Project Report: Land Cover-Adjusted Soil Organic Carbon Mapping for Massachusetts

1. Executive Summary

Overview

Massachusetts now has its first comprehensive, land cover-adjusted soil organic carbon (SOC) maps, providing decision-makers with the data needed to evaluate carbon impacts of land use choices. This project, funded by the Massachusetts Healthy Soils Initiative, delivers high-resolution (1-meter) maps showing current soil carbon stocks across the Commonwealth, and tools to predict how land use changes will affect these stocks. The resulting dataset reveals that Massachusetts soils store approximately 425 million metric tons of carbon, with the amount and vulnerability of this carbon varying dramatically by land use and context.

Project Goals

The Massachusetts Healthy Soils Action Plan identified soil organic carbon as a critical metric for soil health and climate mitigation, but data available at the time could not answer fundamental questions about how land use drives the distribution of soil carbon across the state. This project aimed to fill that gap by creating tools to enable state agencies, municipalities, planners, and land users to incorporate soil carbon into decision making. These tools include:

- Maps showing current soil carbon under existing land uses
- Maps showing maximum potential soil carbon under optimal management, i.e. as forest
- A conversion factor table quantifying carbon changes from land use transitions

These tools help to transform soil carbon from a concept into actionable data that supports evidence-based land use planning and climate policy.

Background

Soils is the Earth's second largest carbon reservoir (after the ocean), storing more carbon than the atmosphere and plant life combined. Soil carbon is dynamic and strongly affected by land cover and land use—the physical features on Earth's surface such as forests, farms, wetlands, and developed areas. When forests are cleared or wetlands drained, stored carbon is released to the atmosphere as carbon dioxide, contributing to climate change. Conversely, restoring forests or improving agricultural practices can pull carbon from the atmosphere and store it in soil.

Data and maps that show how soil carbon is distributed across land uses and across the state are lacking. The USDA's Soil Survey Geographic Database (SSURGO)—our best source of soil information—typically provides carbon estimates for only one land use per soil mapping unit. However, when overlaid with contemporary land cover data, each soil mapping unit in Massachusetts contains an average of twelve different land uses. The Commonwealth has not had tools to accurately assess current soil carbon stocks or predict changes from development, conservation, or restoration activities. This could lead to ignoring or miscalculating the climate effects of those changes.

This project bridges that gap by systematically estimating soil carbon for every combination of soil type and land use across Massachusetts, creating the first dataset capable of supporting a broad range of carbon-informed land use decisions.

Outcomes

Key Deliverables

The project delivered all planned products, plus methodological enhancements developed through collaboration with NRCS soil scientists:

1. **Current and Potential SOC Maps**: High-resolution maps at three depths (0-30cm surface layer, 0-100cm standard assessment depth, and full soil profile) showing carbon density for every square meter of Massachusetts under both current and optimal land use scenarios.
2. **Land Use Conversion Factors**: A comprehensive table quantifying carbon impacts of land transitions. For example, converting forest to development results in a 62% loss of soil carbon in the top meter, while converting pasture to forest increases soil carbon by 79%. These factors include confidence intervals, providing planners with bounded estimates for decision-making.
3. **Spatial Data Products**: Digital files compatible with state GIS systems, enabling integration into existing planning workflows and carbon accounting frameworks.

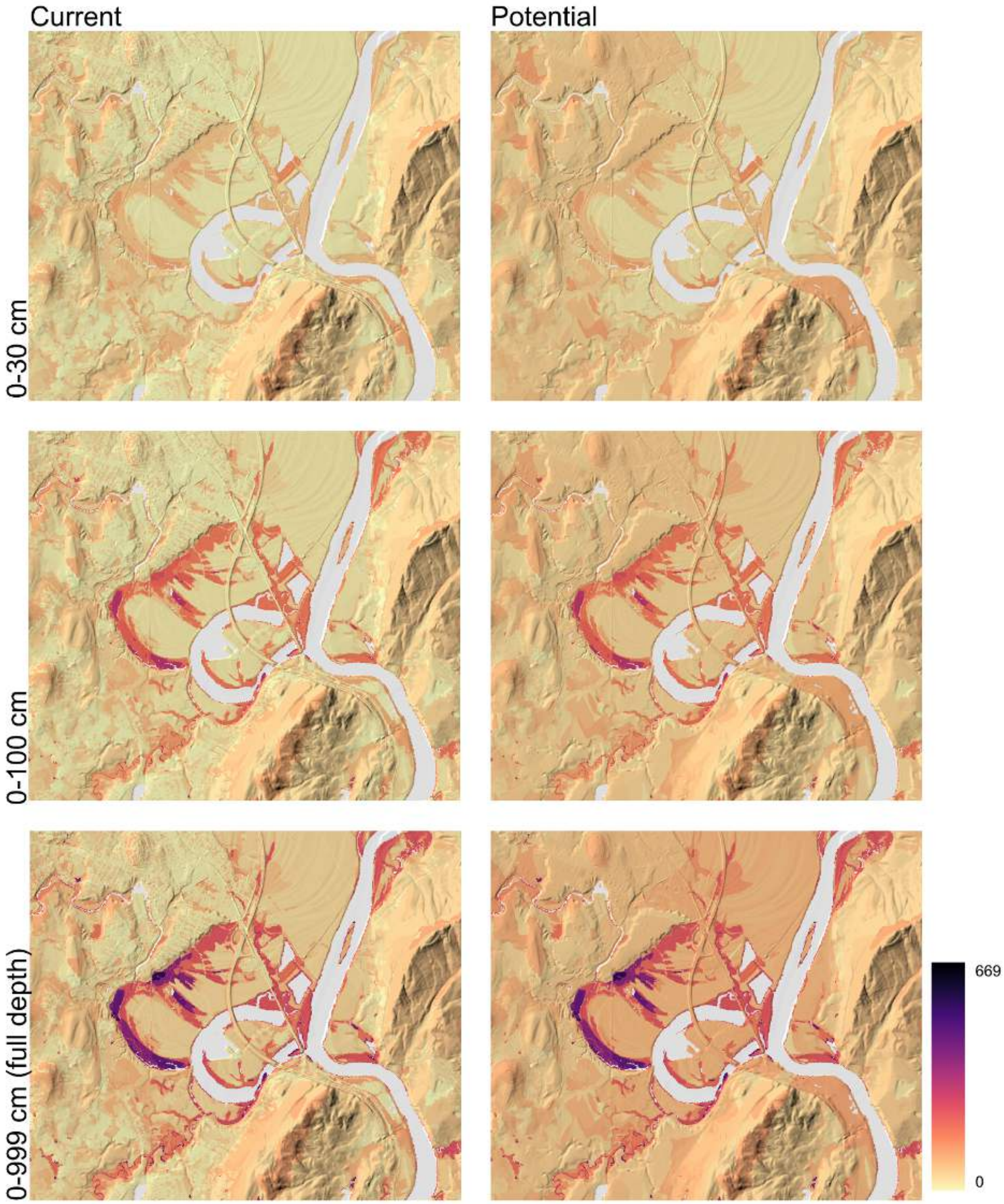

Figure 1. Comparison of the Current and Potential outputs at 3 depths in the Northampton oxbow region. Values are in metric tons per acre. This is shown over a hillshaded elevation relief.

Major Findings

The analysis reveals several critical insights for Massachusetts policy makers:

- **Total Carbon Storage**: Massachusetts soils store 252 million metric tons of carbon in the surface layer (0-30cm) and 427 million metric tons through the full soil profile—a massive carbon reservoir that is vulnerable to land use decisions.
- **Disproportionate Wetland Importance**: Though covering only 12% of the state's area, wetlands store 35% of soil carbon, with some wetland soils containing over 300 metric tons per acre compared to 74-78 for typical forests.
- **Development Impact**: Conversion to development represents one of the largest carbon losses, with even "green" developed spaces storing only 38% of the soil carbon of the forests they often replace.

Limitations

Users should understand several important caveats when applying these data:

Current Limitations:

- The analysis combines soil data collected over multiple decades with 2016 land cover data, introducing temporal uncertainty
- Some land cover categories, like "Cultivated Land," include diverse uses with highly variable soil carbon storage (annual crops, orchards, nurseries)
- Validation of our estimates relied primarily on comparison with existing data ranges rather than extensive field verification
- Conversion factors are not applicable in every context, e.g. conversion to cultivation on very shallow or stony soils. Wetlands in particular are not suitable for land cover conversion factors, and have therefore not been included in the conversion factor table
- Conversion factors between Cultivated Land and Developed Open Space raise concerns, as they do not match expectations based on established scientific understanding, which suggests that annual vegetation (Cultivated Land) should have lower SOC than perennial (Developed Open space – typically sod). Our analysis shows the reverse, with Cultivated Land having 29% higher SOC at 0-100cm depth. We believe this is due to (1) The broad range of land covers included in Cultivated Land, as noted above; and (2) Bias in land user preferences that are reflected in the SSURGO dataset, i.e. the best soils with the highest SOC are selected for farming. Contrary to the conversion factor estimated in this project, these high carbon agricultural soils would likely increase further if converted to perennial vegetation such as sod.

Future Improvements:

- Update with newer land cover data as it becomes available
- Validate and calibrate as new soil carbon data becomes available
- Integrate with carbon monitoring programs to track changes over time
- Enhance ability to distinguish agricultural subcategories
- Develop user-friendly interfaces for non-technical users
- Develop a method for crosswalking landcover categories from the 2016 high resolution categories (used in this analysis) to the 30m National Land Cover Data (released annually) in order to track annual change

This dataset represents the best available information on land cover-specific soil carbon in Massachusetts and provides a robust foundation for carbon-informed decision-making. The transparent methodology allows for continuous improvement as new data and techniques become available.

2. Technical Report

Overview

This project develops a spatially explicit, high-resolution (1m) soil organic carbon (SOC) density map for Massachusetts by integrating USDA-NRCS SSURGO soil survey data with 2016 land cover information. The methodology addresses critical gaps in SSURGO's land cover-specific SOC values through a multi-path imputation framework, producing SOC estimates for all land cover types for each soil mapping unit. The resulting dataset provides SOC density values (metric tons per acre) at three depth intervals (0-30cm, 0-100cm, and 0-999cm) for each unique combination of soil mapping unit and land cover type.

Project Goals

The central objective was to extend the utility of existing SSURGO data by creating a current, land cover-adjusted SOC map that:

- Provides spatially explicit SOC estimates at 1-meter resolution
- Accounts for the heterogeneity of land cover within soil mapping units
- Fills gaps where SSURGO lacks land cover-specific SOC data
- Leverages the improved data quality from MLRA-harmonized soil components
- Produces the best available SOC estimates to support carbon accounting, conservation planning, and climate change mitigation efforts

Background and Rationale

SSURGO represents the most detailed soil survey information available for the United States; however, the database presents several challenges for SOC mapping:

1. **Limited land cover representation**: Most SSURGO mapping units in Massachusetts contain SOC values for only a single land cover type. When mapping units are overlaid with land cover data, however, mapping units contain an average of over 12 distinct land covers each.

2. **Data quality variability**: SSURGO data originates from across different time periods, field offices, estimation methods, and individual soil scientists, creating inconsistencies across the dataset.

3. **Missing data**: Many soil components lack land cover classifications and SOC values entirely.

4. **Spatial vagueness**: Component data provides only percentage coverage estimates across mapping units, without spatially explicit locations.

Despite these limitations, SSURGO remains uniquely valuable as a large-scale soils dataset. This project leverages SSURGO's strengths while addressing its gaps through systematic imputation and integration with high-resolution land cover data.

Methodology

Data Sources and Preparation

Primary Data Sources:

- SSURGO database for Massachusetts (gSSURGO_MA, December 2023)
- 2016 Massachusetts Land Cover/Land Use dataset (1m resolution)
- MLRA Harmonization Project (2010-2015) designation, identifying upgraded components
- 2016 USDA CropScape data layer for distinguishing cranberry cultivation

Initial Data Processing:

1. Extracted soil component data focusing on major components
2. Compiled soil formation variables (parent material, deposit type, texture, drainage class) for hierarchical imputation
3. Developed a crosswalk table linking SSURGO land cover categories to 2016 land cover classes

SOC Estimation Framework

The methodology employs a multi-tiered approach to estimate SOC values for every land cover type within each mapping unit.

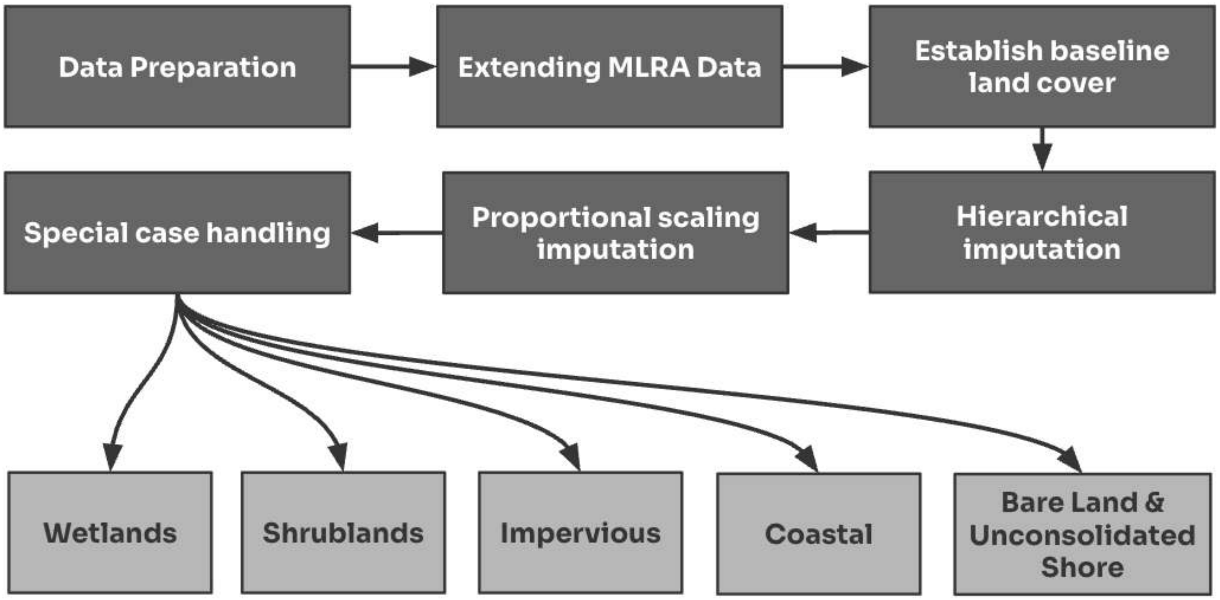

Figure 2. Soil Organic Carbon Estimation Methodology

1. MLRA Data Extension

SSURGO soil components updated during the 2010-2013 MLRA data harmonization project (MDHP) were identified through communications with RDG by NRCS officers as having superior data quality. SOC values from these MLRA components were extended to nonMLRA components sharing the same component name and land cover type, effectively propagating high-quality estimates across the dataset (Fig. 2).

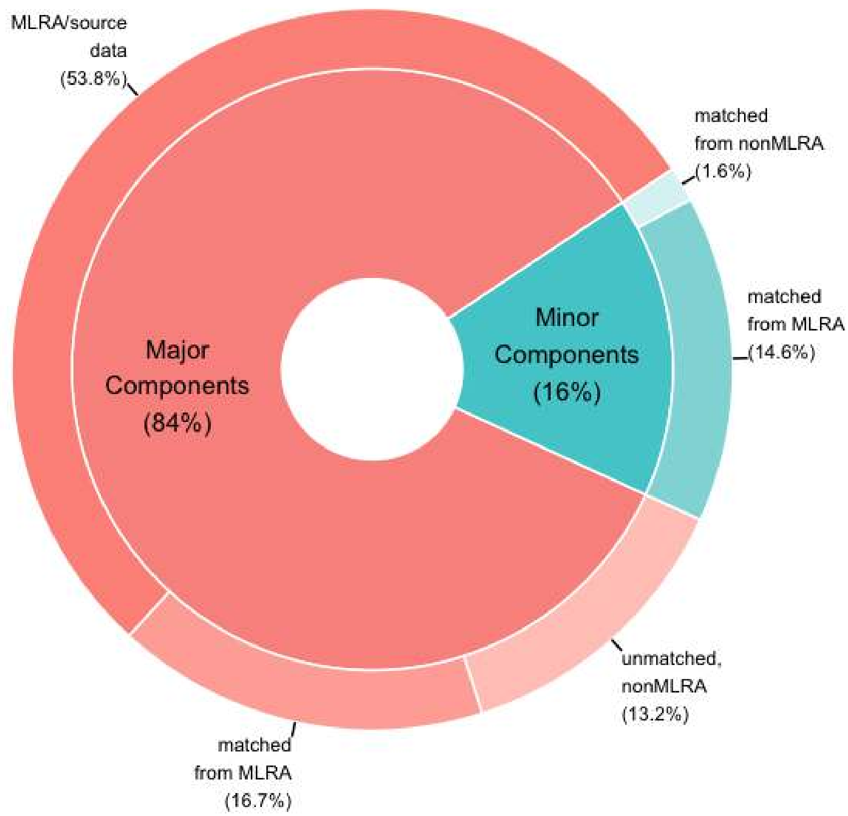

Figure 3. Extension of MLRA harmonized component data across the dataset. Figure shows the distribution of major and minor component data from within and without the data MLRA harmonization project, by percent of total land area in the dataset. Note that all minor component values are based on matching with major components by component name, as minor component data were not directly used in any SOC estimation.

2. Baseline Land Cover Identification

For each soil component, a baseline land cover was established using the following hierarchy of preference:

- Forest (if available)
- Pooled herbaceous vegetation (if forest unavailable)
- Cultivated lands (if neither forest nor herbaceous available)
- Imputed forest values (if no suitable baseline present)

3. Hierarchical Imputation for Missing Forest Values

When forest SOC values were needed but unavailable, a 36-level hierarchical search was implemented. This search identified similar soils based on increasingly generalized combinations of:

- Deposit type (detailed to broad categories)
- Parent material (specific minerals to broad rock types)
- Texture (specific to 3-bin classifications)
- Drainage class (6 classes to 3-bin to 2-bin)
- Temperature regime
- Geographic context: MLRA (Major Land Resource Area) to LRR (Land Resource Region) membership

The hierarchical search used area-weighted geometric means to aggregate SOC values at each level. Geometric means are well suited to the lognormal distribution of SOC data and are more resistant to outliers (compared to the standard, arithmetic mean), thereby providing a more reliable representative estimate of SOC.

4. Proportional Scaling

State-level proportions were calculated between different land cover types and applied to baseline values to estimate missing land cover-specific SOC values. State-level proportions were used rather than local proportions because:

- Local proportions appear to be, and logically are, heavily influenced by land use decisions (e.g., farmers selecting deeper, better soils for cultivation)
- State-level calculations averaged out these confounding factors
- Resulting proportions aligned better with established literature values

5. Special Case Handling

Wetlands: Wetland formation is complex and wetland SOC is unlikely to be as affected by land cover to the extent of upland. For this reason, proportional imputation is not appropriate for wetlands. Direct hierarchical imputation was applied to wetland land covers (forested wetlands, shrub wetlands, herbaceous wetlands) using the same 36-level hierarchy of soil characteristics and region.

Shrublands: There is almost no data for upland shrub cover in the SSURGO database for Massachusetts, as it is a rare and understudied land cover type in the northeastern United

States. Therefore, shrubland SOC was estimated as the mean of forest and pasture/hay values for each location, reflecting the intermediate nature of shrub cover.

Impervious surfaces: Derived from SOC values for Developed Open Space with the top 30cm SOC subtracted (thereby setting surface SOC to zero).

Coastal mapunits: The "Beaches" component common to coastal mapunits lacked sufficient data for reliable imputation. These components were replaced with values from "Hooksan" components (dunes on lowlands), which provided analogous sandy soil conditions with available SOC data. (These results are assigned to the *Derived* column in Table 1, below.)

Bare land and unconsolidated shore: Assigned minimal SOC values based on state medians.

Table 1. Land Cover and SOC Data by Source and Imputation Method

		% of Landcover Data By Area				
Land Cover	**Acres**	**Original**	**Proportional with Original Baseline**	**Hierarchical**	**Proportional with Imputed Baseline**	**Derived**
Evergreen Forest	1,198,133	77	3.9	19.1	0	0
Deciduous Forest	1,820,707	74.5	3.8	21.7	0	0
Scrub/Shrub	66,808	0	0	0	0	100
Cultivated Land	61,784	24.4	51.5	0	24.2	0
Cultivated Land (Cranberry)	11,831	83.5	0	16.5	0	0
Pasture/Hay	132,250	31.4	41.3	0	27.3	0
Grassland	132,300	49.7	22.6	0	26.7	0.9
Developed Open Space	436,578	58.9	18	0	23	0.1
Impervious	475,711	0	0	0	0	100
Unconsolidated Shore	31,291	0	3.6	0	86.8	9.6
Bare Land	48,037	4.2	56.1	0	29.3	10.3
Estuarine Forested	69	46.2	0	53.4	0	0.4
Palustrine Forested	407,599	60.8	0	39.2	0	0
Estuarine Scrub/Shrub	403	65	0	34.1	0	0.9
Palustrine Scrub/Shrub	50,366	47.5	0	52.5	0	0.1
Estuarine Emergent	45,579	79.8	0	19.3	0	1

Palustrine Emergent	89,439	7.4	0	92.5	0	0

Integration with Land Cover Data

The next step was intersecting SSURGO map unit geometry with the 2016 land cover raster data:

1. Each mapping unit was overlaid with land cover data to determine actual land cover composition;
2. SOC density values were assigned based on the map unit × land cover combination;
3. Results were aggregated to produce 1m resolution rasters for each depth interval.

Key Innovations

1. **Systematic use of highest quality data**: Prioritizing MLRA-harmonized data and major components improved reliability.
2. **Comprehensive imputation framework**: The 36-level hierarchy ensures that even data-poor components receive the best, most plausible available SOC estimates.
3. **Component name matching**: Extending estimates for major components to values, to minor components with identical names, maximized data utility.
4. **State-level proportion calculation**: Avoided local biases from land use patterns while maintaining ecological relevance.

Land Cover Conversion Factor Development

A key deliverable of this project is a Land Cover Conversion Factor Table that quantifies the SOC changes associated with land cover transitions. This table enables users to estimate carbon flux implications of land use changes across Massachusetts.

Conversion Factor Methodology

For most land cover pairs, conversion factors were calculated through a systematic bootstrap approach:

1. **Mapunit-level proportions**: The proportional difference in estimated SOC between land cover types was calculated for each mapping unit.
2. **Area weighting**: Weights were assigned as the square root of the product of the two land cover areas ($\sqrt{(Area_1 \times Area_2)}$), ensuring that conversions between extensive land covers received appropriate emphasis.
3. **Bootstrap estimation**: Using weighted geometric means with bootstrap resampling, we calculated the central tendency, and 95% confidence intervals, of the conversion factor for each land cover pair.

Special Handling for Forest Types

Deciduous and Evergreen Forest required a modified approach:

- SSURGO predominantly provides data for "Mixed Tree Cover" without distinguishing between forest types;
- Therefore, at the mapping unit level Deciduous and Evergreen forests receive identical SOC values;
- Due to their different spatial distributions the forest types, and associated soil conditions, modest differences emerge at the state level;
- Conversion factors between these forest types were therefore calculated at the state level using the same bootstrap methodology.

This conversion factor table provides critical information for:

- Evaluating carbon impacts of land use planning decisions;
- Assessing carbon sequestration potential of reforestation or afforestation projects;
- Estimating carbon losses from development or agricultural conversion.

Outputs

The project produces four primary deliverables:

1. **Land Cover Conversion Factor Table**: Bootstrap-derived conversion factors with 95% confidence intervals for all land cover pairs, enabling carbon flux calculations
2. **GeoPackage layers** containing

- **1m resolution rasters** for three depth intervals
    1. 0-30cm depth (surface layer, commonly used for carbon accounting);
    2. 0-100cm depth (standard depth for SOC assessments)
    3. 0-999cm depth (full soil profile available in SSURGO)
- **Mapping unit × land cover SOC estimates**: SOC density (metric tons/acre) for every combination at three depths
    1. **Source attribution**: Documentation of data sources (original, proportional, hierarchical, derived)
    2. **Component tracking**: Records which soil components contributed to each estimate

Results

Total Soil Organic Carbon Stocks

Our process estimates total statewide landcover-adjusted soil organic carbon at approximately 250 million metric tons at 0-30cm, 363 million metric tons at 0-100cm, and 425 million metric tons at 0-999cm.

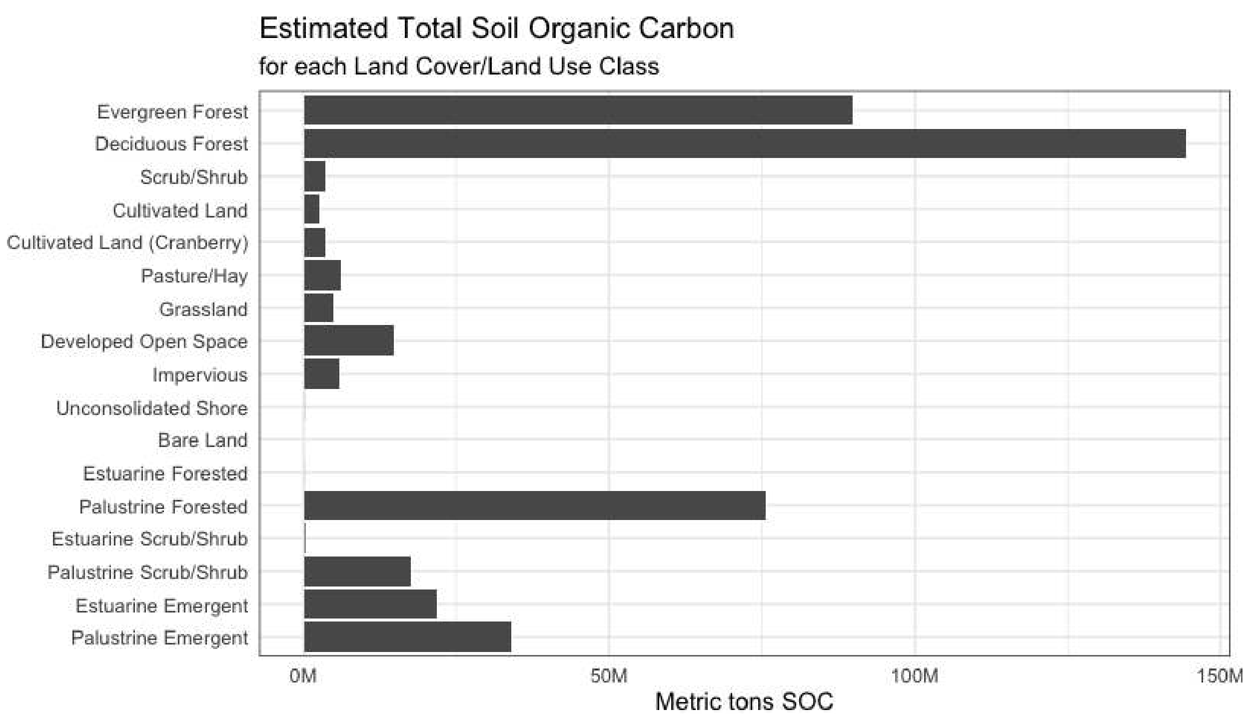

Figure 4. Total Soil Organic Carbon by Land Cover/ Land Use Class

Soil Carbon Density and Land Cover

The carbon density of each land cover class is shown in Figure 5, below. Density is shown for each of three depths, 0-30cm, 0-100cm, and 0-999cm (i.e. the whole soil profile). The colored bars represent the simple “top down” arithmetic mean for that landclass: the total estimated carbon at the state level, divided by the total acres. The black points and error bars represent a “bottom up” approach of a bootstrap geometric mean, and 95% confidence intervals, respectively. We believe the bootstrap geometric mean is a better representation of typical values for the land cover class, as it naturally accounts for variation and resists the influence of a small minority of high-density outliers.

It’s important to note the importance of carbon stored deep in the soil column. Soil Organic Carbon measurement often takes into account only the top 30cm, due to the resources needed to test for carbon at greater depths. Figure 5 shows how accounting for greater depth dramatically shifts our understanding of carbon density across land covers. At a state level, this amounts to 44% of SOC stored deeper than 30cm depth, and 17% stored deeper than 1m.

The wide error bars for some of the wetland land covers–e.g. Cranberry and Estuarine Forested–reflects the relative rarity of those land covers in the state (as reflected in the 2016 Land Cover dataset).

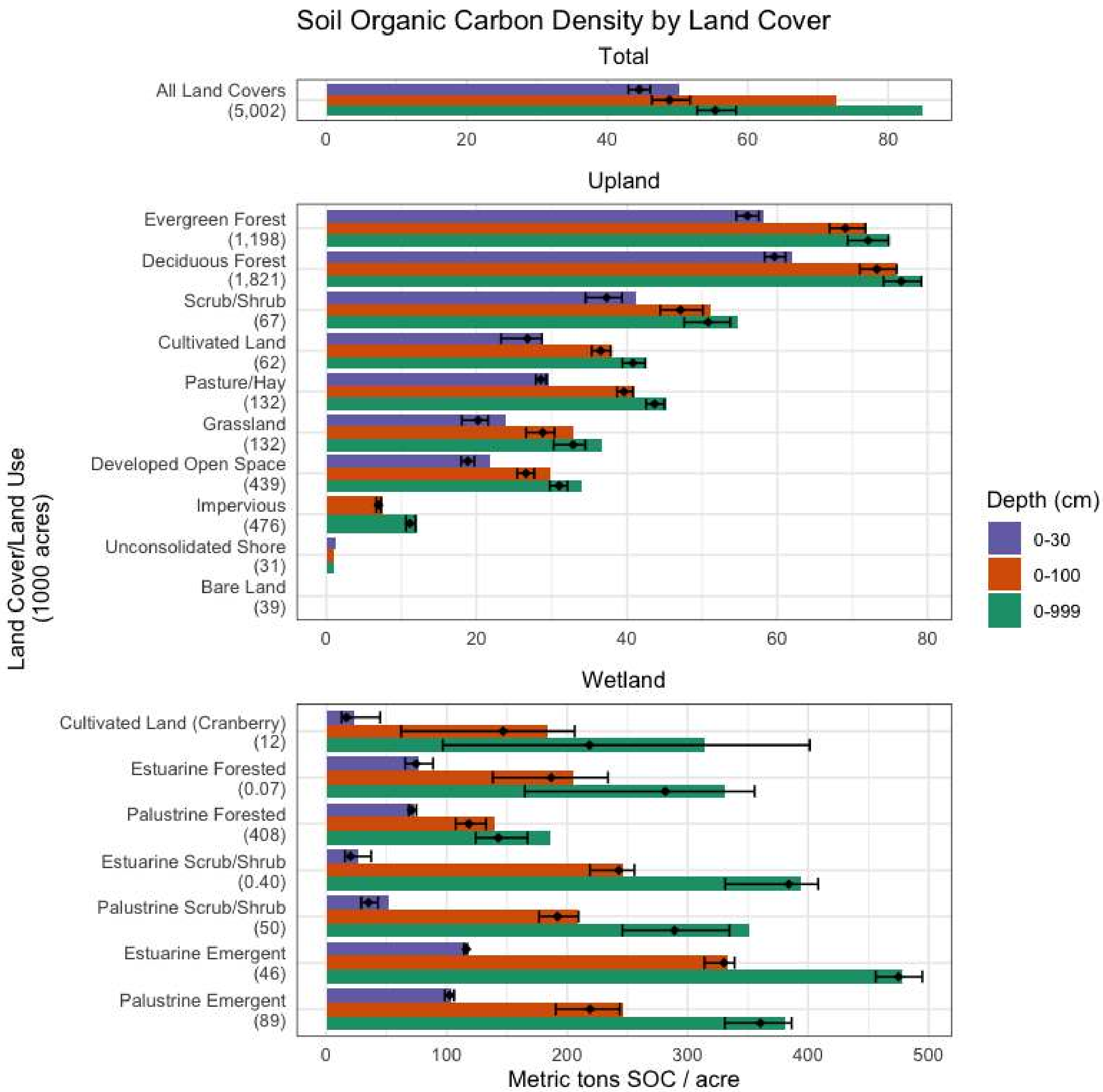

Figure 5. Soil Organic Carbon Density (metric tons/acre) for each Land Cover / Land Use Class

Land Cover Conversion Factors

The conversion factors (or multipliers) for upland land cover classes are shown below in Figure 6. For ease of display, they are shown here as percent change (-100% to 165%) rather than as factors (0 - 2.6).

Each panel shows the conversion factors for a different starting state, i.e. the ‘from’ land cover. Within each panel, the rows delineate the different end states, or ‘to’ land covers. For example, conversion from Cultivated to Forest results in an 82% gain. Based on carbon density described above, this amounts to approximately 30 mt/acre of SOC at 100cm depth. Similarly, conversion from Forest to Developed Open Space results in a 62% loss, or approximately -44 mt/acre of SOC.

Wetland land covers are not included in the conversion values for the same reason that they were treated as a special case for imputation (above):the dynamics of wetland formation are complex and SOC is unlikely to be driven by land cover to the extent of uplands. Conversion from Impervious and Bare Land are not shown here, as the negligible amounts of carbon present in these land covers changes the scale and makes Figure 6 harder to interpret, e.g. conversion from Impervious to Evergreen Forest yields a change of 889%.

Note that conversion factors between Cultivated Land and Developed Open Space raise concerns, as they do not match expectations based on established scientific understanding, which suggests that annual vegetation (Cultivated Land) should have lower SOC than perennial (Developed Open space – typically sod). Our analysis shows the reverse, with Cultivated Land having 29% higher SOC at 0-100cm depth. We believe this is due to (1) The broad range of land covers included in Cultivated Land, as noted above; and (2) Bias in land user preferences that are reflected in the SSURGO dataset, i.e. the best soils with the highest SOC are selected for farming. Contrary to the conversion factor estimated in this project, these high carbon agricultural soils would likely increase further if converted to sod (or other perennial vegetation).

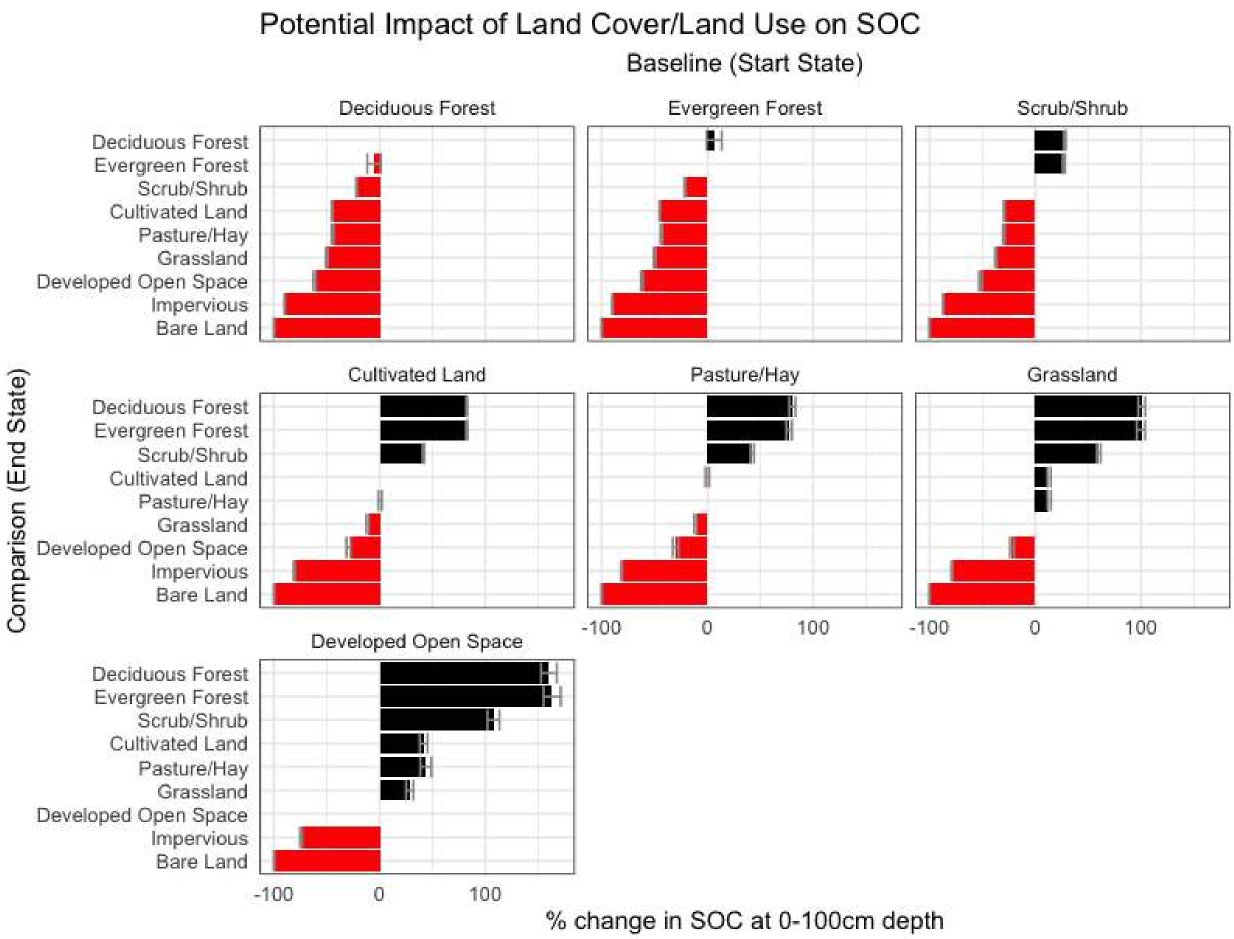

Figure 6. Land Cover Conversion Factors: % change in SOC at 0-100cm depth

Validation

Validation focused on ensuring imputed values remained within plausible ranges:

- The vast majority of imputed SOC values fell within the range of original SSURGO values
- In the small number of cases where proportional scaling produced values outside the original range, these were reviewed and deemed plausible based on known soil-vegetation relationships
- The hierarchical imputation successfully provided values for all components, with less than 1% of land area requiring the most generalized (Level 36) estimates (see Tables A3 and A4).

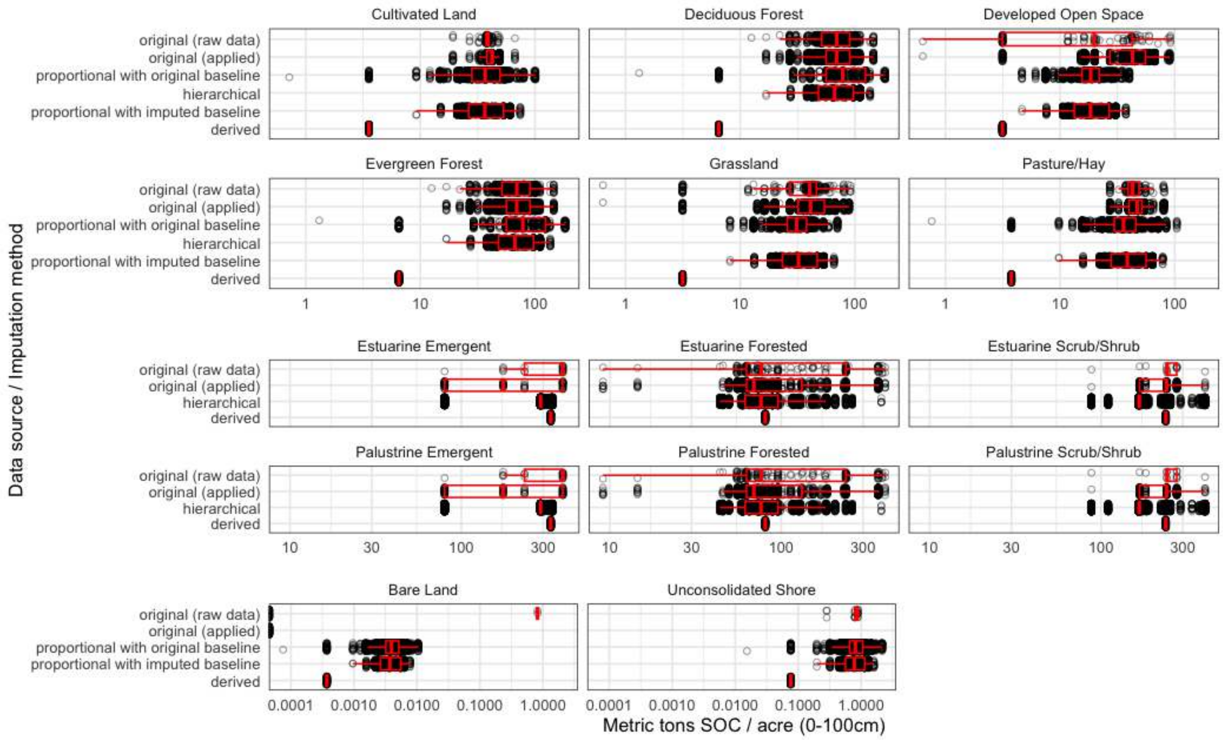

Figure 7. Impact of Imputation on SOC Values: Comparing Original, Applied, and Imputed Distribution Across Land Use / Land Cover Classes

Limitations

1. **Imputation uncertainty**: While imputed values are constrained by observed data, they remain estimates subject to uncertainty, particularly for components requiring high-level hierarchical generalization.
2. **Limited validation data**: Without extensive field measurements, validation relies primarily on range-checking against existing SSURGO values.
3. **Nonharmonized data:** Approximately 15% of the data in the dataset was not part of the MLRA data harmonization project described above, and is therefore subject to inconsistencies originating from different time periods, field offices, estimation methods, and individual soil scientists.
4. **Component spatial ambiguity**: SSURGO provides only percentage estimates for component distribution within mapping units, not explicit locations.
5. **Ambiguity of land cover classes:** The definitions of land cover classes include areas which are expected to have very different SOC values.
    a. **Cultivated Land** encompasses annual agriculture, woody perennial crops such as orchards and vineyards, nurseries, as well as cranberry bogs and aquaculture.

While CropScape data was used to separate out cranberry bogs, woody perennial crops are not represented in the MA SSURGO database, so have to remain lumped in with tilled annual agriculture.

   b. **Unconsolidated Shore** includes both very low carbon intertidal zones of sand and gravel, and high carbon tidal flats of silt and clay.
   c. **Bare land** includes both permanently and temporarily unvegetated landscapes, potentially encompassing everything from areas of bare rock to recently tilled agricultural fields.

6. **Scale dependencies in proportional imputation**: State-level proportions may not capture local variations in SOC-land cover relationships.
7. **Temporal misalignment**: The methodology combines SSURGO data from various time periods with 2016 land cover data. Land cover changes since data collection may affect accuracy.
8. **Potential implausible conversions:** Potential for use of conversion factors to model incompatible land covers for a given context, e.g. cultivation on very shallow or stony soils.
9. **Uncertainty due to interpretive focus**: In the SSURGO dataset, the NRCS estimates of organic matter for each soil component are influenced by the selection of a single land use as the interpretive focus for each mapunit.
10. **Conflation of soil potential with land user choice**: As noted above, it is not possible to fully disentangle the joint effects of land user choice and soil potential. Typically the most fertile soils are selected for growing crops, while less fertile soils are left to forest. For this reason our figures likely underestimate the SOC gain from converting agricultural land to forest, and underestimate the loss from converting forest to cropland.

Applications and Significance

This enhanced SOC dataset enables:

- Improved carbon accounting for conservation and climate programs
- Improved estimates of carbon sequestration potential
- Better targeting of soil health interventions
- Support for natural climate solution initiatives
- Baseline data for monitoring SOC changes over time

By providing SOC estimates for all land covers within each mapping unit, this dataset represents a substantial improvement over existing SSURGO-only estimates, particularly for the heterogeneous landscapes of Massachusetts, where multiple land uses occur within single soil mapping units.

Next Steps

Immediate next steps:

- Make the data layers available on Regenerative Design Groups MassHealthySoils.org website.
- Work with the EOEEA and MassGIS to make the GIS data layers available on MassGIS.
- Present uses and limitations of the datasets to practitioners.

Additional potential long-range next steps:

- Update with newer land cover data as it becomes available
- Validate and calibrate as new soil carbon data becomes available
- Integrate with carbon monitoring programs to track changes over time
- Enhance ability to distinguish agricultural subcategories
- Develop user-friendly interfaces for non-technical users
- Develop a method for crosswalking landcover categories from the 2016 Land Use Land Cover categories (used in this analysis) to the 30m National Land Cover Data (released annually) in order to track annual change.

Conclusion

This project successfully addresses critical gaps in SSURGO's land cover-specific SOC data through innovative imputation methods and integration with high-resolution land cover information. The resulting 1m resolution SOC maps provide the most detailed and comprehensive estimates currently available for Massachusetts, offering valuable data for carbon management, conservation planning, and climate change mitigation efforts. While limitations exist, particularly regarding temporal alignment and imputation uncertainty, the methodology provides a robust framework that could be adapted for other states seeking to enhance their SOC mapping capabilities.

Funding

This project was funded by a Healthy Soils Challenge Grant (Ref# C140) from the Massachusetts Executive Office of Energy and Environmental Affairs.

Acknowledgements

This project would not have been possible without the generous collaboration of Maggie Payne (USDA-NRCS), David Zimmerman (USDA-NRCS), Morgan McKee (USDA-NRCS), Dominique Pahlavan (EOEEA), and Tom Anderson (EOEEA).

Technical Appendix - Reference Tables

The following table shows total acreage, carbon stock, and carbon density at 3 depths for each land cover class in Massachusetts. The arithmetic mean is the simple standard mean for that landclass: the total estimated carbon at the state level, divided by the total acres. We believe the bootstrapped geometric mean is a better representation of typical values for the land cover class, as it naturally accounts for variation and resists the influence of a small minority of high-density outliers.

Table A1. Soil Carbon Density and Total Soil Carbon by Land Cover and Depth

				Carbon Density (mt/acre)		
Land Cover	**Acres**	**Depth (cm)**	**Carbon (mt)**	**Arithmetic mean**	**Bootstrapped geometric mean**	**Bootstrapped 95% CI**
Evergreen Forest	1,198,133	0-30	69,746,933	58	56	55-58
		0-100	86,109,928	72	69	67-72
		0-999	89,847,635	75	72	69-75
Deciduous Forest	1,820,707	0-30	112,861,749	62	60	58-61
		0-100	138,313,887	76	73	71-76
		0-999	144,378,710	79	76	74-79
Scrub/Shrub	66,808	0-30	2,757,325	41	37	34-39
		0-100	3,419,455	51	47	44-50
		0-999	3,660,834	55	51	48-54
Cultivated Land	61,784	0-30	1,786,770	29	27	23-29
		0-100	2,345,436	38	37	35-38
		0-999	2,629,635	43	41	39-42
Cultivated Land (Cranberry)	11,831	0-30	276,667	23	17	13-45
		0-100	2,175,923	184	147	62-206
		0-999	3,714,852	314	218	97-401
Pasture/Hay	132,250	0-30	3,921,909	30	29	28-29
		0-100	5,427,225	41	40	39-41
		0-999	5,992,825	45	44	43-45
Grassland	132,300	0-30	3,149,393	24	20	18-22
		0-100	4,335,029	33	29	27-30
		0-999	4,860,565	37	33	30-34

Developed Open Space	438,578	0-30	9,603,170	22	19	18-20
		0-100	13,057,001	30	27	25-28
		0-999	14,900,931	34	31	30-32
Impervious	475,711	0-30	0	0		
		0-100	3,619,990	8	7	6.7-7.3
		0-999	5,784,778	12	11	11-12
Unconsolidated Shore	31,291	0-30	39,937	1	0.95	0.44-1.3
		0-100	34,834	1	0.91	0.56-1.1
		0-999	34,200	1	0.93	0.55-1.1
Bare Land	38,703	0-30	133	0	0	0-0
		0-100	126	0	0	0-0
		0-999	129	0	0	0-0
Estuarine Forested	69	0-30	5,228	76	75	65-89
		0-100	14,084	205	187	138-234
		0-999	22,758	331	281	165-356
Palustrine Forested	407,599	0-30	29,965,225	74	71	69-75
		0-100	56,726,242	139	118	107-133
		0-999	75,553,599	185	143	124-167
Estuarine Scrub/Shrub	403	0-30	10,653	26	20	16-37
		0-100	99,189	246	243	219-256
		0-999	158,941	394	384	331-408
Palustrine Scrub/Shrub	50,366	0-30	2,585,086	51	35	29-43
		0-100	10,480,537	208	192	177-209
		0-999	17,681,388	351	289	246-335
Estuarine Emergent	45,579	0-30	5,305,379	116	116	114-117
		0-100	15,214,881	334	330	314-339
		0-999	21,812,875	479	475	456-495
Palustrine Emergent	89,439	0-30	9,297,923	104	103	98-106
		0-100	22,061,641	247	219	191-244
		0-999	34,065,113	381	360	331-386
TOTAL	5,001,550	0-30	251,313,479	50	45	43-46
		0-100	363,435,410	73	49	46-52
		0-999	425,099,766	85	55	53-58

The table below shows the crosswalk categories used to connect the SSURGO soil component data with the 2016 Land Cover classes (similar to the classes used in the National Land Cover Dataset). The SSURGO identifiers are drawn largely from the earthcovkind1 and earthcovkind2, representing general and specific land cover classes, respectively. The Cropscape data layer was used only to distinguish cultivated wetlands – i.e. cranberry bogs – from other forms of agriculture. Cranberry bogs have distinct identifiers within the MA SSURGO data but not in the 2016 Land Cover.

Table A2. SSURGO to Land Use/Land Cover Crosswalk

2016 Landcover/NLCD Category	**Cropscape**	**Crosswalk category**	**SSURGO Identifiers**
Impervious	not used	**imperv**	Derived: pooled_herb minus 30 cm
Developed Open Space	not used	**other_herb**	Grass/herbaceous cover:Other grass/herbaceous
Cultivated Land	agriculture except cranberries	**cultivated**	Grassland/herbaceous cover:Row crop
Cultivated Land (Cranberry)	cranberry bog	**wet_cult**	Shrub cover:Crop vines
Pasture/Hay	not used	**pasture_hay**	Grassland/herbaceous cover:Tame pastureland with Grassland/herbaceous cover:Hayland
Grassland	not used	**pooled_herb**	Grass/herbaceous (all) and not Very Poorly Drained
Deciduous Forest	not used	**forest**	Tree cover & not Very Poorly Drained
Evergreen Forest	not used	**forest**	Tree cover & not Very Poorly Drained
Scrub/Shrub	not used	**shrub**	Derived: mean of forest and pasture_hay
Palustrine Forested Wetland (C-CAP)	not used	**wet_tree**	Tree cover & Poorly or Very Poorly Drained
Palustrine Scrub/Shrub	not used	**wet_shrub**	Shrub cover & Very Poorly Drained

Wetland (C-CAP)			
Palustrine Emergent Wetland (C-CAP)	not used	**wet_herb**	Grass/herbaceous cover & Very Poorly Drained
Estuarine Forested Wetland (C-CAP)	not used	**wet_tree**	Tree cover & Poorly or Very Poorly Drained
Estuarine Scrub/Shrub Wetland (C-CAP)	not used	**wet_shrub**	Shrub cover & Very Poorly Drained
Estuarine Emergent Wetland (C-CAP)	not used	**wet_herb**	Grass/herbaceous cover & Very Poorly Drained
Unconsolidated Shore	not used	**beach**	Beach (from compname), cover is Barren:Sand and Gravel
Bare Land	not used	**barren**	Barren, except for identified as beach from compname, above
Open Water	not used	**Drop**	None
Palustrine Aquatic Bed (C-CAP)	not used	**Drop**	None
Estuarine Aquatic Bed (C-CAP)	not used	**Drop**	None

The table below shows the levels from which data was drawn in the hierarchical imputation process. Values for each level are in percent of the total land cover area for the state. The soil and regional characteristics corresponding to each level are shown in Table A4. Cells have conditional formatting within each land cover column to highlight levels from which data was more extensively derived for each land cover.

Table A3. Hierarchical Imputation: Source Level of Data By Land Cover and Percent Area

	Evergreen Forest	Deciduous Forest	Cultivated Land (Cranberry)	Estuarine Forested	Palustrine Forested	Estuarine Scrub/Shrub	Palustrine Scrub/Shrub	Estuarine Emergent	Palustrine Emergent
Acres (this method)	228,769	395,936	1,948	37	159,739	137	26,427	8,782	82,748
% of land cover (this method)	19.0	21.7	16.3	53.2	39.3	34.1	52.5	19.2	92.8
Level 1	2.2	2.7	0.0	32.3	2.4	0.0	4.5	0.2	6.9
Level 2	0.2	0.3	0.0	0.0	0.4	4.2	2.2	0.0	7.7
Level 3	0.9	1.1	0.0	0.9	1.3	0.3	5.6	0.0	0.2
Level 4	0.1	0.0	0.0	0.0	0.0	0.8	0.8	0.0	0.0

Level 8	0.8	0.8	0.1	0.0	0.0	0.0	0.0	1.3	34.4
Level 9	0.0	0.0	0.0	0.5	0.1	0.0	0.0	0.0	0.0
Level 10	0.0	0.0	0.0	0.5	0.0	0.2	2.9	0.0	2.0
Level 11	8.4	8.6	4.8	14.6	27.3	1.1	2.2	0.0	0.7
Level 12	0.0	0.1	0.0	0.0	0.0	0.1	0.2	0.0	0.8
Level 13	0.6	0.6	0.0	0.6	0.7	0.1	1.9	0.7	1.3
Level 15	0.0	0.0	0.0	0.0	1.1	0.0	0.0	0.0	0.0
Level 17	0.0	0.0	0.0	0.0	0.0	0.0	0.2	0.0	0.0
Level 19	0.0	0.0	0.0	0.0	0.7	0.0	0.0	0.0	0.0
Level 23	0.0	0.0	0.0	2.6	2.0	0.2	0.7	0.0	0.0
Level 25	0.1	0.2	0.0	0.0	0.2	0.0	0.0	0.0	0.0
Level 26	0.1	0.1	0.8	0.0	0.1	5.7	7.3	0.2	6.1
Level 27	0.0	0.0	0.2	0.0	0.0	0.0	0.0	0.0	0.0
Level 28	0.0	0.0	0.3	0.0	0.0	0.0	0.0	0.4	0.7
Level 29	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.2
Level 30	0.0	0.0	0.0	0.0	0.0	1.4	10.2	0.6	10.4
Level 32	2.7	4.8	0.0	0.2	0.6	2.7	4.3	0.0	0.1
Level 33	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.4	1.2
Level 34	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.1
Level 35	2.9	2.4	10.1	1.0	2.4	17.3	9.5	15.4	15.8
Level 36	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	4.2

The table below shows the origin of data in the hierarchical imputation process, expressed as a percent of total land area and cumulative percent. Relevant soil and regional grouping variables are shown for each level. The percent by area column has conditional formatting to highlight levels from which more data was derived. The single most heavily used level in the hierarchy was Level 11, from which data was drawn for 370,822 acres, based on the grouping variables deposit_detail, parent_detail, and texture. Individual levels for each grouping variable are shown in Table A5

Table A4. Hierarchical Imputation: Source Level by Percent of Total Land Area

Level	Acres	% by area	Cumulative %	Grouping Variables
1	94,566	1.89	1.89	deposit_detail, parent_detail, texture, dc, taxtempregime
2	17,186	0.34	2.23	deposit_detail, parent_detail, texture, dc
3	39,901	0.8	3.03	deposit_detail, parent_detail, text_3bin, dc
4	2,482	0.05	3.08	deposit_detail, parent_broad, texture, dc
5	0	0	3.08	deposit_detail, parent_broad, text_3bin, dc

6	0	0	3.08	deposit_broad, parent_detail, texture, dc
7	0	0	3.08	deposit_broad, parent_detail, text_3bin, dc
8	55,119	1.1	4.18	deposit_broad, parent_broad, texture, dc
9	277	0.01	4.19	deposit_broad, parent_broad, text_3bin, dc
10	3,333	0.07	4.26	deposit_detail, parent_detail, texture, dc_3bin
11	370,822	7.4	11.66	deposit_detail, parent_detail, texture
12	2,117	0.04	11.7	deposit_detail, parent_detail, text_3bin, dc_3bin
13	23,504	0.47	12.17	deposit_detail, parent_detail, text_3bin
14	0	0	12.17	deposit_detail, parent_broad, texture, dc_3bin
15	4,640	0.09	12.26	deposit_detail, parent_broad, texture
16	0	0	12.26	deposit_detail, parent_broad, text_3bin, dc_3bin
17	270	0.01	12.27	deposit_detail, parent_broad, text_3bin
18	0	0	12.27	deposit_broad, parent_detail, texture, dc_3bin
19	2,786	0.06	12.33	deposit_broad, parent_detail, texture
20	0	0	12.33	deposit_broad, parent_detail, text_3bin, dc_3bin
21	0	0	12.33	deposit_broad, parent_detail, text_3bin
22	0	0	12.33	deposit_broad, parent_broad, texture, dc_3bin
23	8,601	0.17	12.5	deposit_broad, parent_broad, texture
24	0	0	12.5	deposit_broad, parent_broad, text_3bin, dc_3bin
25	5,896	0.12	12.62	deposit_broad, parent_broad, text_3bin
26	12,395	0.25	12.87	text_3bin, dc
27	28	0	12.87	deposit_broad, dc
28	820	0.02	12.89	taxtempregime, dc
29	144	0	12.89	lrr_name, dc
30	14,709	0.29	13.18	text_3bin, dc_3bin
31	0	0	13.18	deposit_broad, dc_3bin
32	124,156	2.48	15.66	deposit_broad
33	1,257	0.03	15.69	mlra_name, dc_3bin
34	118	0	15.69	lrr_name, dc_3bin
35	115,634	2.31	18	mlra_name
36	3,763	0.08	18.08	lrr_name

The below table shows the individual levels within each grouping variable, including missing values. The number of times each level appears in the dataset follows each level name in parentheses.

Table A5. Hierarchical Search Variables and Values

Variable	Values
deposit_detail (21)	Ablation till (72), Alluvium (270), Basal till (219), Beach sand (54), Eolian deposits (1713), Estuarine deposits (21), Glaciofluvial deposits (1434), Glaciolacustrine deposits (174), Glaciomarine deposits (3), Herbaceous organic material (105), Human-transported material (165), Lacustrine deposits (27), Lodgment till (1704), Marine deposits (48), Melt-out till (561), Organic material (168), Outwash (114), Supraglacial meltout till (177), Supraglacial till (432), Till (21), NA (951)
deposit_broad (8)	Anthropogenic (165), Freshwater_Deposits (270), Glacial_Direct (3186), Glacial_Water (1722), Marine_Deposits (48), Organic_Deposits (273), Wind_Deposits (1713), NA (1056)
parent_detail (41)	Ablation till (60), Alluvium (252), Basal till (78), basalt (24), basalt_gneiss_granite (12), basalt_sandstone_shale (12), Beach sand (54), conglomerate (27), conglomerate_sandstone (24), dolomite_limestone_schist (3), Eolian deposits (1710), Estuarine deposits (21), Glaciofluvial deposits (612), Glaciolacustrine deposits (141), Glaciomarine deposits (3), gneiss (120), gneiss_granite (288), gneiss_granite_mica_phyllite_schist (147), gneiss_granite_mica_schist (12), gneiss_granite_quartzite (114), gneiss_granite_quartzite_schist (6), gneiss_granite_schist (2694), gneiss_schist (30), granite (39), granite_mica_phyllite_schist (99), Herbaceous organic material (105), Human-transported material (165), Lacustrine deposits (27), limestone (27), Lodgment till (18), Marine deposits (45), mica_schist (117), Organic material (168), phyllite (33), phyllite_schist (162), quartzite (6), schist (159), slate (3), Supraglacial meltout till (123), Supraglacial till (6), NA (687)
parent_broad (14)	Anthropogenic (165), Freshwater_Deposits (240), Glacial_Direct (267), Glacial_Water (699), Igneous (372), Igneous_Metamorphic (3237), Igneous_Sedimentary (15), Marine_Deposits (45), Metamorphic (651), Mixed_All (36), Organic_Deposits (273), Sedimentary (78), Wind_Deposits (1710), NA (645)

texture (15)	beach sand (54), Clayey (9), Coarse-loamy (3147), Coarse-silty (213), Fine-loamy (99), Fine-silty (12), Gravelly (12), Loamy (1596), organic (273), Sandy (1275), Sandy and gravelly (594), Sandy and loamy (6), Silty (132), Silty and clayey (42), NA (969)
text_3bin (6)	beach sand (54), coarse (1887), fine (9), moderate (5241), organic (273), NA (969)
dc (7)	Excessively drained (1026), Moderately well drained (1383), Poorly drained (522), Somewhat excessively drained (639), Very poorly drained (540), Well drained (3174), NA (1149)
dc_3bin (4)	excessively (1665), poorly (1062), well (4557), NA (1149)
dc_2bin (3)	poorly (1062), well (6222), NA (1149)
taxtempregime (3)	frigid (843), mesic (6489), NA (1101)
mlra_name (5)	Connecticut Valley (1155); Long Island-Cape Cod Coastal Lowland (1044); New England and Eastern New York Upland, Northern Part (936); New England and Eastern New York Upland, Southern Part (5139); Northeastern Mountains (159); NA (0)
lrr_name (2)	Northeastern Forage and Forest Region (7392), Northern Atlantic Slope Diversified Farming Region (1041), NA (0)

The table below shows the estimated SOC land cover conversion factors for upland land cover classes in Massachusetts. Note concerns about the Cultivated-Developed Open Space factors discussed in the limitations section above. This table also appears as a separate deliverable, “Land Cover SOC Conversion Table.csv” and is reproduced here for convenience.

Table A6. Land Cover Conversion Factors

Land Cover: From	Land Cover: To	Conversion Multiplier	95% CI (low)	95% CI (high)
Impervious	Evergreen Forest	9.9	9.6	10.2
Impervious	Deciduous Forest	9.8	9.5	10.1
Impervious	Scrub/Shrub	7.7	7.5	7.9
Impervious	Pasture/Hay	5.4	5.2	5.6
Impervious	Cultivated Land	5.2	5.0	5.5
Impervious	Grassland	4.8	4.6	4.9
Impervious	Developed Open Space	3.9	3.8	4.0
Developed Open Space	Evergreen Forest	2.6	2.6	2.7

Developed Open Space	Deciduous Forest	2.6	2.5	2.7
Developed Open Space	Scrub/Shrub	2.1	2.0	2.1
Grassland	Deciduous Forest	2.0	2.0	2.0
Grassland	Evergreen Forest	2.0	2.0	2.0
Cultivated Land	Deciduous Forest	1.8	1.8	1.8
Cultivated Land	Evergreen Forest	1.8	1.8	1.8
Pasture/Hay	Deciduous Forest	1.8	1.8	1.8
Pasture/Hay	Evergreen Forest	1.8	1.8	1.8
Grassland	Scrub/Shrub	1.6	1.6	1.6
Developed Open Space	Pasture/Hay	1.4	1.4	1.5
Pasture/Hay	Scrub/Shrub	1.4	1.4	1.4
Cultivated Land	Scrub/Shrub	1.4	1.4	1.4
Developed Open Space	Cultivated Land	1.4	1.4	1.5
Developed Open Space	Grassland	1.3	1.2	1.3
Scrub/Shrub	Deciduous Forest	1.3	1.3	1.3
Scrub/Shrub	Evergreen Forest	1.3	1.3	1.3
Grassland	Pasture/Hay	1.1	1.1	1.1
Grassland	Cultivated Land	1.1	1.1	1.1
Evergreen Forest	Deciduous Forest	1.1	1.0	1.1
Cultivated Land	Pasture/Hay	1.0	1.0	1.0
Pasture/Hay	Cultivated Land	1.0	1.0	1.0
Deciduous Forest	Evergreen Forest	0.9	0.9	1.0
Cultivated Land	Grassland	0.9	0.9	0.9
Pasture/Hay	Grassland	0.9	0.9	0.9
Evergreen Forest	Scrub/Shrub	0.8	0.8	0.8
Deciduous Forest	Scrub/Shrub	0.8	0.8	0.8
Grassland	Developed Open Space	0.8	0.8	0.8
Scrub/Shrub	Cultivated Land	0.7	0.7	0.7
Cultivated Land	Developed Open Space	0.7	0.7	0.7
Scrub/Shrub	Pasture/Hay	0.7	0.7	0.7
Pasture/Hay	Developed Open Space	0.7	0.7	0.7
Scrub/Shrub	Grassland	0.6	0.6	0.6

Evergreen Forest	Pasture/Hay	0.6	0.6	0.6
Deciduous Forest	Pasture/Hay	0.6	0.5	0.6
Deciduous Forest	Cultivated Land	0.5	0.5	0.6
Evergreen Forest	Cultivated Land	0.5	0.5	0.6
Deciduous Forest	Grassland	0.5	0.5	0.5
Evergreen Forest	Grassland	0.5	0.5	0.5
Scrub/Shrub	Developed Open Space	0.5	0.5	0.5
Deciduous Forest	Developed Open Space	0.4	0.4	0.4
Evergreen Forest	Developed Open Space	0.4	0.4	0.4
Developed Open Space	Impervious	0.3	0.2	0.3
Grassland	Impervious	0.2	0.2	0.2
Cultivated Land	Impervious	0.2	0.2	0.2
Pasture/Hay	Impervious	0.2	0.2	0.2
Scrub/Shrub	Impervious	0.1	0.1	0.1
Deciduous Forest	Impervious	0.1	0.1	0.1
Evergreen Forest	Impervious	0.1	0.1	0.1