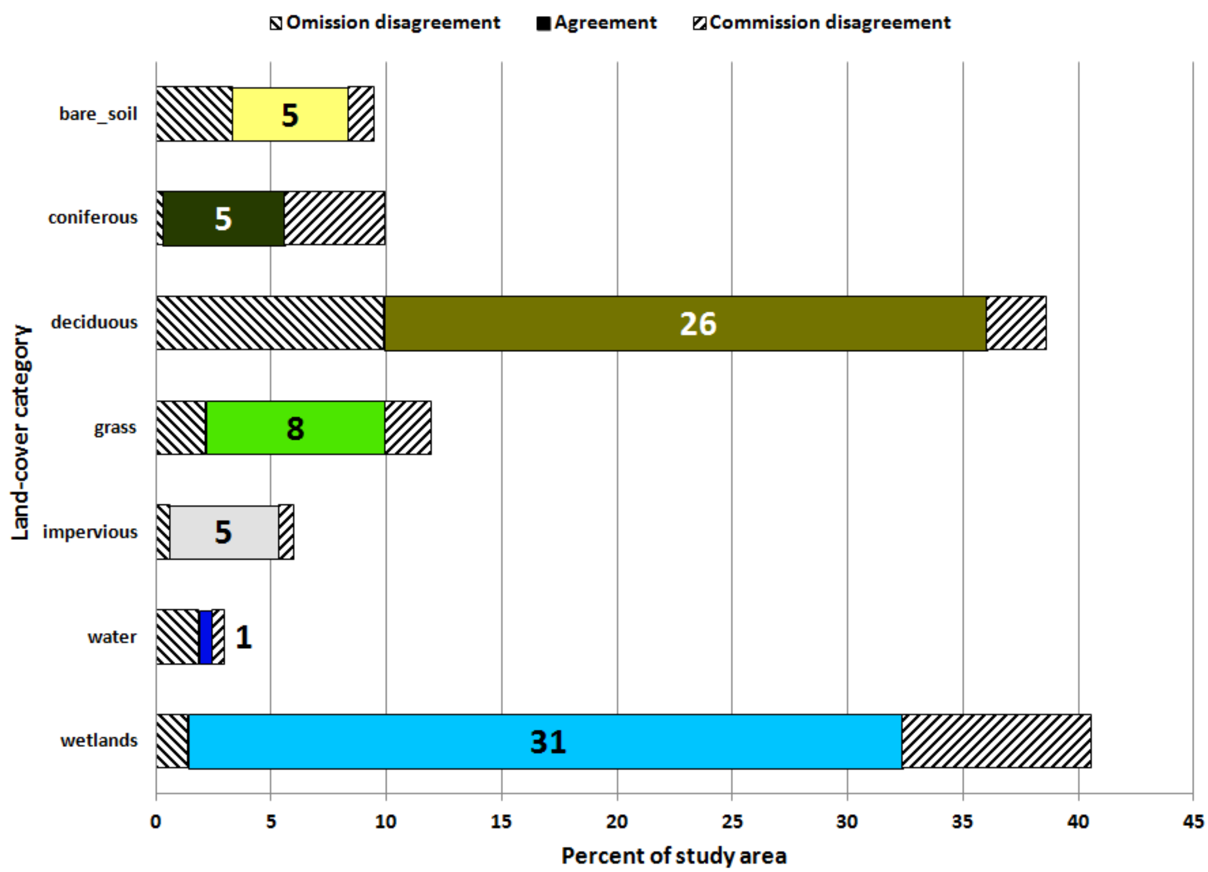


Figure 52. Figure of merit for the town of Newbury. The colored segments of the bars indicate the percentage of that land-cover category in the town correctly classified.

Table 17. Estimated population error matrix for the town of Newburyport.

Newburyport error matrix (entries in percent of the study area)									
	Reference map								
Land-cover map	Bare soil	Coniferous	Deciduous	Grass	Impervious	Water	Wetlands	Map total	Commission disagreement
Bare soil	**6.75**	0.04	0.65	1.09	0.40	0.00	0.02	8.95	2.20
Coniferous	0.02	**9.68**	1.24	0.07	0.02	0.00	0.00	11.02	1.34
Deciduous	0.23	0.27	**19.61**	5.17	1.24	0.00	0.00	26.53	6.92
Grass	2.65	0.58	1.01	**11.87**	0.87	0.00	0.00	16.99	5.11
Impervious	0.00	0.08	0.06	0.05	**20.73**	0.00	0.00	20.91	0.18
Water	0.03	0.04	0.06	0.04	0.00	**0.91**	0.20	1.29	0.38
Wetlands	2.21	0.40	2.74	0.03	0.00	1.11	**7.81**	14.30	6.50
Reference total	11.90	11.08	25.38	18.33	23.26	2.02	8.03	100.00	22.64
Omission disagreement	5.15	1.40	5.77	6.45	2.53	1.11	0.22	22.64	

Overall agreement = 77%

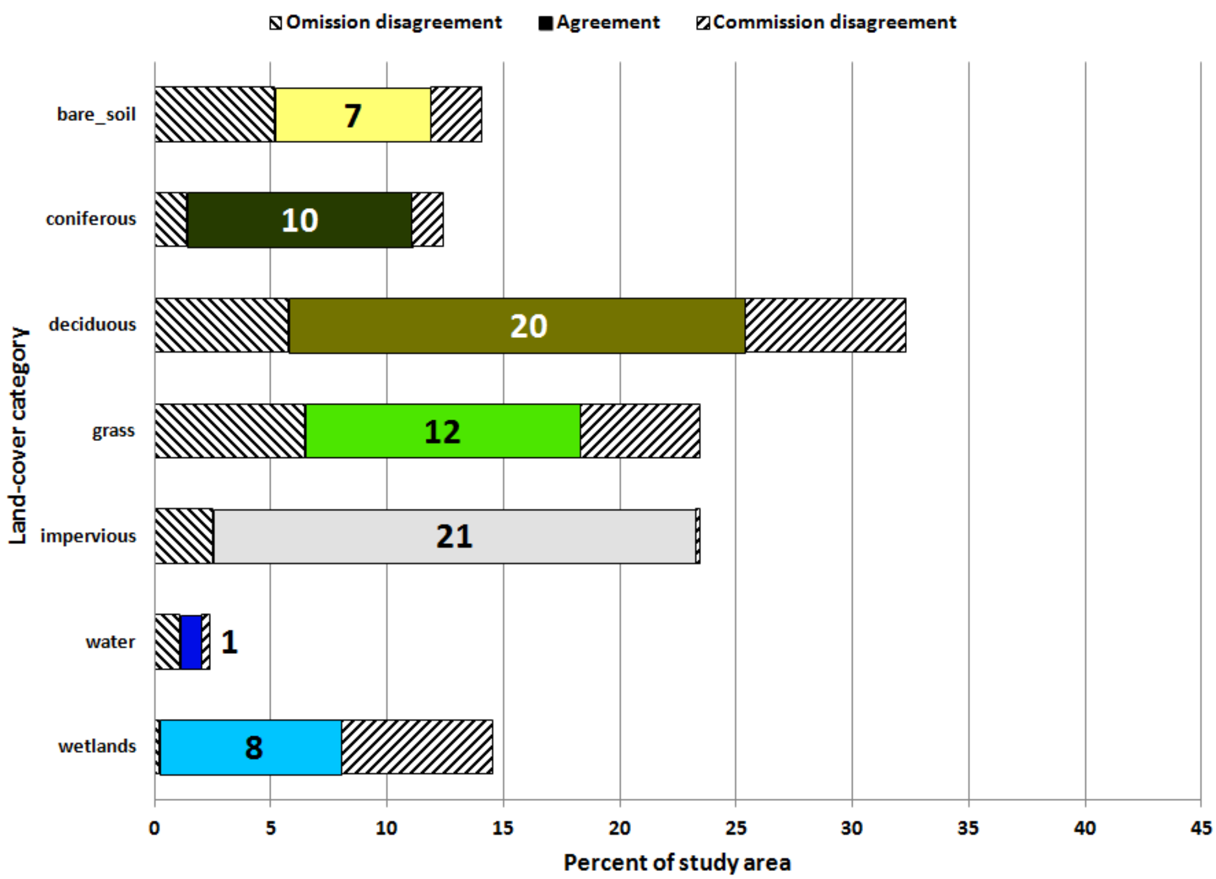


Figure 53. Figure of merit for the town of Newburyport. The colored segments of the bars indicate the percentage of that land-cover category in the town correctly classified.

Table 18. Estimated population error matrix for the town of North Andover.

North Andover error matrix (entries in percent of the study area)									
	Reference map								
Land-cover map	Bare soil	Coniferous	Deciduous	Grass	Impervious	Water	Wetlands	Map total	Commission disagreement
Bare soil	**0.63**	0.01	0.00	0.07	0.08	0.00	0.00	0.79	0.16
Coniferous	0.01	**8.37**	0.77	0.34	0.00	0.02	0.02	9.52	1.15
Deciduous	0.07	1.53	**36.52**	2.43	0.29	0.90	2.04	43.78	7.26
Grass	4.36	0.57	0.23	**7.75**	0.32	0.00	0.00	13.22	5.48
Impervious	0.17	0.01	0.27	0.13	**13.17**	0.00	0.00	13.75	0.57
Water	0.02	0.01	0.19	0.00	0.00	**4.51**	0.02	4.76	0.24
Wetlands	0.00	1.33	4.99	0.00	0.00	0.48	**7.38**	14.19	6.80
Reference total	5.24	11.82	42.98	10.72	13.87	5.92	9.46	100.00	21.66
Omission disagreement	4.62	3.45	6.46	2.97	0.69	1.41	2.08	21.66	

Overall agreement = 78%

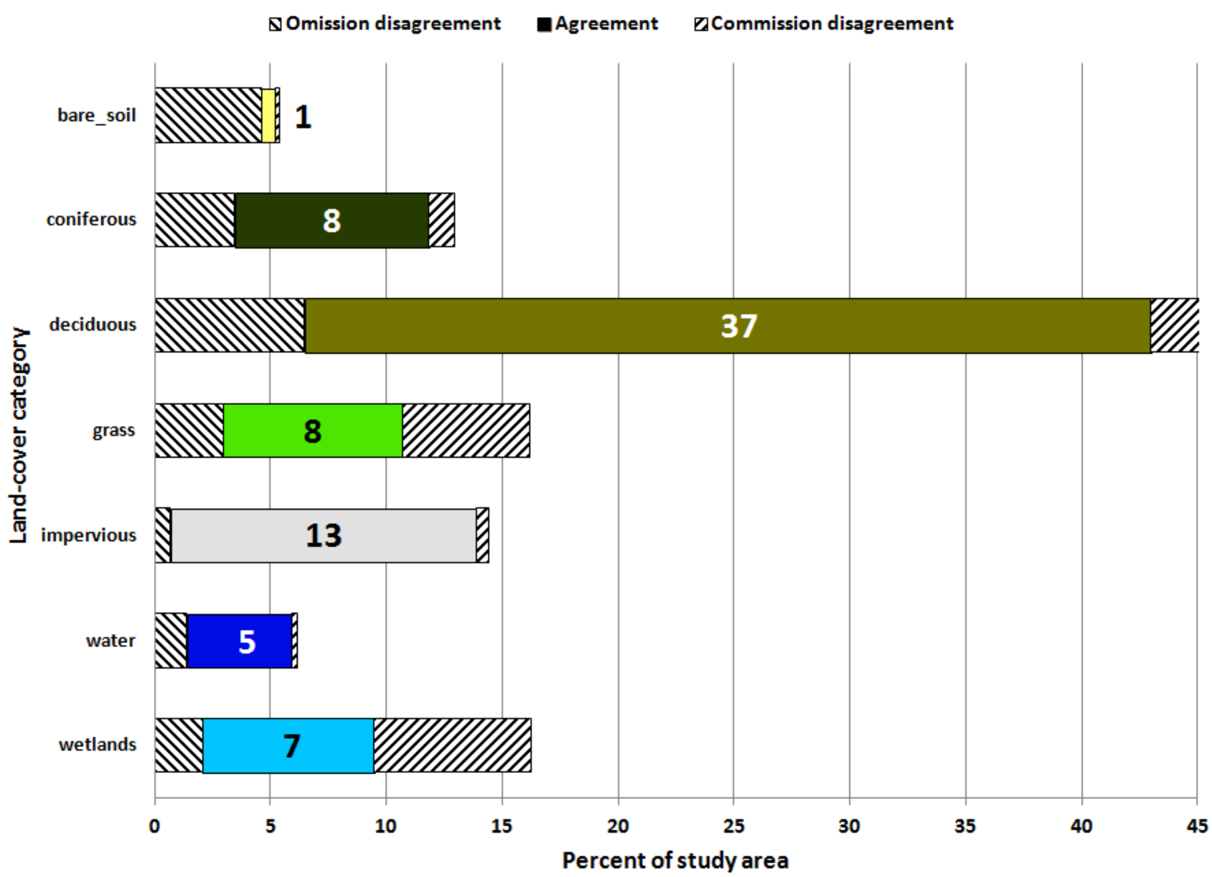


Figure 54. Figure of merit for the town of North Andover. The colored segments of the bars indicate the percentage of that land-cover category in the town correctly classified.

Table 19. Estimated population error matrix for the town of North Reading.

North Reading error matrix (entries in percent of the study area)									
	Reference map								
Land-cover map	Bare soil	Coniferous	Deciduous	Grass	Impervious	Water	Wetlands	Map total	Commission disagreement
Bare soil	**1.17**	0.12	0.07	0.72	0.05	0.00	0.68	2.81	1.64
Coniferous	0.01	**14.39**	2.67	0.63	0.16	0.00	0.37	18.22	3.83
Deciduous	0.53	1.38	**27.51**	1.37	0.09	0.00	0.00	30.88	3.37
Grass	0.60	0.08	0.69	**7.57**	0.34	1.47	0.00	10.75	3.18
Impervious	1.29	0.07	0.56	0.13	**14.27**	0.00	0.00	16.32	2.05
Water	0.00	0.07	0.07	0.02	0.00	**2.28**	0.43	2.86	0.58
Wetlands	0.13	0.61	6.68	0.20	0.01	0.46	**10.06**	18.16	8.09
Reference total	3.73	16.73	38.25	10.64	14.91	4.21	11.54	100.00	22.75
Omission disagreement	2.56	2.33	10.73	3.07	0.64	1.93	1.48	22.75	

Overall agreement = 77%

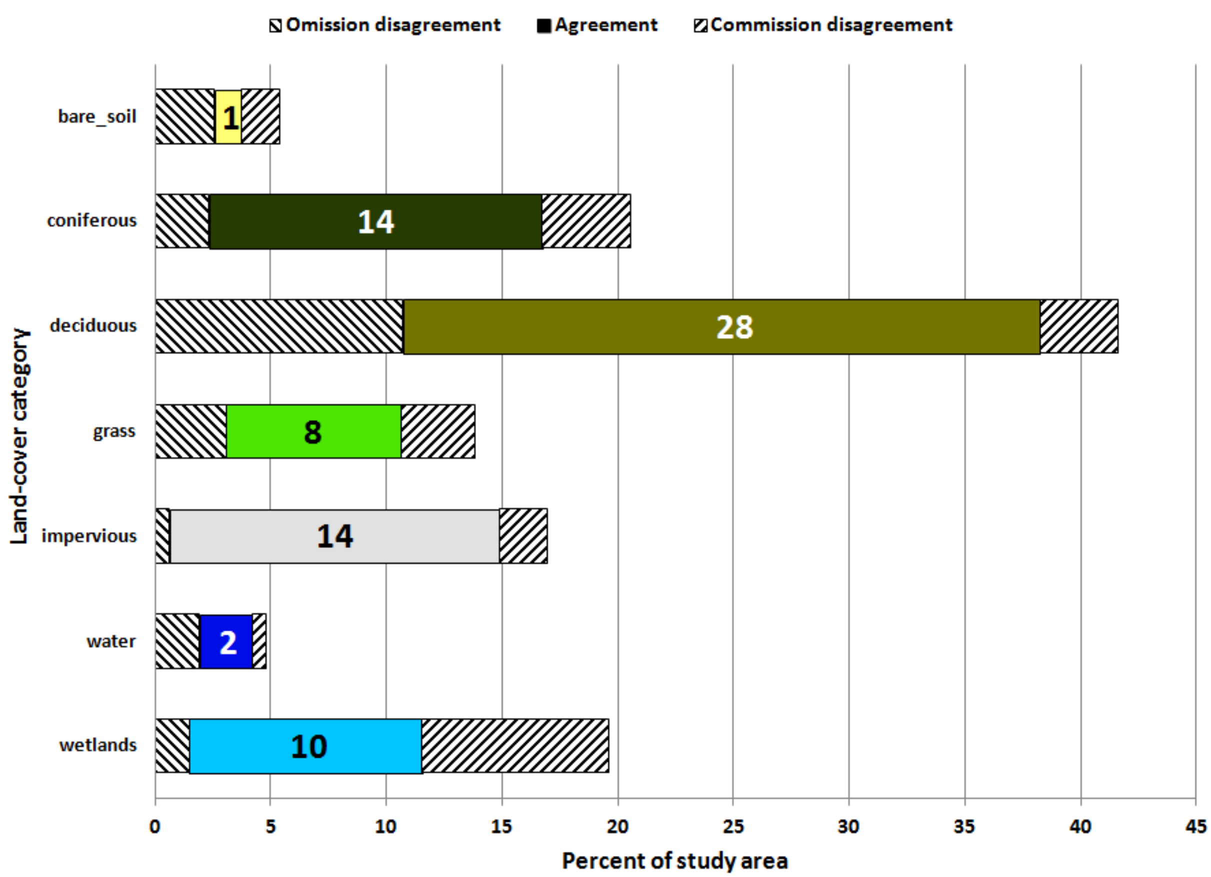


Figure 55. Figure of merit for the town of North Reading. The colored segments of the bars indicate the percentage of that land-cover category in the town correctly classified.

Table 20. Estimated population error matrix for the town of Peabody.

Peabody error matrix (entries in percent of the study area)									
	Reference map								
Land-cover map	Bare soil	Coniferous	Deciduous	Grass	Impervious	Water	Wetlands	Map total	Commission disagreement
Bare soil	**1.34**	0.00	0.31	1.09	0.47	0.00	0.00	3.21	1.87
Coniferous	0.00	**3.10**	2.43	0.36	0.33	0.05	0.00	6.27	3.17
Deciduous	0.06	0.17	**27.79**	1.20	0.64	0.00	0.00	29.86	2.07
Grass	0.04	0.02	1.55	**14.90**	2.28	0.00	0.00	18.78	3.89
Impervious	1.48	0.02	0.01	0.03	**29.94**	0.00	0.00	31.48	1.54
Water	0.00	0.02	0.07	0.01	0.00	**4.03**	0.06	4.19	0.16
Wetlands	0.26	0.04	2.44	0.07	0.00	0.25	**3.16**	6.21	3.06
Reference total	3.19	3.36	34.59	17.66	33.65	4.33	3.21	100.00	15.74
Omission disagreement	1.84	0.27	6.80	2.76	3.71	0.30	0.06	15.74	

Overall agreement = 84%

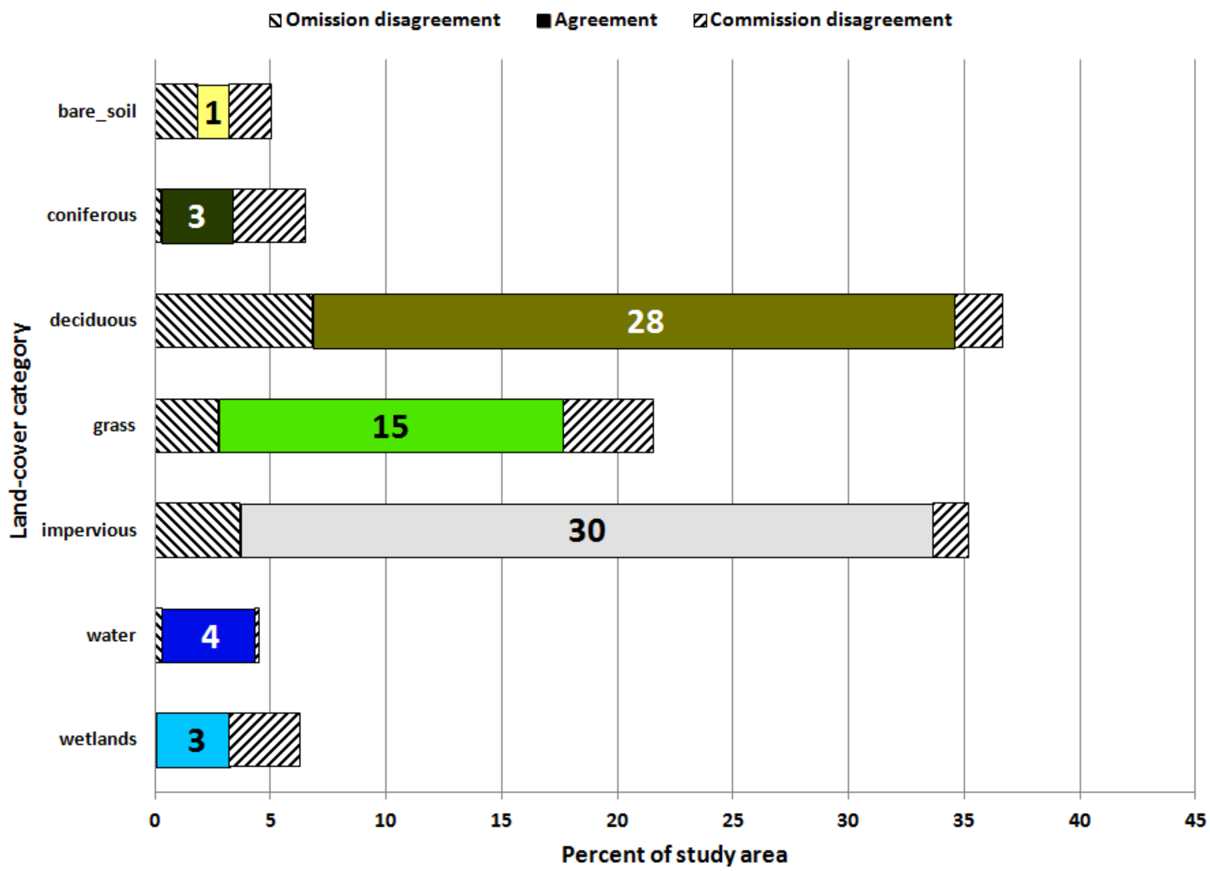


Figure 56. Figure of merit for the town of Peabody. The colored segments of the bars indicate the percentage of that land-cover category in the town correctly classified.

Table 21. Estimated population error matrix for the town of Reading.

Reading error matrix (entries in percent of the study area)									
	Reference map								
Land-cover map	Bare soil	Coniferous	Deciduous	Grass	Impervious	Water	Wetlands	Map total	Commission disagreement
Bare soil	**1.56**	0.00	0.35	1.67	0.45	0.00	0.00	4.02	2.47
Coniferous	0.00	**7.85**	1.42	1.07	0.20	0.00	0.00	10.55	2.69
Deciduous	0.00	2.05	**16.67**	2.94	0.10	0.00	0.00	21.77	5.10
Grass	0.43	0.43	3.70	**11.29**	0.76	0.00	0.00	16.60	5.31
Impervious	1.08	0.11	0.52	0.27	**19.09**	0.00	0.00	21.06	1.97
Water	0.02	0.01	0.02	0.00	0.00	**0.08**	0.05	0.18	0.10
Wetlands	0.00	0.95	13.77	0.02	0.00	0.00	**11.07**	25.82	14.75
Reference total	3.08	11.40	36.44	17.27	20.60	0.08	11.12	100.00	32.39
Omission disagreement	1.52	3.55	19.77	5.98	1.51	0.00	0.05	32.39	

Overall agreement = 68%

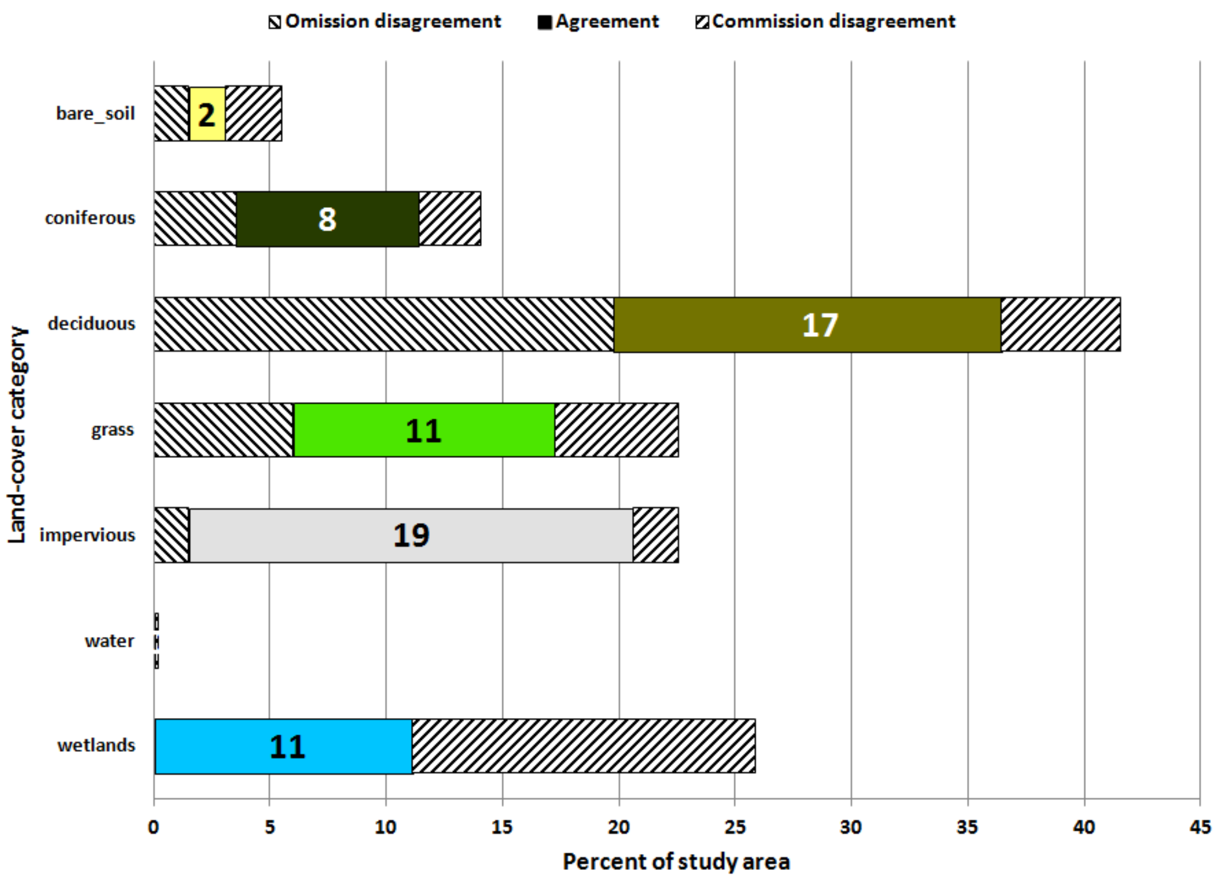


Figure 57. Figure of merit for the town of Reading. The colored segments of the bars indicate the percentage of that land-cover category in the town correctly classified.

Table 22. Estimated population error matrix for the town of Rowley.

Rowley error matrix (entries in percent of the study area)									
	Reference map								
Land-cover map	Bare soil	Coniferous	Deciduous	Grass	Impervious	Water	Wetlands	Map total	Commission disagreement
Bare soil	**5.61**	0.00	0.00	0.25	0.17	0.00	0.00	6.03	0.43
Coniferous	0.01	**12.83**	0.17	0.05	0.00	0.00	0.00	13.07	0.24
Deciduous	0.00	1.39	**32.81**	0.49	0.12	0.02	1.48	36.31	3.50
Grass	0.00	0.03	0.44	**6.18**	0.09	0.00	0.00	6.75	0.57
Impervious	0.00	0.04	0.20	0.22	**5.68**	0.00	0.00	6.15	0.47
Water	0.00	0.00	0.02	0.00	0.00	**0.78**	0.02	0.82	0.04
Wetlands	0.00	2.01	1.13	0.00	0.00	1.58	**26.16**	30.88	4.72
Reference total	5.62	16.31	34.79	7.19	6.07	2.38	27.65	100.00	9.97
Omission disagreement	0.01	3.48	1.98	1.01	0.39	1.60	1.50	9.97	

Overall agreement = 90%

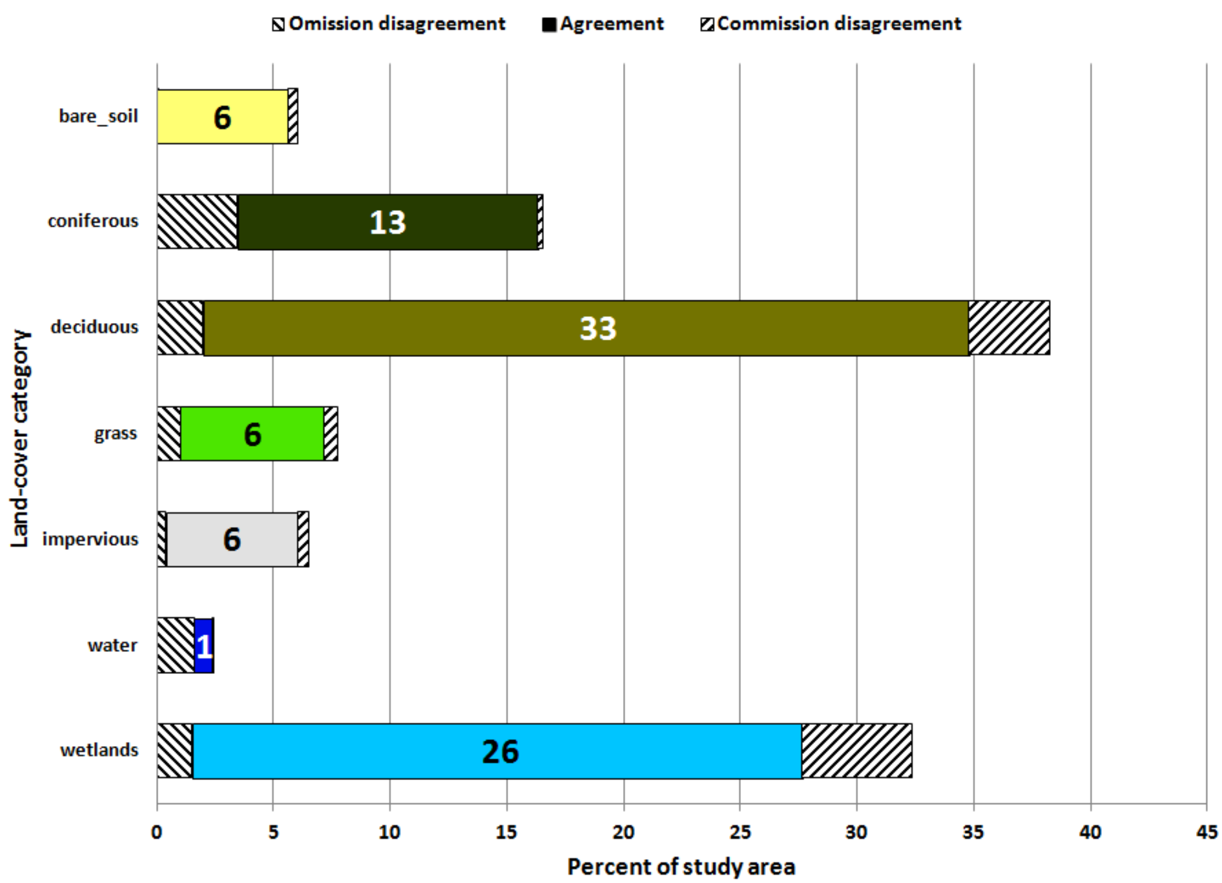


Figure 58. Figure of merit for the town of Rowley. The colored segments of the bars indicate the percentage of that land-cover category in the town correctly classified.

Table 23. Estimated population error matrix for the town of Tewksbury.

Tewksbury error matrix (entries in percent of the study area)									
	Reference map								
Land-cover map	Bare soil	Coniferous	Deciduous	Grass	Impervious	Water	Wetlands	Map total	Commission disagreement
Bare soil	**1.21**	0.04	0.66	0.56	0.25	0.00	0.00	2.73	1.51
Coniferous	0.00	**11.63**	1.24	0.56	0.09	0.00	0.01	13.53	1.90
Deciduous	0.67	1.41	**22.63**	3.62	0.34	0.00	0.23	28.90	6.28
Grass	0.88	0.42	1.28	**14.53**	0.33	0.00	0.01	17.44	2.92
Impervious	0.99	0.10	0.16	0.28	**15.77**	0.00	0.00	17.31	1.54
Water	0.00	0.01	0.03	0.00	0.00	**1.74**	0.07	1.86	0.12
Wetlands	0.48	0.65	4.78	0.31	0.01	1.52	**10.48**	18.22	7.75
Reference total	4.24	14.26	30.78	19.87	16.79	3.26	10.80	100.00	22.01
Omission disagreement	3.02	2.63	8.15	5.34	1.02	1.52	0.33	22.01	

Overall agreement = 78%

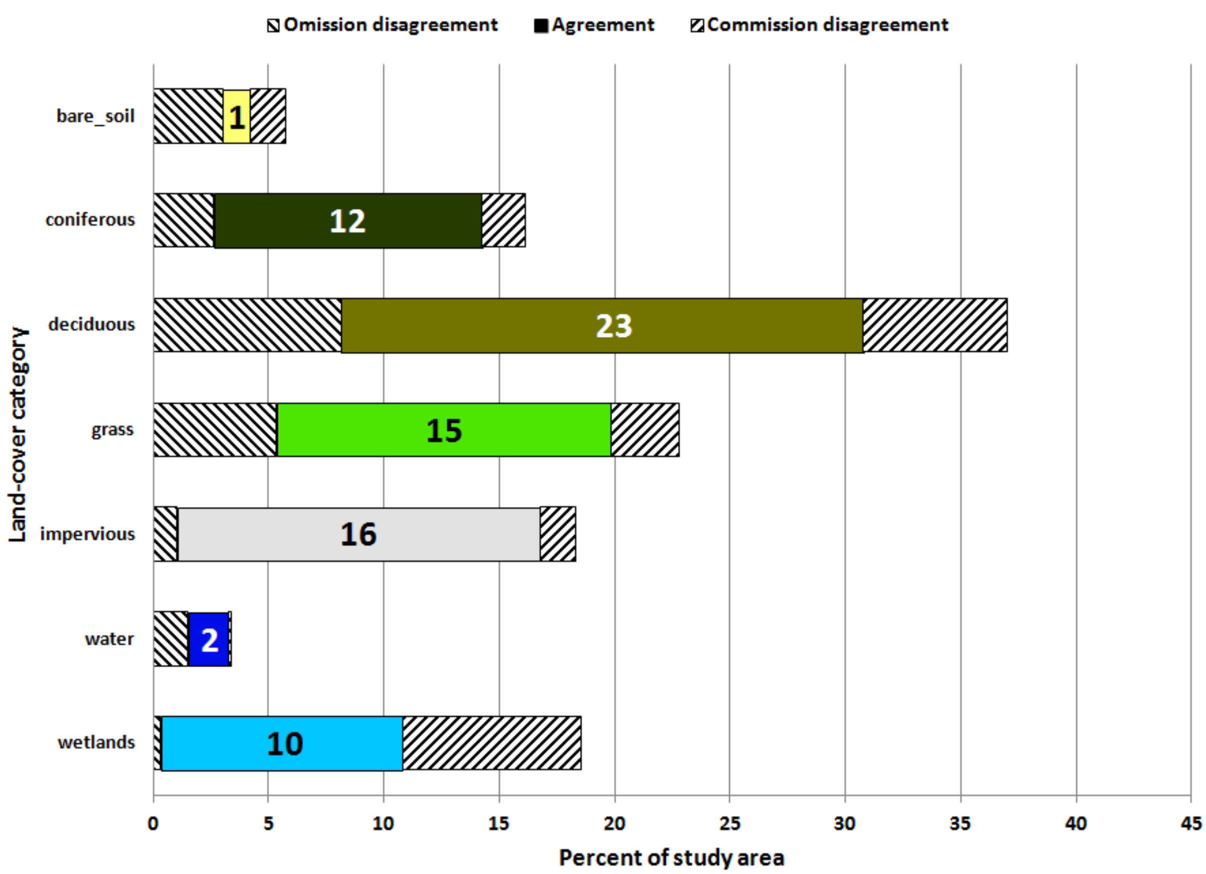


Figure 59. Figure of merit for the town of Tewksbury. The colored segments of the bars indicate the percentage of that land-cover category in the town correctly classified.

Table 24. Estimated population error matrix for the town of Topsfield.

Topsfield error matrix (entries in percent of the study area)									
	Reference map								
Land-cover map	Bare soil	Coniferous	Deciduous	Grass	Impervious	Water	Wetlands	Map total	Commission disagreement
Bare soil	**0.85**	0.00	0.09	0.15	0.00	0.00	0.00	1.10	0.25
Coniferous	0.00	**11.17**	3.81	0.00	0.05	0.00	0.00	15.03	3.86
Deciduous	0.54	1.52	**34.60**	1.23	0.48	0.00	0.03	38.41	3.81
Grass	1.79	0.09	1.62	**12.61**	0.46	0.00	0.00	16.56	3.95
Impervious	0.00	0.00	0.12	0.18	**7.80**	0.00	0.00	8.11	0.31
Water	0.00	0.01	0.11	0.01	0.00	**1.59**	0.49	2.22	0.63
Wetlands	0.18	0.56	2.42	0.00	0.00	0.01	**15.38**	18.57	3.18
Reference total	3.37	13.36	42.78	14.19	8.79	1.61	15.91	100.00	15.99
Omission disagreement	2.52	2.19	8.17	1.58	0.99	0.01	0.53	15.99	

Overall agreement = 84%

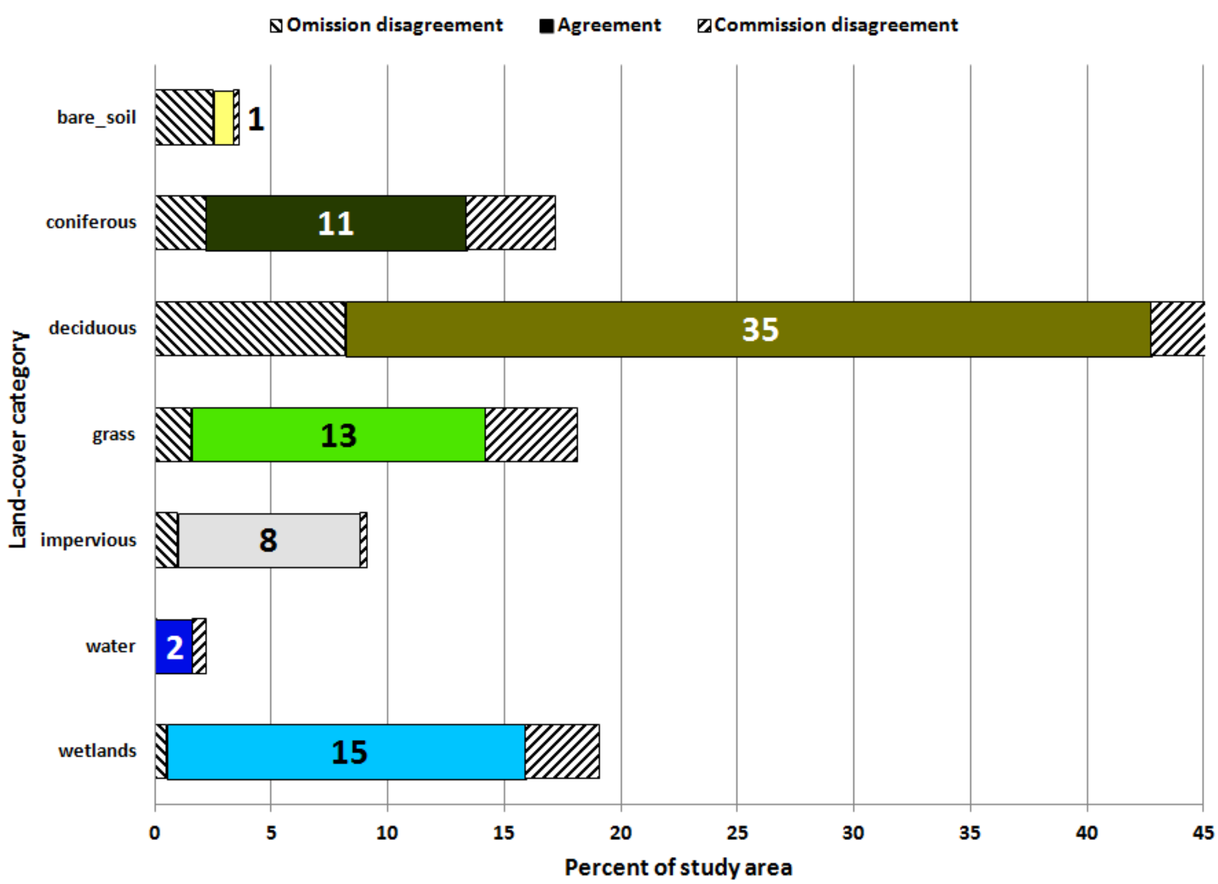


Figure 60. Figure of merit for the town of Topsfield. The colored segments of the bars indicate the percentage of that land-cover category in the town correctly classified.

Table 25. Estimated population error matrix for the town of Wenham.

Wenham error matrix (entries in percent of the study area)									
	Reference map								
Land-cover map	Bare soil	Coniferous	Deciduous	Grass	Impervious	Water	Wetlands	Map total	Commission disagreement
Bare soil	**0.48**	0.00	0.04	0.03	0.00	0.00	0.00	0.56	0.07
Coniferous	0.05	**13.75**	1.26	0.14	0.01	0.00	0.00	15.20	1.45
Deciduous	0.81	1.66	**26.39**	2.21	0.39	0.00	1.18	32.65	6.26
Grass	3.76	0.34	0.80	**8.57**	0.24	0.00	0.00	13.70	5.14
Impervious	1.70	0.02	0.30	0.22	**5.71**	0.00	0.00	7.95	2.24
Water	0.01	0.00	0.03	0.00	0.00	**5.70**	0.00	5.74	0.04
Wetlands	0.17	1.79	12.64	0.11	0.42	1.02	**8.06**	24.20	16.14
Reference total	6.98	17.56	41.46	11.28	6.76	6.72	9.24	100.00	31.35
Omission disagreement	6.50	3.81	15.07	2.71	1.05	1.02	1.18	31.35	

Overall agreement = 69%

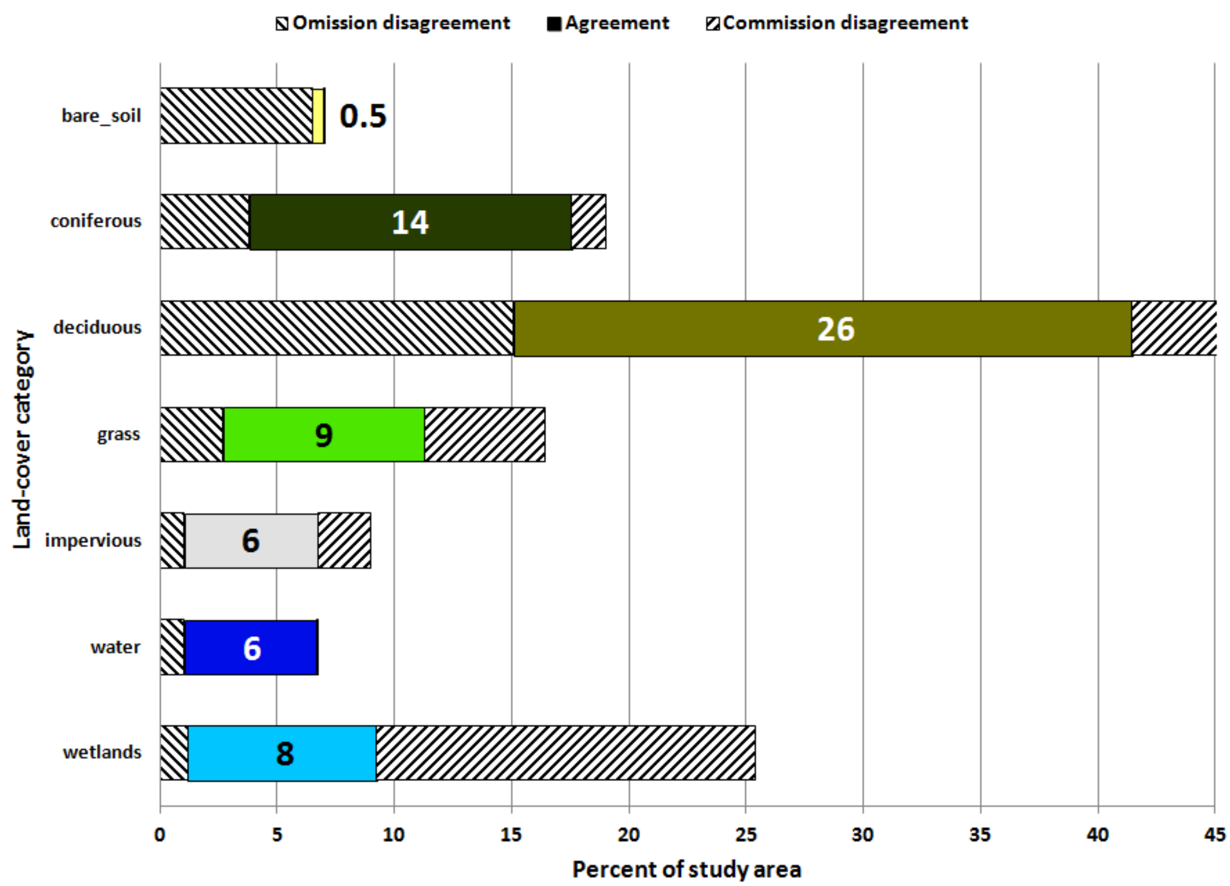


Figure 61. Figure of merit for the town of Wenham. The colored segments of the bars indicate the percentage of that land-cover category in the town correctly classified.

Table 26. Estimated population error matrix for the town of West Newbury.

	West Newbury error matrix (entries in percent of the study area)								
	Reference map								
Land-cover map	Bare soil	Coniferous	Deciduous	Grass	Impervious	Water	Wetlands	Map total	Commission disagreement
Bare soil	**3.03**	0.00	0.06	0.73	0.02	0.02	0.00	3.86	0.83
Coniferous	0.00	**4.48**	1.56	0.34	0.02	0.00	0.06	6.45	1.98
Deciduous	0.21	0.74	**40.50**	1.66	0.11	0.18	1.55	44.96	4.46
Grass	1.81	0.05	1.16	**12.45**	0.21	0.00	0.00	15.68	3.23
Impervious	0.09	0.00	0.07	0.04	**5.16**	0.00	0.00	5.35	0.19
Water	0.00	0.03	0.22	0.01	0.01	**7.86**	0.27	8.40	0.54
Wetlands	0.52	0.08	3.96	0.00	0.00	0.41	**10.32**	15.29	4.97
Reference total	5.66	5.39	47.52	15.23	5.53	8.47	12.20	100.00	16.20
Omission disagreement	2.63	0.91	7.02	2.78	0.37	0.61	1.88	16.20	

Overall agreement = 84%

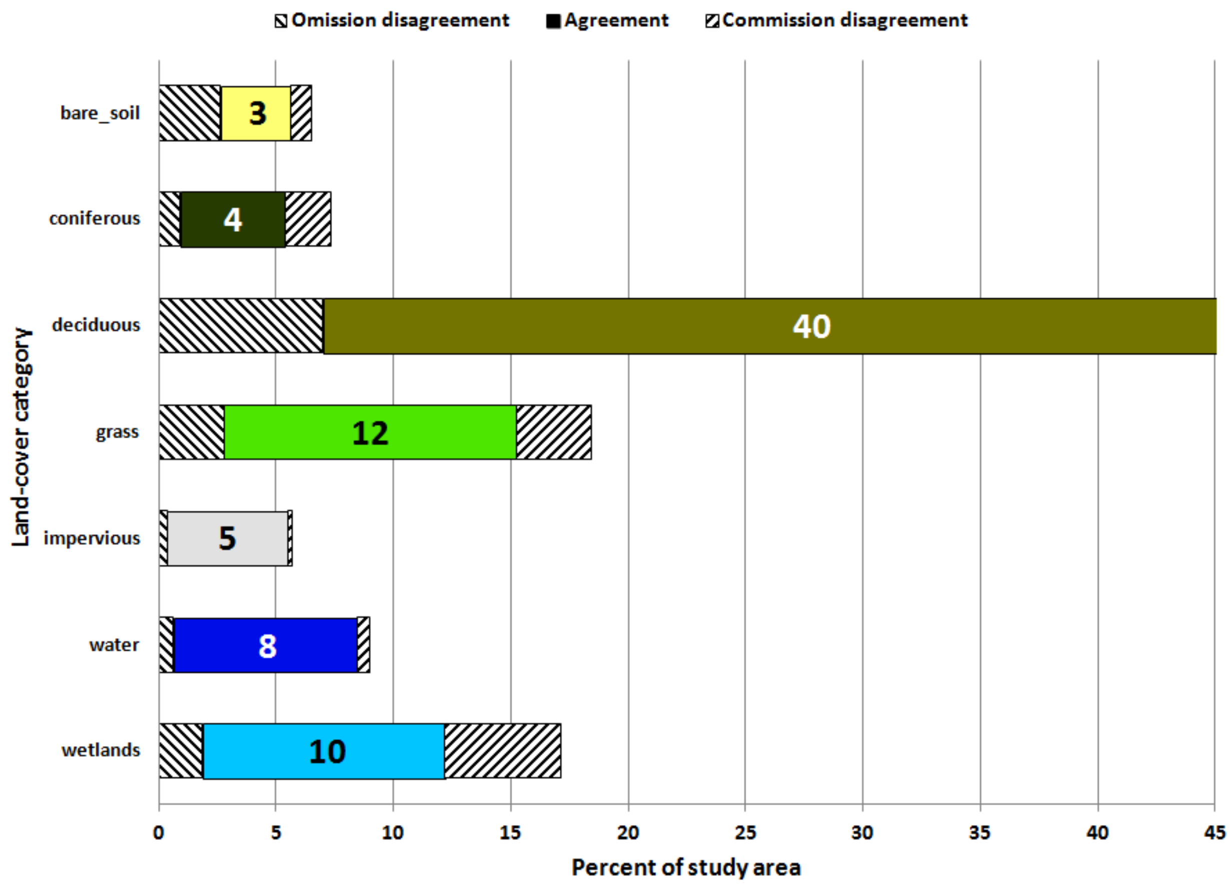


Figure 62. Figure of merit for the town of West Newbury. The colored segments of the bars indicate the percentage of that land-cover category in the town correctly classified.

Table 27. Estimated population error matrix for the town of Wilmington.

Wilmington error matrix (entries in percent of the study area)									
	Reference map								
Land-cover map	Bare soil	Coniferous	Deciduous	Grass	Impervious	Water	Wetlands	Map total	Commission disagreement
Bare soil	**0.70**	0.01	0.23	0.11	0.01	0.00	0.00	1.06	0.35
Coniferous	0.00	**10.36**	2.71	0.26	0.10	0.60	0.09	14.12	3.75
Deciduous	0.33	0.73	**26.08**	2.11	0.20	0.00	0.00	29.45	3.37
Grass	1.00	0.08	0.71	**11.50**	0.28	0.00	0.00	13.57	2.06
Impervious	0.78	0.02	0.39	0.16	**20.89**	0.00	0.00	22.23	1.34
Water	0.00	0.00	0.05	0.02	0.00	**0.93**	0.28	1.29	0.36
Wetlands	0.00	1.41	6.51	0.02	0.00	0.29	**10.06**	18.29	8.23
Reference total	2.81	12.61	36.68	14.18	21.47	1.82	10.43	100.00	19.47
Omission disagreement	2.11	2.25	10.59	2.68	0.58	0.89	0.37	19.47	

Overall agreement = 81%

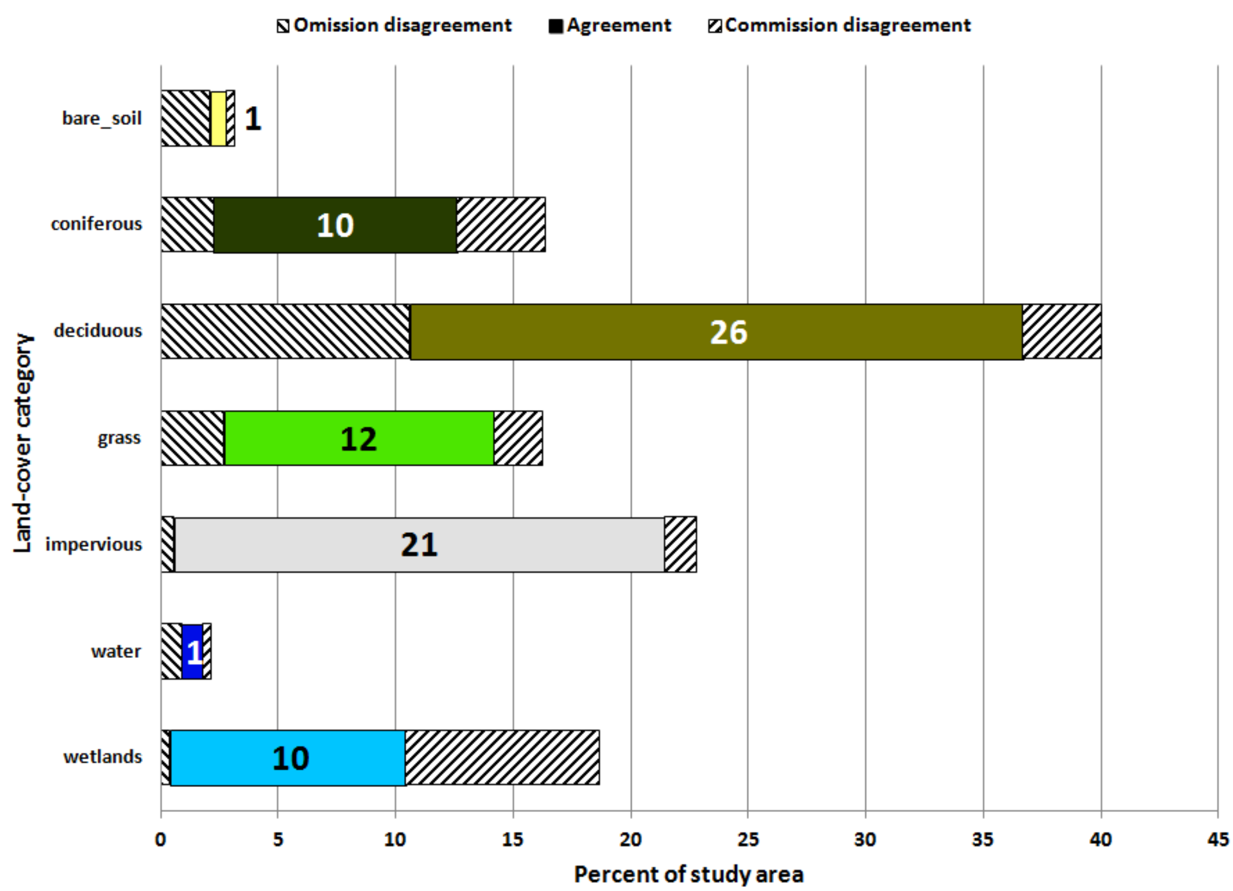


Figure 63. Figure of merit for the town of Wilmington. The colored segments of the bars indicate the percentage of that land-cover category in the town correctly classified.

Table 28. Estimated population error matrix for the town of Woburn.

Woburn error matrix (entries in percent of the study area)									
	Reference map								
Land-cover map	Bare soil	Coniferous	Deciduous	Grass	Impervious	Water	Wetlands	Map total	Commission disagreement
Bare soil	**0.36**	0.00	1.42	1.19	3.43	0.00	0.00	6.39	6.03
Coniferous	0.00	**3.57**	3.20	0.76	0.11	0.00	0.00	7.64	4.07
Deciduous	0.00	0.27	**23.06**	2.65	0.34	0.17	0.00	26.49	3.43
Grass	0.20	0.12	2.24	**12.72**	0.64	0.00	0.00	15.91	3.20
Impervious	0.50	0.31	0.87	0.18	**33.13**	0.00	0.00	34.98	1.86
Water	0.00	0.00	0.10	0.00	0.00	**1.95**	0.18	2.24	0.29
Wetlands	0.00	0.03	3.39	0.00	0.03	0.12	**2.77**	6.34	3.57
Reference total	1.06	4.30	34.28	17.49	37.68	2.24	2.96	100.00	22.44
Omission disagreement	0.70	0.73	11.22	4.77	4.55	0.29	0.18	22.44	

Overall agreement = 78%

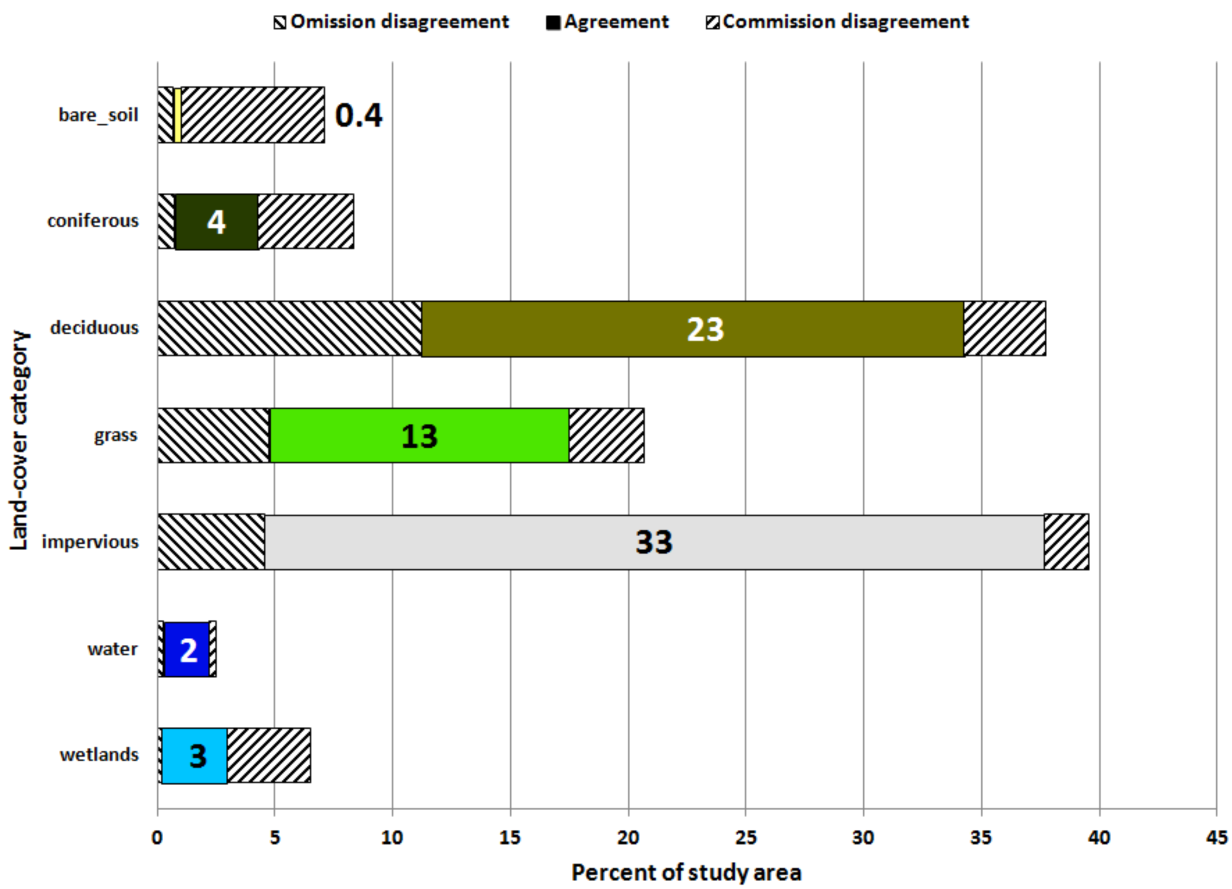


Figure 64. Figure of merit for the town of Woburn. The colored segments of the bars indicate the percentage of that land-cover category in the town correctly classified.

References

Anderson, J. R., Hardy, E. E., Roach, J. T., Witmer, R. E., 1976. A land use and land cover classification system for use with remote sensor data. *Geological Survey Professional Paper 964,* USGS, Reston, VA.

Giner, N.M., 2013. Using GIS and remotely sensed data to validate, analyze, and predict lawn land-cover maps: A case study in suburban Boston, Massachusetts, USA. Unpublished PhD Dissertation, Geography, Clark University, Worcester, Massachusetts.

MassGIS, 2011. Massachusetts Geographic Information Systems office. http://www.mass.gov/mgis/. Accessed December 16, 2011.

Milesi, C., S. W. Running, C. D. Elvidge, J. B. Dietz, B. T. Tuttle, and R. R. Nemani. 2005. Mapping and Modeling the Biogeochemical Cycling of Turf Grasses in the United States. . Environmental Management 36:426-438.

Milesi, C., C. Elvidge, and R. R. Nemani. 2009. Assessing the Extent of Urban Irrigated Areas in the United States. In Remote Sensing of Global Croplands for Food Security, 217-236.

Plum Island Ecosystems LTER, 2012. URL: http://ecosystems.mbl.edu/pie/. Last accessed: 5 December 2012.

Polsky, C., Rogan, J., Pontius, Jr., R.G., Turner, B.L., 2007. Undergraduate GIScience Research at Clark University: The HERO Program. *Council on Undergraduate Research Quarterly* 27 (3):124-129.

Pontius, Jr., R.G., Boersma, W., Castella, J.C., Clarke, K., de Nijs, T., Dietzel, C., Duan, Z., Fotsing, R., Goldstein, N., Kok, K., Koomen, E., Lippitt, C.D., McConnell, W., Mohd Sood, A., Pijanowski, B., Pithadia, S., Sweeney, S., Ngoc Trung, T., Veldkamp, A.T., Verburg, P.H., 2008. Comparing the input, output, and validation maps for several models of land change. *Annals of Regional Science*, 42: 11-37.

Pontius Jr., R.G., Millones, M., 2011. Death to kappa: Birth of quantity disagreement and allocation disagreement for accuracy assessment. *International Journal of Remote Sensing,* 32: 4407-4429.

Zhou, W., Troy, A.R., Grove, M. 2008. Object-based land cover classification and change analysis in the Baltimore Metropolitan area using multitemporal high resolution remote sensing data. *Sensors*. 8: 1613-1636.

Zhou, W., Troy, A.R., 2008. An object-oriented approach for analysing and characterizing urban landscape at the parcel level. *International Journal of Remote Sensing*. 29 (11): 3119-3135.

Appendix

holmes understanding of the landscape 67
Choice of land-cover categories 67
Meaning of land-cover categories 67
Conceptions of the landscape 67
Pre-processing 68
Input data 68
Quality assurance 70
Object-based image analysis 70
What is OBIA? 70
How to train users – or yourself 71
holmes OBIA ruleset 73
Post-processing 74
Accuracy assessment 75
How to Create Samples 75
How to Install Prerequisites 75
Enumerator's Task 76
How to make the error matrix 77
Data backup 77
SQL queries 77
Delivering our products 78
Glossary 79

holmes understanding of the landscape

Choice of land-cover categories

The land-cover categories were chosen to capture land-cover variation in vegetation life forms and human impacts in a way that mimics Anderson Level I classifications (Anderson, 1976) on a suburban, rather than continental scale. For instance, agriculture and tundra do not appear in our classification of suburban Boston. While agriculture is present on the landscape, it is characterized as either fine green grass or bare soil depending on its spectral characteristics. The decision not to include agriculture was influenced by this study's focus on the suburban environment and on land cover rather than land use. Impervious surfaces are relevant because they have a distinct and quantifiable influence on storm-water runoff. The *holmes* land-cover categories are specific to the study area and may not be applicable elsewhere.

Meaning of land-cover categories

- Bare soil
  - Exposed areas of rock or soil that are devoid of vegetation, including overhanging tree canopy. Examples include quarries, unpaved parking lots, and construction sites. Dead or fallow grasses are also included.
- Coniferous
  - Coniferous tree canopy when it does not cover impervious surfaces.
- Deciduous
  - Deciduous tree canopy when it does not cover impervious surfaces.
- Impervious
  - Any paved or roofed area through which water cannot easily penetrate. Examples include roads, sidewalks, paved parking lots, driveways, and any rooftop. This category includes surfaces covered by tree canopy when possible.
- Fine green grass
  - Healthy grassy cover. This includes lawn grass, agricultural grass, cemeteries, parks, schools, athletic fields, and any place else that grass is found.
- Water
  - Deep bodies of open water that are not wetlands.
- Wetlands
  - Areas defined by the Massachusetts Department of Environmental Protection (DEP) (MassGIS, 2011) as wetlands that do not have impervious cover over them.

Conceptions of the landscape

Our classification is synoptic, meaning every part of the study area is characterized by a pixel which has an associated land cover. However, our classification is not volumetric, meaning it does not concern itself with 3D spaces, and it also does not have a Z dimension, meaning it does not concern itself with altitude. Because we describe the landscape as a single flat grid of pixels we are necessarily missing some of the richness with which that landscape might be characterized. One challenge is that when a remote sensing instrument, in this case a camera mounted on a plane, captures an image, it is viewing the Earth from the top-down. It cannot see beneath what is covered up. In the case of some sensors, it is possible to determine what lies beneath the tree canopy or even the soil. However, the aerial photographs we use are incapable of seeing through landscape features. We use the following rules to handle conflicts between potential descriptions of the landscape that result from a top-down view:

- Water takes precedence over all other land covers
  - We want to keep in mind the difference between land and water areas
- Impervious surfaces take precedence over tree canopy and wetlands
- Tree canopy takes precedence over most land-cover categories most of the time – just because the sensing instruments cannot see beneath it
  - Most importantly, the sensors cannot see beneath canopy to lawns
- Shadows are assumed to be the same land cover as the features that they share the most border with
  - For example, shadows along the edge of a house are assumed to be impervious, while shadows along the edge of a tree are assumed to be tree canopy
  - However, frequently during the manual editing of image classifications, more careful considerations were taken into account by human beings, who are more adept at image interpretation than computers. Our team engaged in many hundreds of hours of manual edits.

Ultimately this design was the result of being more interested in fine green grass – meaning lawns and other grasses – than in tree canopy. Note that our project may suffer from an underestimation of lawns because of tree canopy overlapping grasses in homeowner's yards. There is no way to know the exact amount of lawn given the data available, but we can guess at reasons we may have missed some lawns.

Pre-processing

Imagery and ancillary data must be prepared before they can be used in OBIA to produce a land-cover classification. This is known as preprocessing, and requires the following steps:

1. Find data sources for the study area
2. Choose appropriate datasets to aid in classification
   a. Ensure relevance
   b. Determine quality and consistency
      i. Make changes if necessary and feasible given project needs and resources
         1. Note that this is possible in an OBIA environment as additional rules before classification
3. If useful, produce derivative datasets from existing data or combinations of existing data

Input data

All of our input datasets came from MassGIS, the state GIS agency of Massachusetts (MassGIS, 2011). A dataset used to aid classification is referred to as a thematic layer. This is the terminology used by eCognition. How these datasets were used in OBIA will be described in greater detail in the next section. We used the following data to aid in our classification.

- Aerial photographs
  - Orthorectified, 0.5 m spatial resolution, R,G,B, NIR
  - The image used by OBIA
  - http://www.mass.gov/mgis/colororthos2005.htm
- Massachusetts Department of Environmental Protection (DEP) wetlands
  - Water and wetlands
  - A thematic layer used by OBIA
  - http://www.mass.gov/mgis/wetdep.htm
- Land-use 2005
  - Forest and non-forest
  - A thematic layer used by OBIA
  - http://www.mass.gov/mgis/lus2005.htm
- Impervious surfaces
  - Impervious and non-impervious surfaces
  - A thematic layer used by OBIA
  - http://www.mass.gov/mgis/impervious_surface.htm
- Massachusetts Department of Transportation (DOT) roads
  - To help make grid boundaries by hand
  - Finished grid boundaries were used in OBIA
  - http://www.mass.gov/mgis/eotroads.htm

Changes sometimes needed to be made by hand before these data could be used. The water and wetlands datasets were examined and corrected by hand in order to include small bodies of water, remove wetland patches located in conspicuous places, and correct water body borders for larger errors. All changes were made with the help of the 0.5 m aerial photographs. While such edits are at the discretion of the editor, quality assurance was used to ensure that editors worked in a similar fashion.

Derivative datasets were produced from existing input datasets. Our OBIA techniques could only be run on small subsets of the landscape at a time, which we referred to as grids. In order to avoid an edge matching problem when the grids were merged back together, we decided to make the edges of all grids have the same land cover on both sides. Grids are named by town and then assigned a number. The most natural choice for a grid-boundary in our suburban study area was to use roads. The DOT roads layer and the impervious surfaces layer were used to find appropriate boundaries for these grids. However, in more rural parts of the study area the landscape was nearly devoid of roads so it became necessary to use wetland and water centerlines as boundaries.

Note that the data available on the MassGIS website changes frequently. Not only are new layers being added, but existing layers are being updated. For this reason our inputs may be slightly different from what layers are available today.

Quality assurance

As the project progresses, quality assurance (QA) is extremely important. It is also important to make sure that each person's work is checked over by someone else, and that any concerns are brought up. Sometimes work does have to be redone, but the faster a mistake is caught the less time will be wasted. How to conduct a QA:

1. Ensure that all tasks are understood as discrete steps
2. Each step is assigned to a different person
3. As the steps are completed, they are reviewed by a person who did not do the work under examination
4. Ensure that a single standard for quality is understood by all
    a. This should be understood in terms of an acceptable minimum of quality, and also as a limit above which it is unnecessary to check quality
        i. No work is perfect, and at some point imperfections are so minor that they must be tolerated for the project to proceed
    b. Making sure a single standard is understood is easy to do through pictures of good and bad examples, and through training people to assess quality and reviewing their example assessments
5. Be sure to mix up who checks who's work
6. Avoid creating an atmosphere of resentment between employees by stressing the importance of constructive assessments

When the assessor is not pleased with another team member's work it is important that that work be redone. However, it may not be necessary to redo the entire task; rather, the team member can find ways to fix only the problem that occurred. This is especially easy in the OBIA environment. For instance, if the problem was that small dark objects with a high relational border to houses are classified as coniferous trees when they are actually shadows, the assessor can go back to the OBIA project and make the appropriate rule. If the rule performs well enough and the problem it solves is seen frequently, then the rule should be added to the official eCognition ruleset.

Object-based image analysis

What is OBIA?

The land-cover maps have been made using object-based image analysis (OBIA), a relatively new mapping technique. It is distinguished by using a set of rules custom-made by an expert. These rules operate on groups of pixels which can be created and modified with a variety of algorithms. For example, OBIA can classify pixels by asking questions about the relative border between one group of pixels from another, or whether one group of pixels is surrounded by the other, or the distance between the groups of pixels. Standard GIS classifying procedures are also available as rules which can be incorporated by the expert, such as vegetation indices or the reflectance values of any of the bands available in the imagery. Note that OBIA still views these qualities as belonging to the group of pixels, known as the object, rather than to individual pixels. However, it is also possible to use pixel-based classifications as an input to OBIA, or even to perform some operations in your classification on the pixel rather than the object-level. Note that the outputs of an OBIA classification are standard GIS products –

the output can still be, for instance, a raster that uses pixels to describe the attributes of the landscape. The software used in our work was eCognition, which used to be known by the name Definiens.

How to train users – or yourself

This guide assumes that the OBIA trainee knows pixel-based classification techniques and basic GIS. OBIA can also be executed in many platforms, and could be performed very painstakingly in a traditional GIS. This guide assumes the use of eCognition, which is currently the richest toolset for OBIA.

The important difference between OBIA and pixel-based classifiers is that OBIA understands the image as being composed on objects. These objects are in turn composed of pixels. It is also possible for objects to exist in a series of hierarchical levels, in which there are sub- and super-objects. The properties of the pixels in the objects, and potentially sub-objects in super-objects, are what allow for OBIA algorithms to make decisions in classification. There are hundreds of such properties used by eCognition, and potentially many more that you could define yourself. The following is a list of the kinds of properties that an object can have:

- Spectral
    - Mean band values
    - Standard deviations of band values
    - Map algebra on the object using bands
        - Such as mean blue divided by mean red
    - Minimum and maximum pixel values
        - For example, what is the lowest value for the red band in any pixel in the object?
- Shape
    - Surface area
    - Perimeter
    - Surface area divided by perimeter
    - Length of major and minor axes
- Texture
    - Various metrics for smoothness and coarseness
        - Coarseness is defined by an uneven gradient in the transition of digital numbers across an objects
            - Note that the digital numbers tell us about reflectance values
- Relationship to other objects
    - Relative border to other objects
        - What portion of this objects' border is with each of its neighboring objects?
    - Relative reflectance values to other objects
        - Does this object have higher mean red than its neighbors within 100m?
    - Relative values for any metric
- Direction
    - Is the object oriented north/south, east/west?

- Position
  - Where is an object in the image? For instance, is it further to the southwest?

When an OBIA expert begins to build a ruleset for classifying imagery they ought to follow these steps for the sake of efficiency:

- Make a subset of your imagery that contains all of the kinds of features that you will have to map
  - This is much more than land-cover categories
    - If "impervious" is a land cover category, then features that must be mapped include large flat buildings with dark roofs, small red buildings with driveways, tall buildings with long shadows, sidewalks, pools, dockyards and so on
    - What distinguishes kinds of features to map is based on how the OBIA classifying will understand the image
      - This includes concepts like "has a long thin adjacent shadow" - houses, "is very bright in all bands and is long and thin" - sidewalk, "is dark but surrounded by light" - pool, "is triangular and has a triangular shadow" – conifer
  - Make this subset as small as possible to avoid long computation times
- Mask first
  - Never spend any time on parts of the image that you do not want to classify
    - This includes both out of study area and out of grid area – if you have another OBIA project that classifies that spot you should not spend time doing it now
- Make use of thematic layers
  - The more ancillary data that you can incorporate immediately the better
  - You do not need to keep these data as they are – they may only be a simple mask with rules like "tree classes will not be found outside the tree mask" or "tree classes will not be found inside the water mask"
- Begin creating broad temporary classes that will either be whittled down to your final class or be added to until they are complete
  - In other words, attempt to avoid starting with confusion of both omission and commission – it is best to make your task either removing commission or finding the omitted objects
- Do not be afraid to constantly merge all objects in a class and segment again, as long as the computation time is low
  - Keep computation time down by working on the smallest possible class with intensive algorithms, and using the least intensive algorithms for large classes

Once the ruleset is complete it will serve as a description of how to understand the landscape in terms of image-object properties and their relations to each other. For instance, some of the impervious surfaces might be large, square, dark-roofed warehouses surrounded by bright parking lots. Ultimately, these features will be folded into the impervious category, and may not all be warehouses, but the important thing is that they are different from other impervious surfaces and may therefore require unique rules.

You may find that your ruleset does not solve certain kinds of problems. Maybe there were features that you neglected to include in your test dataset, or maybe there is a problem with the imagery in some locations. These can be fixed along the way by incorporating additional rules that are activated by if/then statements, or else are only run as needed.

When all of your grids are complete and you have exported the classifications, you can move on the post-processing.

holmes OBIA ruleset

The specifics of the *holmes* OBIA ruleset as used in eCognition follow:

1. Delete any samples from a previous grid
2. Mask areas not to be classified
3. Classify thematic layers
    a. In this order of precedence: impervious, water, wetlands, forest
    b. Fix persistent problems with thematic layers
        i. Remove objects in thematic layers small than the project's minimum mapping unit
        ii. Pull wetland boundaries away from impervious because often these areas are non-wetland
        iii. Pull forest boundaries away from everything else because they are too generous
4. Segment forest
    a. Make chessboard objects just in forest
    b. Merge these objects if the difference between their NDVI and their neighbor's NDVI is below a threshold
        i. Loop this until no more candidates are possible
5. Segment unclassified
    a. Perform a multi-resolution on all bands for all areas that remain unclassified
        i. Multi-resolution allows for objects of varying sizes. The greater the variance in band values over a given distance, the smaller the output image object.
6. User is warned to save the project
7. User takes samples for many classes
    a. Note that some of these classes will not appear in the final product but are instead merged into each other.
8. Perform nearest neighbor classification given these samples
    a. Variable used by nearest neighbor are sum of all band values, mean band values, standard deviation of NIR, and maximum difference between band values

For a complete list of steps used by the *holmes* OBIA ruleset described in a format prepared by eCognition, see an attached document.

Post-processing

After OBIA has been performed certain steps still need to be taken to complete the classification.

1. Ensure that these meet your project standards
    a. Map projection
    b. Georeferencing
        i. Data may mysteriously be in the wrong place on the surface of the Earth
    c. spatial scale of data
    d. Map attributes
        i. Ensure that No Data values and values for all categories are appropriate. No extraneous values, no needed values missing
    e. Mask is appropriate
        i. Ensure that you do not have small No Data errors from slivers or, or anything missing
        ii. Note that it is possible to have missing categories, missing areas, etc
    f. File name
    g. File location
2. Mosaic grids
    a. If the above standard are met then you can proceed with mosaicking
3. Manually edit
    a. Be sure that all manual editors understand certain standards
        i. For example, do not examine the image at too fine a scale
            1. Avoid spending excessive time in minutia such as the straightness of a sidewalk's outline
        ii. Do not make edits smaller than the MMU – these will be eliminated at a later stage in post-processing
    b. Be sure to gather up all the most common features that must be manually edited and change your ruleset accordingly. In many cases, it is possible to make a ruleset that will avoid the need for that particular kind of edit altogether. For instance, you can eliminate the need to manually merge shadows with houses that they are adjacent to by adding an OBIA rule to merge any shadow with its adjacent house.
4. Edge match grids
    a. This is another form of manual editing, but it is meant to address inter- rather than intra- grid classification problems
    b. May not be necessary depending on method

Accuracy assessment

Please refer to Giner (2013) for an in-depth description of how the accuracy assessment was performed. The following sections describe ways to automate the creation of the accuracy assessment reference dataset and to perform the accuracy assessment using industry-standard GIS software.

How to Create Samples

The script that accompanies this document automatically creates the polygon samples for areal accuracy assessment. Making the samples by hand is possible in almost any GIS software as it requires only the most common tools. However, this would be a very tedious task for many strata or for a large land-cover dataset.

The script is a GRASS GIS module, in the way that IDRISI has modules. Modules are tools that are called from the GIS software to execute a task. This script requires a Linux machine that can run the Bash shell, with GRASS GIS and GDAL/OGR installed. Note that GDAL/OGR, and an additional script, are only required if the input raster is huge and needs to be diced into pieces first. Fortunately all the tools required by the script are free and open source and available online. They can be installed through a system common on Linux machines called package management. The installation procedures described here use APT package management, common to Debian-based Linux distributions.

How to Install Prerequisites

$ sudo apt-get install grass gdal-bin

How the Script Works

The user inputs a raster land-cover map, names of directories for output sample polygons, the number of samples per stratum, the growing radius to use on the samples, and the minimum mapping unit used by their input map. Additionally, two optional inputs can be set. The user can specify whether the raster is so large that it must be diced into pieces, and if so what the size of these pieces should be. This will have no effect on the output samples – all it does is put less stress on the machine's memory in exchange for more computation time. The other optional input is whether to calculate the raster categories as being the same as the set of integers between the maximum and minimum values. This considerably speeds up the first steps of the analysis on large rasters, but may be inaccurate for your dataset. For instance, your map may have categories 1, 3, and 4, rather than 1, 2, 3, and 4. This option must be used with caution.

1. GRASS module options and flags
    a. GRASS modules understand user input as either options or flags. Options can include paths to files, such as an input raster. Flags are Boolean. For instance, whether or not to assume that 0 is a null value.
2. Check whether in GRASS
    a. The script will exit if the user is found to be using it outside of the GRASS environment. This is because the GRASS tools used by the script cannot be called outside the GRASS environment. To enter GRASS, use a terminal to call either $ grass64 –text or $ grass64 -gui
3. Accept the options and flags described by the user
4. Remove all raster and vector files in the current GRASS mapset

    a. GRASS organizes GIS files under a database, then locations, then mapset. Databases are directories under which all the GIS data is stored. Locations restrict the projection and the extent of the dataset. Mapsets exist under locations, and can be used to logically organize files by project or to allow for different user permissions. The user must be careful to enter a mapset that is either new or that contains no important data.
5. If dicing a large raster
    a. Make temporary directories to place the diced files into
6. Import raster provided by user
7. Determine the list of categories to use as strata
    a. Do not include any categories listed as null by the user
8. Determine whether the number of categories is one or greater than one
    a. Later on in the script if multiple categories were used then the outputs will be merged into a single file. If only one category was used this step will not take place.
9. Remove temporary files that might have been created by a previous run of the script
    a. These will be overwritten
10. For every stratum selected, perform the following
    a. Isolate the stratum as a single file
    b. If dicing, dice the isolated stratum
    c. Buffer the isolated stratum by the distance specified by the user
        i. This is so that the probability of selection of any given point inside the final samples is not biased by distance to edge of stratum
    d. Inside the buffered stratum, randomly select the number of sample points specified by the user
    e. For these points, buffer by the distance specified by the user
    f. Clip these buffered areas to the stratum
    g. Convert these to vectors
    h. If the area of any vector polygon is below the MMU specified by the user then discard it
11. If there is more than one category, merge all the categories
12. Randomly sort the features so that the enumerator cannot tell what the original strata were
    a. This is necessary to ensure that the enumerator is blinded to the categories found in the map. So far GRASS has kept all the features organized by stratum.
13. Export the merged and randomly sorted vector
    a. Export one copy as enumerator – this will not contain any land-cover information
    b. Export one copy as master – this will be identical except that it will contain land-cover information
    c. Ensure that the enumerator does not see the master copy, at least until they are joined after hand digitization

Enumerator's task

Now that the script has produced polygon samples, the enumerator takes the enumerator shapefile and overlaps it with aerial photography or whatever other imagery was used in the land-cover classification. Then they hand digitize just the area inside the sample polygons using a tool such as "cut polygon feature" in ArcGIS, or some similar tool in another GIS software. They must provide a land-cover attribute for each new polygon that they make. By the end of their analysis, all of the sample polygons will be broken into new polygons with their own land-cover attributes. It is crucial that the enumerator neither add to nor subtract from the sample polygons created by the automated method described above.

How to make the error matrix

Once the enumerator has finished hand-digitizing and providing attributes for the polygon samples, that shapefile and the old master shapefile with the land-cover information from the map can be used as inputs to another *holmes* GRASS module, vcontingency.sh. This script produces a table with three columns: the enumerator land cover, the map land cover, and the surface area in map units that is shared between these two categories. This works on all of the polygons simultaneously. There is no need to consider sample polygons one at a time.

Once a pivot table has been made in Microsoft Excel from the output of that script, it can be summarized by sum of the surface area column. The table is now a traditional error matrix used for map accuracy assessment.

Data backup

Do not leave data backup entirely in the hands of a third party. They may do a good job but backup is about redundancy. You should make sure to always have your own copy. The ideal backup system would have these properties:

1. Backs up on a schedule
2. Retain old data inside folders with that day's date
    a. Never lose your data – it may become relevant later
3. Exists in multiple physical locations
    a. Protect against disasters like fire and flooding
    b. It is sufficient to back up your data with a third party to meet this criteria – but if you do consider data sensitivity. You may want to encrypt your data
4. Exists in multiple media
    a. Such as hard drive and compact disc
5. Does not backup trash
    a. Much of your data is probably useless – see the folder hierarchy section for the concept of scratch folders
    b. This minimizes the space needed for the data – the less space taken by the backup the more backups you can have
    c. Consider simply buying some cheap external drives to keep additional copies of your data
6. Encrypted if it is sensitive, but remains retrievable
    a. Make sure to not get stuck with an encrypted backup that you cannot decrypt – this is just as useless as no backup at all
7. Backs up incrementally
    a. This is for speed. Incremental backups only transfer those bytes of data which are different from the old copy to the new copy. This is sometimes called delta copying.

SQL queries

ArcGIS essentially uses SQL to make queries, but they are extremely limited compared to what can be done in a SQL database. Also note that OGR is capable of using a SQL driver to interface with a shapefile's attribute table – DBF. This makes OGR a powerful tool for using SQL queries on plain

shapefiles. Even better than either ArcGIS or OGR of course, is a real SQL database with a spatial backend, like PostgreSQL with PostGIS. This will allow you to enter shapefiles as tables, to make joins, summarize fields, order results, choose one field by another, and so on. Any SQL query can be made into a new table which can be exported as either a CSV file or a shapefile.
Kinds of queries best performed in

- ArcGIS or another desktop GIS
  - one time, one file
  - output is selection of features
    - problematic because it is prone to crashing and does not last unless exported as a new file
    - slow
- OGR
  - simple query, many files, many times
  - output is new file
    - can be performed on any number of files with equal effort
    - fast
- PostgreSQL with PostGIS
  - Complex query, many files, many times
  - Output is new table in the database, new CSV, or new shapefile
    - Can be performed on any number of files with equal effort
    - fast

Joining spatial and tabular data is best done using a field that either has

- spatial information, like latitude and longitude of points
- some unique identifying field, like parcel ID or block group ID

With a field with spatial information you will really just make a new spatial file and perform a spatial join. With a field with a unique identifier you will have to get the two unique identifying fields to agree and then join the tables. Note that you should check the field to make sure the identifiers are actually unique, for instance, by reporting the number of unique values in the field and comparing that to the number of rows in the table. Note that it is likely that the fields do not quite agree already – for instance, you may have to remove leading zeroes, or perform some other similar task. When this becomes extremely difficult, such as adding and removing various numbers and letters under different conditions, it is very useful to know regular expressions. Regular expressions are a method to explain text string pattern matching to computers using programming.

Delivering our products

End users might make requests like
1. Selecting certain categories
2. Selecting a polygon area
3. Delivering data with different specifications
   a. File format
   b. Spatial scale
   c. Null value
   d. Projection

End users will want access to files far too large to send over email. When this happens File Transfer Protocol (FTP) is a good solution. Filezilla is a simple, free, and open source FTP program commonly used in universities. Tell your end users to get the client version of the program if they do not have it already. The only information that the end user needs is hostname, username, and password. If you do not use the default FTP port (21) they will also need to know the port number.

Note that many issues can prevent your end users from connecting to your FTP server. To troubleshoot network connectivity, be sure to conduct tests in a logical order and determine where the problem lies. For instance, can you connect to the FTP on your own machine? If you can then the server is running. Can you connect to the FTP from another machine inside the same network? Then your port must be open. Can you connect to the FTP from outside the network? Then your port must be exposed to the wider internet – be careful about this. You might prefer to stay in network and request your users to use a VPN to connect to the network.

Common issues include:

1. Your machine's firewall
    a. Open port 21 or relevant port
    b. Configure to use as FTP server – this is a special section of the control panel for Windows
2. Your network
    a. Talk to your IT department

If your data are sensitive and must be secure consider one of the following transfer methods:

- Transfer an encrypted file
    - Ensure that your end user receives the material necessary to decrypt in a secure fashion
    - Consider using private key security rather than symmetrical ciphers
        - Symmetrical ciphers are our most common passwords – the same message is used to encrypt and decrypt
- Use an encrypted tunnel to transfer files
    - Secure shell (ssh) is the most common example
- Or try SFTP, which is not exactly SSH

Glossary

- ArcGIS – common desktop GIS software made by ESRI with a focus on vector datasets
- CSV – comma separated values file. Used like a table.
- OBIA – object based image analysis.
- Shapefile – very common vector file format made by ESRI
- SQL – structured query language. Can be used to make complex queries very easy by asking questions about tables in a database and how those table relate to each other

Cartographic tips:

Establish an easy to read color scheme for categories that is easy to look at, and make it default in your project. Ensure that palette files, such as for ArcGIS and IDRISI, are used each time for the appropriate layers. Note that different color schemes may be appropriate for manually edited products, final land-cover products, files to display on a desktop GIS, and PDF maps meant to be seen by end users.

Make sure that PDF maps have high dot per inch counts (DPI). Around 300dpi is commonly considered adequate.

Do careful about unnecessary clutter. Maps do not necessarily need north arrows – your users understand that north is probably up.

Be aware that some end users, especially males, may be color blind. Entire palettes are designed with the color blind in mind. Consider consulting the work of Cynthia Brewer of Penn State University: http://colorbrewer2.org/

Batching large tasks – to free up machines and person hours, and even the accomplish tasks that are impossible in commercial software like ArcGIS or IDRISI.

All of the following batched tasks require a Linux machine running the Bash shell, and with some combination of GDAL/OGR, GRASS GIS, PostgreSQL, PostGIS, and Quantum GIS installed.
These are all free and open source tools easily installed from package managers like APT.

Here is an example of how to automatically identify all rasters that do not meet a given projection, and to reproject them to whatever is appropriate.

```
#!/bin/bash

searchstring="PROJCS\[NAD83 / Massachusetts Mainland"
EPSG=26986
list_to_proj=$(for i in *.img; do proj=$(gdalinfo $i | grep "PROJCS\[\"" | sed 's:"::g;s:,::g'); echo
"\"$i\"","\"$proj\""; done | grep -v "$searchstring" | sed 's:,$::g' | awk -F, '{ print $1 }' | sed
's:"::g')
for i in $list_to_proj
do
	gdal_translate -a_srs EPSG:$EPSG $i $(basename $i .img)_proj.img
done

exit 0
```

Here is an example of how to make a raster with a built in color table from a Quantum GIS symbology file, known as .qml

```
# first make a virtual file – an XML file describing the rasters' properties
gdalbuildvrt out.vrt in.tif

# now make an intermediary colortable output from the .qml file.  Note that this intermediary
output is sufficient for gdaldem
grep colorRamp raster_foo.qml | sed 1d | awk '{ print $4,$2,$5,$3 }' | sed 's/^value=//g' | sed
's/red=//g' | sed 's/blue=//g' | sed 's/green=//g' | sed 's/"//g' > /tmp/color.txt

# now make a true colortable to be inserted into the virtual file
echo "<ColorInterp>Palette</ColorInterp>"; echo "<ColorTable>"; echo "<Entry c1="0" c2="0"
c3="0" c4="255"/>"; awk '{print "<Entry c1=\42"$2"\42 c2=\42"$3"\42 c3=\42"$4 "\42
c4=\42""255\42""/>" }'  /tmp/color.txt; echo "</ColorTable>" > color-relief.txt
```

```
# insert these values just above </VRTRasterBand> in the .vrt file

gdal_tranlate out.tif out.vrt
```

Here is an example of how to automatically export TIFF files as ASCII using a new no data value, using Bash and GDAL/OGR.

```
for i in *_ratio?.tif
do
	gdal_translate -a_nodata -9999 -of AAIGrid $i $(basename $i .tif).txt
	sed -i 's:-nan:-9999:g' $(basename $i .tif).txt
done
```

We have also produced scripts that handle the following tasks:

1. Determine the proportions of raster categories found in each feature of a vector dataset – eg, proportions of land-cover in lawns
    a. Can also be used for downsampling large rasters
2. Producing polygon samples for human interpretation in an areal accuracy assessment method
    a. Making a contingency table from the enumerator output and the master copy of the sample polygons
3. Dicing large rasters into smaller rectangles
4. rasterize arbitrary number of vectors
5. add column to vectors to be used in rasterization based on an attribute
    a. useful for rasterize
6. PostgreSQL
    a. Make PostgreSQL databases, users, passwords
    b. Send shapefiles to a PostgreSQL database
    c. Select by location
        i. Select all features from one shapefile that overlap with another shapefile
        ii. Select all features from a file that overlap more than a given percentage – very useful for polygons that do not quite match, like parcels or census block groups
    d. Answer questions about the proportions of a field that is held by each row, join these across tables
7. Clipping a large raster by a directory of polygons – one raster per polygon
8. Clipping a vector by a vector – useful for batching when you ask it to work on thousands of pairs
9. Basic Linux admin tasks
    a. Make users, passwords, home directories
        i. Remove same
    b. Apply appropriate permissions to all user files