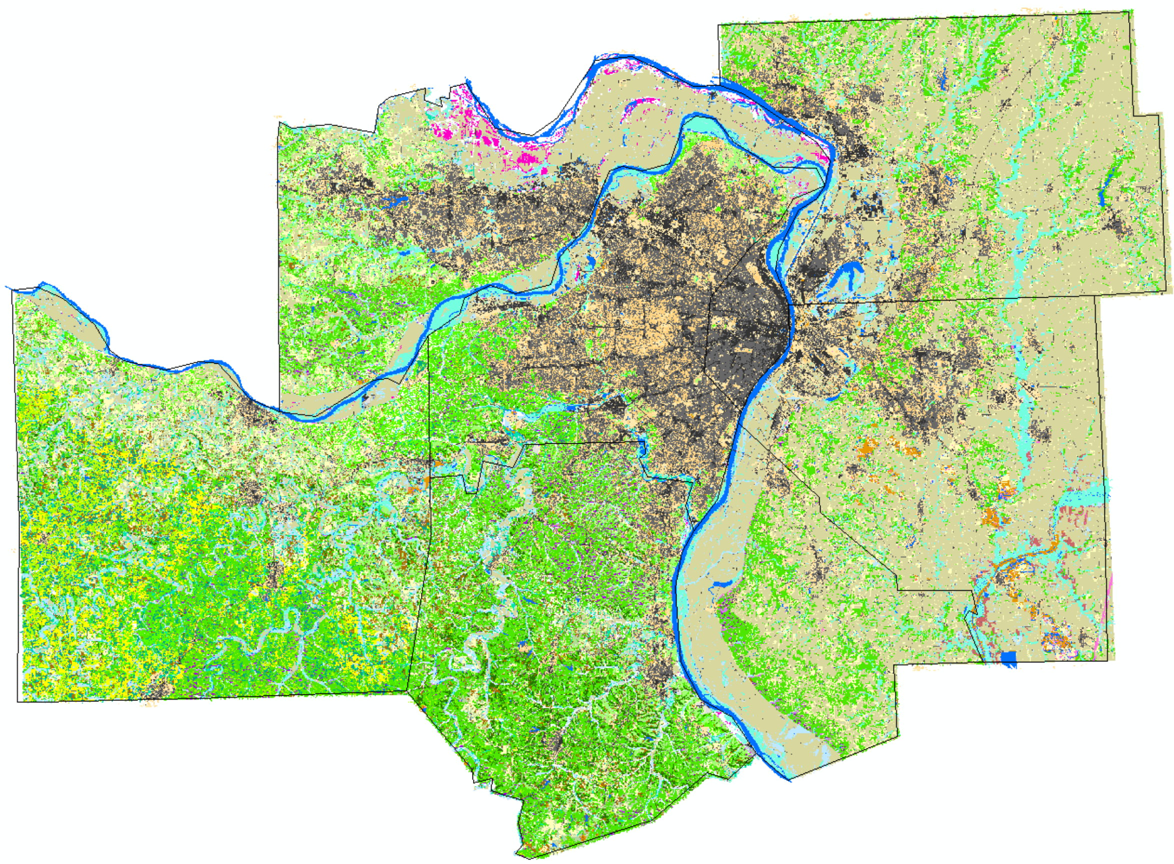
Ecological Approach to Infrastructure Development:
Update and Improvement of St. Louis Regional Vegetation Dataset
Final Report
June 2022

Contact: David Diamond (diamondd@missouri.edu) & Michael Sunde (sundem@missouri.edu)



Citation: Diamond, D.D., M. Sunde, and L.F. Elliott. 2022. Ecological approach to infrastructure development: Update and improvement of St. Louis regional vegetation dataset. Final Report. East-West Gateway Council of Governments, St. Louis.

Cover Image: Ecological Mapping Systems for the East-West Gateway region.

Contents

Figures

Tables

Abstract

A total of 61 terrestrial Ecological Mapping Systems (EMSs) were mapped at 10-meter resolution for the East-West Gateway region. Cropland made up the largest cover type, covering 24.7% of the area, and urban land covers make up 13.5%. Bottomland woodland and forest cover 4.5% of the area. An Ecological Significance (ES) suitability surface (score) ranging from 0 to 10 was also produced for the region. Input datasets used to create the final ES scores included variables related to landscape context, natural community representation, and species representation. The community importance of EMSs was given the highest weight in the preferred ES model, followed by landscape context variables. When the assessing ES model results by class using an equal interval method (10 classes), 14.68% of the region fell within the most important three classes, and 39.13% fell within the least important three classes. Hence, areas of high ecological significance within the St. Louis metropolitan region are uncommon to rare, and are worthy of special consideration.

Methods

Land Cover and Ecological Systems Modeling/Mapping

Land cover was classified using a supervised approach, and the EMS model and map were produced using land cover and geophysical setting information. The general workflow included the following steps:

1. Collection of training data for land use/landcover classification from phot-interpretation of randomly selected points.
2. Remote sensing classification of land use/landcover. Three date mosaics of 10 m resolution Sentinel 2a and 2b imagery and derivatives were used.
3. Development of geophysical setting information, primarily from digital soils datasets (see https://www.nrcs.usda.gov/wps/portal/nrcs/detail/soils/survey/geo/?cid=nrcs142p2_053628) and solar insolation (the amount of sun striking a given spot).
4. Development of image objects (segmentation to form small polygons) from Sentinel 2, 10 m resolution imagery.
5. Attribution of image objects with land use/landcover (from step 3) and geophysical setting information (from step 4).
6. Identification of EMS mapping targets starting with the Ecological Systems Classification (see https://www.natureserve.org/publications/ecological-systems-united-states), followed by iterative adjustments based on work accomplished. This step started with previous work done in the East-West Gateway region.
7. Modeling and mapping of EMS Types using land cover and geophysical setting (Figure 1). The result of this step was a draft final dataset and map.
8. Development of final map and database. This step involved use of ancillary data, especially roads.
9. Development and delivery of final report.

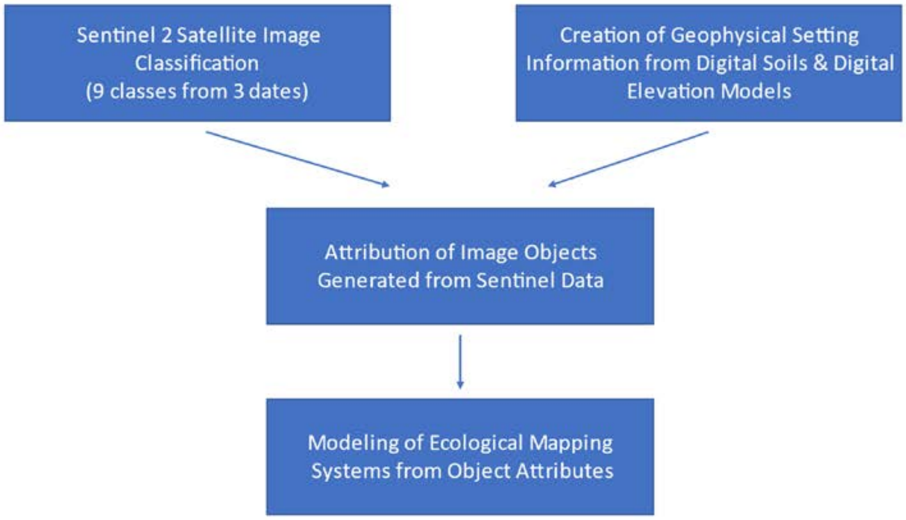


Figure 1. General approach to land cover classification using three dates of imagery, geophysical setting data, and attributed image objects.

Selection and Preparation of Sentinel 2a and 2b Satellite Data

We selected cloud-free imagery, including two growing-season dates and one non-growing season date (leaf on and leaf off), for classification. Dates selected included April 18, 2020; September 16, 2019; and December 9, 2020. We then prepared data-stacks for classification. The stacks included reflectance bands 2, 3, 4, 5, 6, 7, 8, 8A, 11 and 12 for three dates from the Multispectral Instrument (MSI) on the Sentinel-2 satellite (Figure 2). Additionally, the Normalized Difference Vegetation Index (NDVI) and Normalized Difference Water Index (NDWI) were calculated for all three dates. The NDVI is designed to enhance 'greenness' in the vegetation, while the NDWI is complementary to NDVI and enables the classifier to more effectively distinguish moisture differences in vegetation.

Figure 2. Sentinel-2 Multispecral Instrument bands used in data stack for each of three dates of imagery.

Sentinel-2 Bands	Central Wavelength (µm)
Band 1 - Coastal aerosol	0.443
Band 2 - Blue	0.490
Band 3 - Green	0.560
Band 4 - Red	0.665
Band 5 - Vegetation Red Edge	0.705
Band 6 - Vegetation Red Edge	0.740
Band 7 - Vegetation Red Edge	0.783
Band 8 - NIR	0.842
Band 8A - Vegetation Red Edge	0.865
Band 9 - Water vapour	0.945
Band 10 - SWIR - Cirrus	1.375
Band 11 - SWIR	1.610
Band 12 - SWIR	2.190

While bands 2, 3, 4, 8 are native 10-meter resolution bands, 5, 6, 7, 8A, 11 and 12 are captured at 20 meters. To improve mapping accuracy, an image resolution enhancement method, the Sentinel-2 Super-Resolution tool (Brodu, 2017), was used. Each of the image tiles for all three dates were processed using the tool for the six previously mentioned bands. The resultant 10-meter tiles significantly improved image details in those wavelengths (Figure 3).


Figure 2. Comparison of Super-Resolved imagery to raw imagery. Sentinel-1 MSI Band 5 at 20 m (left) and 10 m (right).

Training Data for Supervised Classification

We generated a random sample of 500 points across the study region, and then generated 1701 georeferenced land cover samples from those points using heads-up sampling on-screen (Figure 3). Several photo datasets were viewed to take samples, including Google Satellite Imagery and ESRI World Imagery. These samples were then checked by reference to (viewing of) the original Sentinel 2 imagery by selection of individual 10 m pixels for each sample.

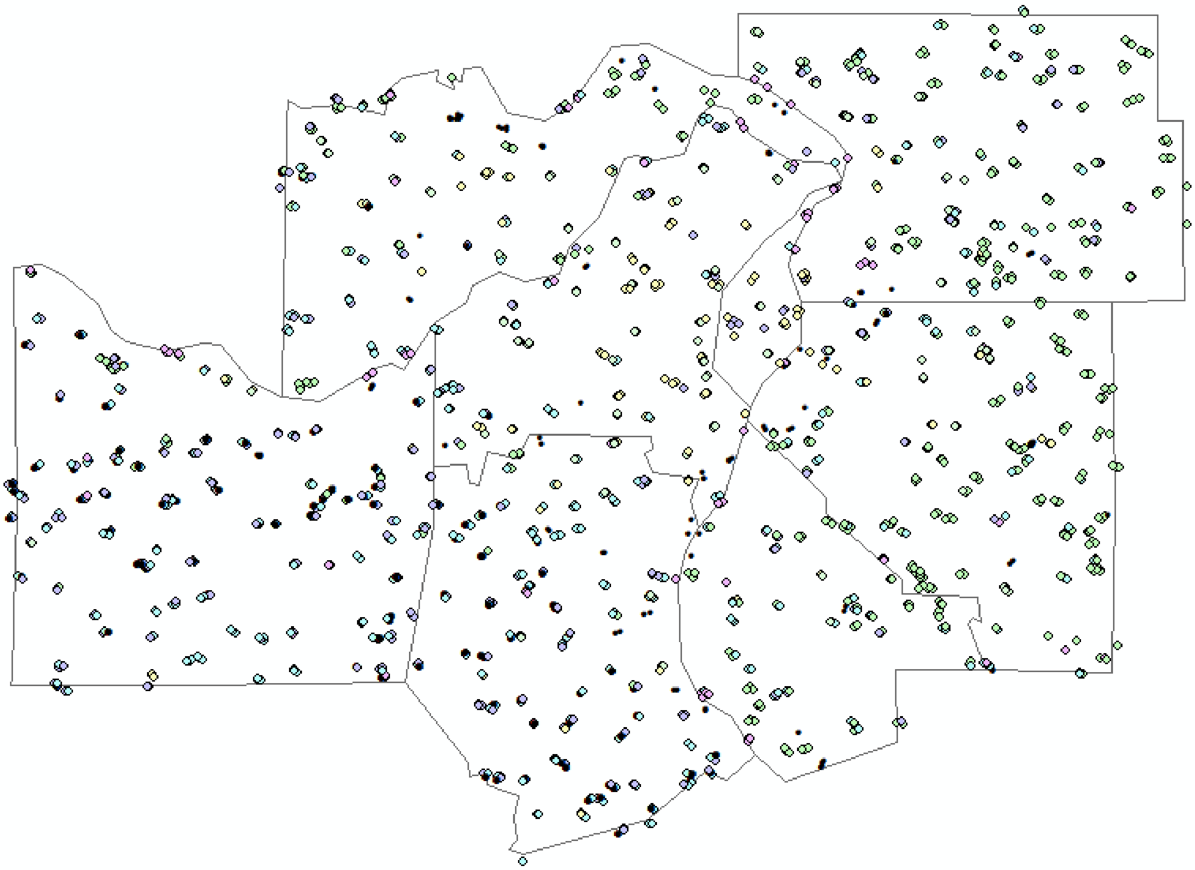


Figure 3. A total of 1701 samples were collected to generate land cover for the East-West Gateway region. Each photo-interpreted sample was checked against Sentinel 2 imagery.

Special attention was paid to NDVI winter values per pixel, as higher values often indicate evergreen cover as opposed to deciduous cover. Hence, each sample was representative of the land cover type attributed to that sample (Table 1). The samples were further viewed over image objects to ensure they were representative for image object classification (see Generating and Refining Land Cover Classification below). Additionally, the sample points were reviewed by a second observer to ensure consistency and eliminate input error that could negatively affect the final classification.

Table 1. Land cover classed for the study area. Developed Wooded and Developed Herbaceous types were modeled based on distance to urban cover. A deciduous shrub type was attempted but was too rare to be accurately mapped.

Code	Land Cover	Description	Area (hectares)	Percent
11	Water	reservoirs, rivers	32091.05	2.6%
21	Urban, Low Intensity	residential roof tops, neighborhood roads	127479.08	10.3%
22	Urban, High Intensity	building tops, parking lots	35958.45	2.9%
31	Barren	quarries, river channel sand and mud	11188.95	0.9%
41	Deciduous Woodland & Forest	cold-deciduous trees	380412.63	30.7%
42	Evergreen Woodland	dense eastern redcedar, pine	42977.63	3.5%
44	Developed, Wooded	trees and shrubs in developed areas (from modeling)	53222.45	4.3%
71	Herbaceous Vegetation	grasslands, forb-dominated areas	164574.68	13.3%
73	Wet Herbaceous Vegetation	wet grasslands, marshes	18298.98	1.5%
74	Developed, Herbaceous	grasses in developed areas (from modeling)	66442.31	5.4%
81	Cropland	Row crops such as wheat or corn	305071.75	24.6%

Generating and Refining Land Cover Classification

We generated two versions of pixel-based land cover classification and one object-based classification. All results are marked improvements over other nationally available datasets, which use 30 m resolution remote sensing data (Figure 4). Pixel-based classification assigns land cover to pixels based on reflectance values of the imagery, NDVI, and NDWI (these two indices are derived from reflectance values). Image object classification was also completed for the region. First image objects were produced using eCognition software (Trimble Geospatial), representing polygons of visually homogeneous areas relative to the satellite imagery. Then information derived from the imagery and indices that characterize each polygon were attributed to the polygons and used in a classification. Both classifications were accomplished using randomForest (Breiman, 2001), one using attributes of each pixel as the predictors, and the other using summarizations of the satellite imagery to polygons as the predictors. The pixel-based classification was selected as more representative of the landscape. While the image object classification was not ultimately used, the results of the pixel-based classification were assigned to image objects based on the majority of pixel classes occurring within each visually homogeneous polygon. This served as a filter to reduce the occurrence of single errant mis-classified pixels, and to provide a polygon framework to facilitate corrections and model application. Building footprints (Microsoft, 2022) and a road network data (IDOT, 2021;

MoDOT, 2021) were burned into the final land-cover classification to improve the overall mapping results. Finally, an urban mask was created from classified urban land cover, and herbaceous and woody land covers within the mask were defined as Developed, Herbaceous and Developed, Wooded land cover (Table 1).

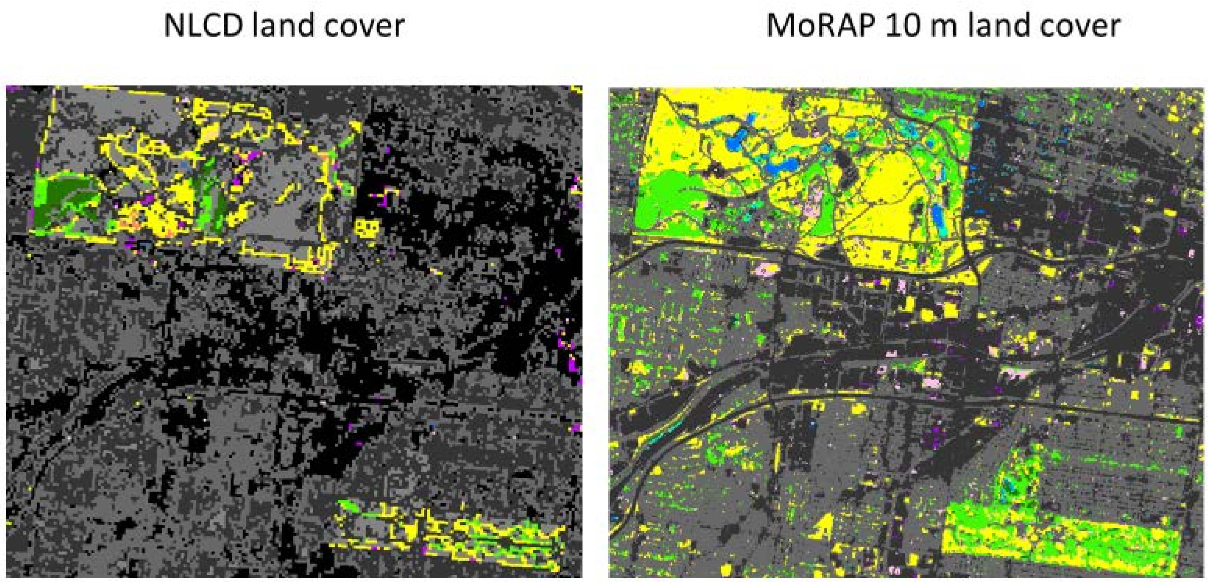


Figure 4. NLDC 30 m land cover (left) versus new 10 m land cover (right). Yellows are herbaceous vegetation, green wooded vegetation, and greys are urban low intensity and urban high intensity.

Geophysical Settings

We used Gridded Soil Survey Geographic (gSSURGO) digital soil map unit (MU) information as the primary source of geophysical setting data (Soil Survey Staff, 2021). MUs are polygons that circumscribe areas of similar soils. Existing ecosite attributes (for Missouri) and other soil MU characteristics (for Illinois. where ecosite information was lacking) were used to group MU polygons into Ecogroups. The Ecogroups represent areas that have similar 'climax' vegetation under natural disturbance regimes. A total of 19 Ecogroups were recognized in Missouri and 11 in Illinois (Figure 5). Fine-resolution digital elevation models (DEMs) were also used to inform geophysical settings. Illinois DEMs were collected in 2012 and 2013, whereas Missouri DEMs are more recent. The DEMs were used to generate solar insolation values. Solar insolation varies with slope, aspect, slope position, and shading (for narrow valley bottoms), with south and west-facing slopes receiving more sun and north and east facing sloped more sun. Solar insolation was calculated ever hour for one day every two weeks between March 15 and October 15 and combined to represent the solar insolation striking a pixel over the growing season. As with land cover, the geophysical setting variables generated were attributed to image objects for Ecological Systems modeling.

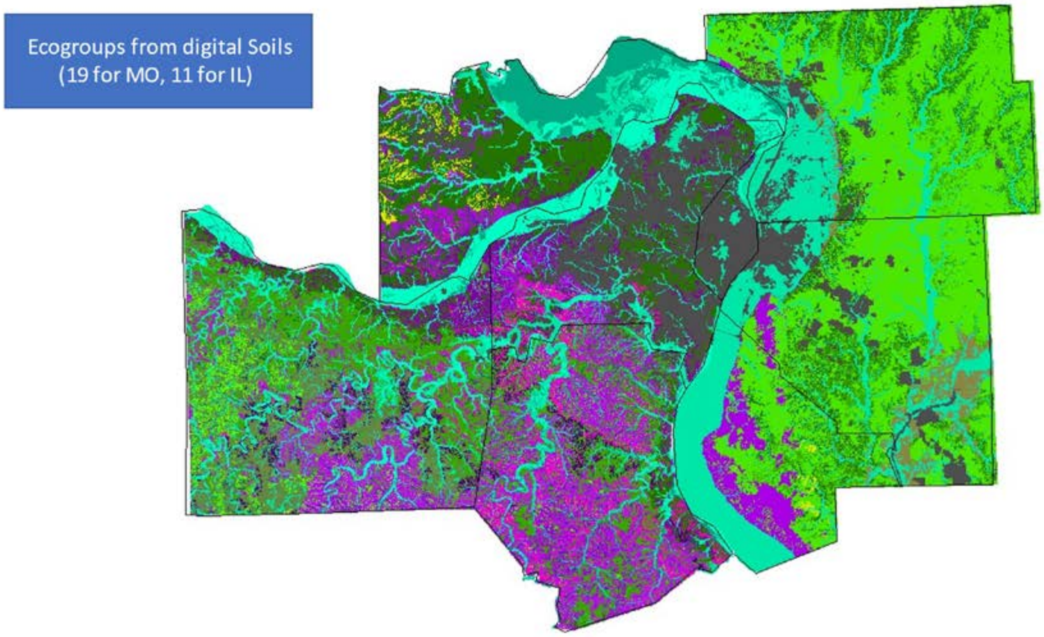


Figure 5. A total of 30 Ecogroups representing areas of broadly similar vegetation under natural conditions were formed from digital soil map unit polygons.

Ecological Systems Modeling, Map Generation, and Interpretation

Ecological Systems were modeled and mapped using land cover plus geophysical setting data attributed to image objects (Figure 6). The primary geophysical setting information was soil map unit Ecogroup, but for some types, solar insolation was also used. For hilly landscapes, the range of solar insolation was defined for a given aggregate EMS type, and the type was further subdivided by assigning, for example, dry, intermediate, and moist types based on solar insolation. The insolation-based types are directly identified in the final EMS classification (see Table 3 in Results). The highest and lowest 5% of solar insolation values for an EMS type were considered dry and moist, respectively, and the rest was considered intermediate. For example, the Illinois Loess and Till: White Oak/Red Oak-Hickory Woodland and Forest was subdivided into three types based on exposure.

Ecological Mapping System Modeling: Land Cover X Ecogroup = EMS Type (cell numbers)

Deciduous Woodland and Forest — Grassland

Ecogroup	11	21	22	31	41	42	71	73
Bottomland Backswamps	21	54	56	58	4	9	6	42
Chert Slopes	21	54	56	29	22	**63**	29	47
Chert Uplands	21	54	56	1	23	52	16	47
Claypan Summit Prairies	21	54	56	1	13	52	17	47
Disturbed Soils	21	54	56	1	50	52	59	47
Fragipan Uplands	21	54	56	1	41	52	16	47
Limestone/Dolomite Glades	21	54	56	29	30	63	29	47
Limestone/Dolomite Slopes	21	54	56	29	26	**63**	29	47
Limestone/Dolomite Uplands	21	54	56	1	28	52	16	47
Loamy Bottomland Forests	21	54	56	58	2	9	6	6
Loamy Bottomland Terraces/Footslopes	21	54	56	58	3	9	6	6
Loamy Upland Woodlands	21	54	56	1	34	52	17	47

Figure 6. Modeling of Ecological Mapping System (EMS) types relied primarily on geophysical setting (Ecogroup, left column) and land cover (top row). Numbers in the cells (two are circled in red) represent EMS type. See Appendix for models and EMS numbers.

Ecological Significance Modeling/Mapping

Ecological Significance (ES) was defined using a multi-criteria evaluation approach (Carver, 1991), wherein 10-meter grids representing different ecological variables were stacked and weighted in order to create a suitability surface. Variables were representative of landscape context, natural community importance, and species representation. A total of 10 datasets were used (Table 2). While the majority of variables used were scaled and scored from 0 to 1 to form a continuous 10 m pixel-based surface, aquatic ecological system priority watersheds and sinkholes were categorical and scored either 0 (not priority) or 1 (priority). For Glade Frequency and Species Element Occurrence Frequency, the number of occurrences of each element (glade or species) were counted within an EMS patch and assigned to that patch. These counts were then scaled from 0 to 1, with 1 representing the maximum value within any EMS patch. Weights were then assigned to the variables in different model runs that took the form of (see Table 5):

(Variable 1 * Weight 1) + (V2 * W2) + …. (V10 * W10) = Ecological Significance Score

A preferred ecological significance model was selected by combining an iterative process with expert opinion by viewing model results on-screen.

Table 2. Datasets used for ecological significance modeling

Dataset	Source	Theme	Notes
Distance to Protected Areas	USGS Protected Area Dataset (PAD-US 2.1)	landscape context	Protected Area GAP Status Code 1 – 3
Cropland Density	MoRAP	landscape context	derived from MoRAP land cover by weighted kernel density method
Natural Vegetation Density	MoRAP	landscape context	derived from MoRAP land cover by weighted kernel density method
Urban Land Density	MoRAP	landscape context	derived from MoRAP land cover by weighted kernel density method
Aquatic Ecological System Priority Watersheds	MDC and IL state efforts	natural communities	selected watershed are best examples of representative landscapes within larger watershed units in MO; Biologically Significant Stream watersheds in IL; assigned 0 or 1
Ecological Mapping System (EMS) Importance	MoRAP	natural communities	mapped at 10 m resolution
Glade Frequency	Nelson in MO; Element Occurrence Records in IL	natural communities	mapped from air photos in MO; ground assessed in IL; frequency assigned to EMS patches
Sinkholes	MO and IL datasets	natural communities	hand-mapped in MO; mapped from Lidar in IL; assigned 0 or 1
Species Element Occurrence Record Frequency	MDC and IL Natural Heritage Datasets	species	frequency assigned to EMS patches; goodness of data depends on how complete the state surveys were
Vertebrate Species Richness	USGS-GAP Analysis Program	species	models depend on 30 m land cover and species range, primarily

Results and Discussion

Ecological Mapping Systems

Deciduous woodland and forest land cover comprised 30.7% of the East-West Gateway region, and cropland covered 24.6% (see Table 3). Sixty-one EMS types were mapped, with 21 types comprising more than 1% of the area (Figure 7, Table 3; see Appendix). Although 40 EMS types each composed less than 1% of the area, 19 of these relatively rare types were among the 22 highest ranked (8, 9) in terms of terrestrial natural community significance (Table 4, see Appendix).

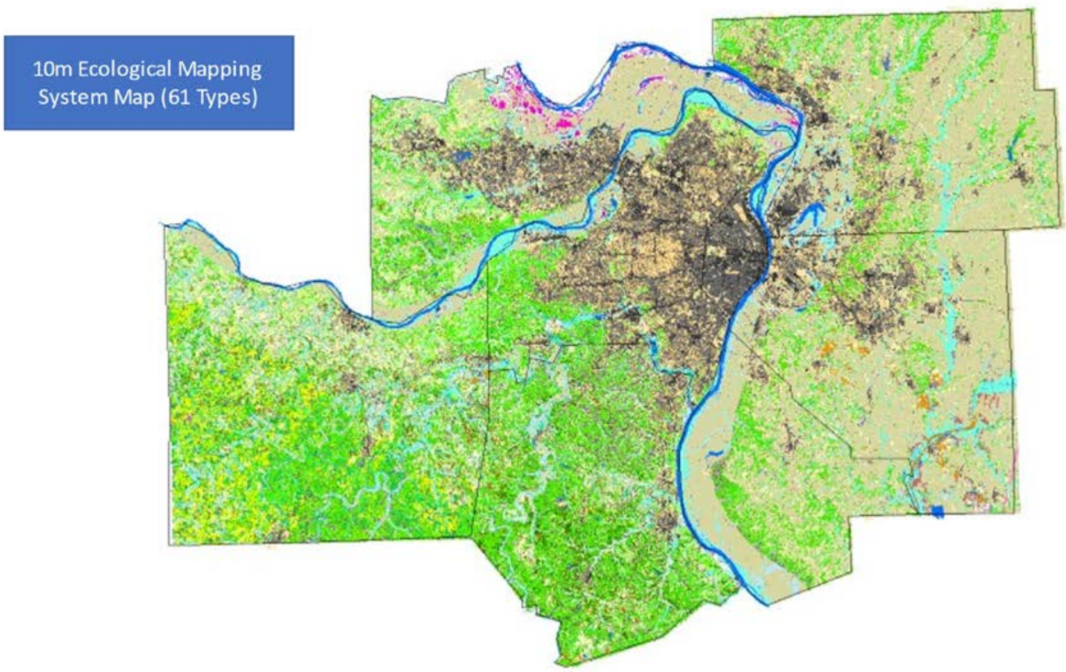


Figure 7. Sixty-one Ecological Mapping System (EMS) types were mapped in the East-West Gateway region. Greens are wooded, tan is cropland, aqua is bottomlands with wooded or herbaceous cover, and purple and pink are other wetland types.

Table 3. Ecological Mapping Systems with more than 1% of the total area for the East-West Gateway region.

Current Vegetation	Area (ha)	Percent
Cropland	294,293	24.7%
Urban High Intensity	125,235	10.5%
Cultural/Disturbance: Upland Loess and Till Grassland	83,106	7.0%
Ozark Highlands: Chert Backslope White Oak/Black Oak-Dogwood Woodland and Forest (intermediate exposure)	74,247	6.2%
Developed Herbaceous	64,360	5.4%
Bottomland Forest: Mixed Bottomland Hardwood Forest	53,545	4.5%
Illinois Loess and Till: White Oak/Red Oak-Hickory Woodland and Forest (intermediate exposure)	52,754	4.4%
Developed Wooded	51,237	4.3%
Bottomland Herbaceous Vegetation	44,866	3.8%
Central Dissected Till Plains: Loess and Till Bur Oak/Post Oak Upland Woodland	37,898	3.2%
Urban Low Intensity	35,541	3.0%
Open Water	30,762	2.6%
Cultural/Disturbance: Upland Limestone/Dolomite and Chert Grassland	23,610	2.0%
Ozark Highlands: Loess and Till Upland Post Oak/White Oak-Black Oak Woodland	19,500	1.6%
Successional Upland Eastern Redcedar Evergreen Woodland and Forest	17,847	1.5%
Ozark Highlands: Loess and Till Backslope White Oak/Black Oak-Hickory Woodland and Forest (intermediate exposure)	16,468	1.4%
Ozark Highlands: Upland Post Oak-Bluestem Flatwoods (wooded)	16,017	1.3%
Ozark Highlands: Chert Upland Mixed Oak Woodlands	15,173	1.3%
Ozark Highlands: Limestone/Dolomite Upland Glade/Chinquapin Oak Woodland Complex (deciduous woods/intermediate exposure)	14,074	1.2%
Ozark Highlands: Limestone/Dolomite Upland Glade/Chinquapin Oak Woodland Complex (juniper or mixed woods/intermediate exposure)	13,983	1.2%
Bottomland Forest: Pin Oak/Bur Oak-Swamp White Oak/Pecan Forest	13,629	1.1%

Table 4. Ecological Mapping Systems that scored as highly important (8 or 9) based on expert opinion. Note that 19 of 22 most important types make up less than 1% of the total area of the region.

Current Vegetation	Area (ha)	Percent	Comm Import
Bottomland Forest: Mixed Bottomland Hardwood Forest	53,545	4.4857%	9
Bottomland Herbaceous Vegetation	44,866	3.7586%	9
Bottomland Forest: Pin Oak/Bur Oak-Swamp White Oak/Pecan Forest	13,629	1.1418%	9
Bottomland Forest: Sycamore, Cottonwood, Elm, Ash Hackberry Riverfront Forest	9,495	0.7954%	9
Bottomland Forest: White Oak/Red Oak-Dogwood/Sycamore Forest	4,302	0.3604%	9
Herbaceous-dominated Wetlands (non-riverine)	3,973	0.3328%	9
Riverine Marsh	3,831	0.3209%	9
Illinois Pin Oak/Post Oak-Hickory Flatwood Forest	2,383	0.1996%	9
Ozark Highlands: Limestone/Dolomite Upland Glade/Chinquapin Oak Woodland Complex (grassy/dry exposure)	29	0.0024%	9
Ozark Highlands: Sandstone Upland Glade/Post Oak Woodland Complex (deciduous woods/dry exposure)	28	0.0023%	9
Ozark Highlands: Sandstone Upland Glade/Post Oak Woodland Complex (grassy/dry exposure)	1	0.0001%	9
Ozark Highlands: Limestone/Dolomite Upland Glade/Chinquapin Oak Woodland Complex (grassy/intermediate to moist exposure)	6,095	0.5106%	8
Ozark Highlands: Limestone/Dolomite Upland Glade/Chinquapin Oak Woodland Complex (deciduous woods/moist exposure exposure)	1,359	0.1139%	8
Ozark Highlands: Sandstone Upland Glade/Post Oak Woodland Complex (grassy/intermediate to dry exposure)	923	0.0773%	8
Ozark Highlands: Limestone/Dolomite Upland Glade/Chinquapin Oak Woodland Complex (deciduous woods/dry exposure)	800	0.0670%	8
Ozark Highlands: Limestone/Dolomite Upland Glade/Chinquapin Oak Woodland Complex (juniper or mixed woods/moist exposure)	598	0.0501%	8
Ozark Highlands: Sandstone Upland Glade/Post Oak Woodland Complex (deciduous woods/intermediate to moist exposure)	256	0.0214%	8
Ozark Highlands: Sandstone Backslope Red Oak/White Oak-/Black Oak-Dogwood Woodland and Forest (dry exposure)	221	0.0185%	8
Ozark Highlands: Sandstone Backslope Red Oak/White Oak-/Black Oak-Dogwood Woodland and Forest (wet exposure)	199	0.0167%	8
Ozark Highlands: Limestone/Dolomite Upland Glade/Chinquapin Oak Woodland Complex (juniper or mixed woods/dry exposure)	138	0.0116%	8
Ozark Highlands: Sandstone Upland Glade/Post Oak Woodland Complex (juniper or mixed woods/dry exposure)	20	0.0017%	8
Ozark Highlands: Sandstone Upland Glade/Post Oak Woodland Complex (deciduous woods/moist exposure)	12	0.0010%	8

Ecological Significance

Seven ES suitability models with different weighting for variables were executed and their results were evaluated on-screen (Table 5). A preferred ecological significance model was selected by expert opinion by viewing model results. In this model, Ecological Mapping System community importance and natural vegetation density were weighted most heavily, whereas the categorical variables had lower weights. When values from the ES suitability surface were grouped into 10 classes, a total of 4.59% of the region was scored within the highest two classes (Table 6). An additional 10.01% was in the third highest class, and 15.21% was in the fourth highest class. A total of 39.13% of the region was in the lowest three classes for ecological significance, and these areas were predominantly urban or cropland land cover (Figure 8).

Table 5. Seven different Ecological Significance suitability models, each with different variables included and different weighting of variables, were executed and viewed on screen. Numbers in cells represent weighting. Cell shading represents theme (see Table 2).

Dataset	Preferred Model	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Distance to Protected Areas	0.10	0	0.10	0	0	0	0
Cropland Density	0.00	0.11	0.10	0.1	0	0	0
Natural Vegetation Density	0.30	0.11	0.10	0.22	0.4	0.35	0.37
Urban Land Density	0.00	0.11	0.10	0.15	0	0	0
Aquatic Ecological System Priority Watersheds	0.03	0.11	0.10	0.03	0.05	0.05	0.03
Ecological Mapping System (EMS) Importance	0.44	0.11	0.10	0.2	0.5	0.45	0.47
Glade Frequency	0.05	0.11	0.10	0.04	0	0.05	0.05
Sinkholes	0.03	0.11	0.10	0.03	0	0.05	0.03
Species Element Occurrence Record Frequency	0.05	0.11	0.10	0.08	0.05	0.05	0.05
Vertebrate Species Richness	0.00	0.11	0.10	0.15	0	0	0

Table 6. Summary of the Ecological Significance suitability surface based on grouping into 10 classes. Higher numbers are more significant.

Significance Class	Value Range	Area (hectares)	Landscape %
1	0.035 – 0.115	16,243	1.37%
2	0.116 – 0.196	158,657	13.34%
3	0.197 – 0.276	290,412	24.42%
4	0.277 – 0.357	137,558	11.57%
5	0.358 – 0.438	91,346	7.68%
6	0.439 – 0.519	139,422	11.72%
7	0.520 – 0.599	180,914	15.21%
8	0.060 – 0.680	119,952	10.09%
9	0.681 – 0.761	50,682	4.26%
10	0.762 – 0.842	3,938	0.33%
	totals	1,189,125	100.00%

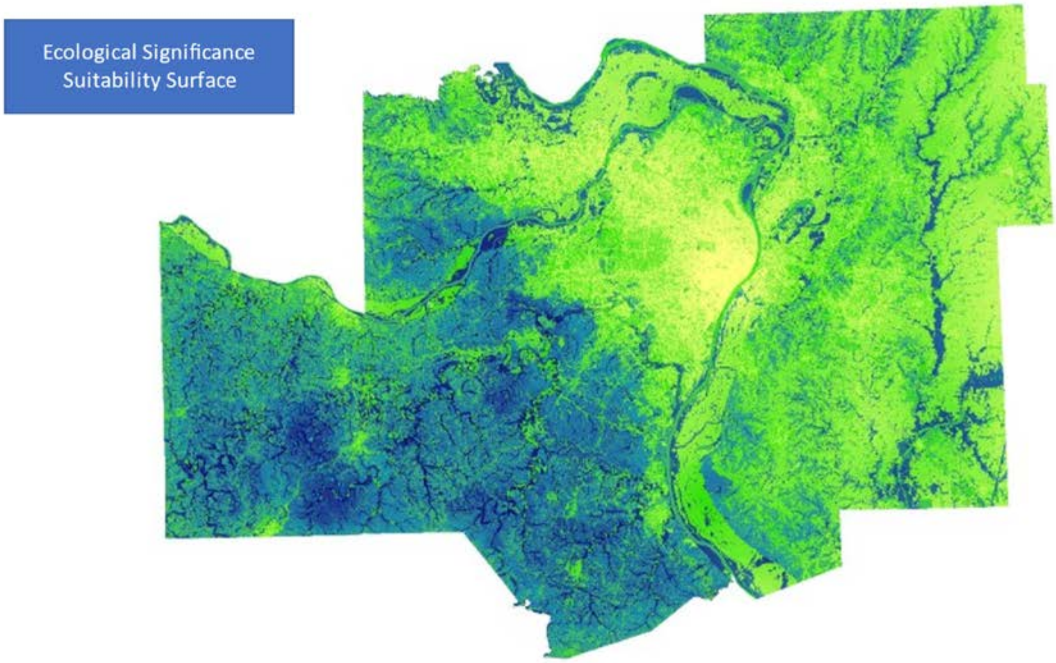


Figure 8. Ecological Significance suitability surface. Dark blues are high significance and yellows are low as scored for the entire region.

Ecological Significance Scoring Caveats

The assessment region and variables used for scoring will influence outcomes of any ecological assessment. For example, variable scoring and categorical grouping (1 to 10) of scores for Ecological Significance relative to Illinois only, excluding Missouri, would show more area in higher significance categories in Illinois. If only urban areas were assessed, non-urban patches within the assessment region would receive higher ranking (Figure 9). Likewise, adding a new scoring variable that, for example, emphasized the importance of natural vegetation patches within an urban setting would up the scores for such patches. A related study designed to emphasize scoring for urban areas is currently underway. The importance of vegetation within an urban setting extends beyond, and possibly is only moderately related to, the ecological significance in terms of native flora and fauna.

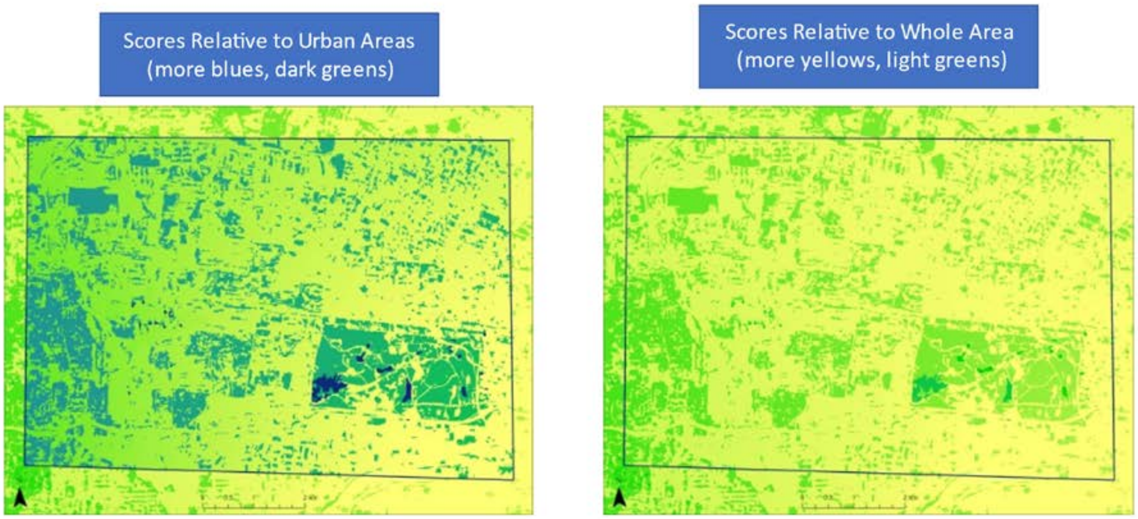


Figure 9. Ecological Significance scores will vary based on the area of concern. For example, natural vegetation in urban areas will score higher when rural areas are excluded from analysis.

References

Breiman, L. (2001). Random forests. Machine Learning. https://doi.org/10.1023/A:1010933404324

Brodu. (2017). Super-Resolving Multiresolution Images With Band-Independent Geometry of Multispectral Pixels. IEEE Transactions on Geoscience and Remote Sensing, 55(8), 4610–4617. https://doi.org/10.1109/TGRS.2017.2694881

Carver, S. J. (1991). Integrating multi-criteria evaluation with geographical information systems. International Journal of Geographical Information Systems, 5(3), 321–339. https://doi.org/10.1080/02693799108927858

Illinois Department of Transportation (IDOT). 2017. District 8 Roads. gis-idot.opendata.arcgis.com/datasets/IDOT::district-8-roads-2017/

Microsoft. Bing Maps Building Footprints. 2022. https://github.com/microsoft/USBuildingFootprints

Missouri Department of Transportation (MoDOT). 2022. Roads/Routes. data-msdis.opendata.arcgis.com/datasets/MSDIS::mo-modot-roads-routes/

Soil Survey Staff. Gridded Soil Survey Geographic (gSSURGO) Database for the Conterminous United States. United States Department of Agriculture, Natural Resources Conservation Service. 2021. Available online at https://gdg.sc.egov.usda.gov/.