Knowledge-Based Soil Inference Modeling for Estimating Soil Productivity and Grain Yield in North-Central Missouri

Fred J. Young*, D. Brent Myers,** and Newell R. Kitchen**
*USDA-NRCS and **USDA-ARS-CSWQRU, respectively, Columbia, Missouri

Abstract

We used ArcSIE (Soil Inference Engine) software to model soils resembling those mapped by NRCS soil survey, for eight 12-digit watersheds in the Central Claypan (MLRA 113) in north-central Missouri. Our source data for modeling was the 10m USGS Digital Elevation Model. Environmental Covariates used in modeling included Relative Position, Slope, Planiform Curvature, Profile Curvature, and Wetness Index. The knowledge-based rules were derived by interpreting the existing SSURGO (soil survey) pattern for the area, and from the expert opinion of the authors. The resulting seven-class model simulated components of soil map units from the local SSURGO data. We assigned the appropriate Missouri Productivity Index value to each modeled soil, and used ArcSIE to create a continuous-surface Productivity Index for each watershed. We compared the SSURGO maps, the modeled soil maps and the continuous-surface PI maps with yield data from combine-mounted yield monitors. Results and discussion focus on comparisons among yield data, SSURGO data, modeling results and environmental covariates, and the implications for knowledge-based inference modeling in future soil survey projects and products.

Objectives

- Use Soil Inference Engine (SIE) to quantitatively model a seven-class soil-landscape pattern in an agricultural area with existing soil survey (SSURGO) data.
- Use the resulting SIE model to create a continuous-surface soil property map, using the Missouri Productivity Index (PI) as the property value.
- Compare the SSURGO, SIE and PI models to thousands of corn yield points randomly selected from yield monitor data on about 25 fields, over 12 years.

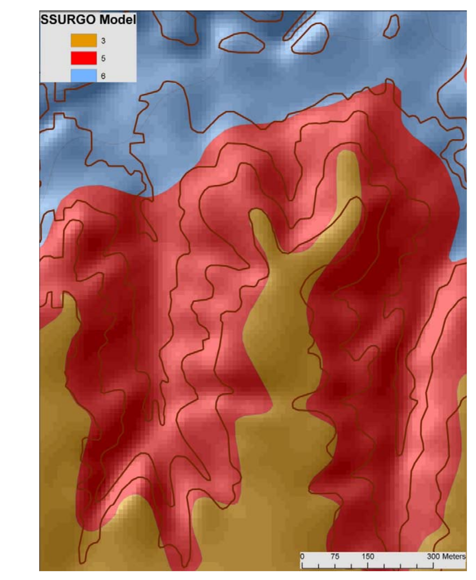
Conclusions

***Soil Models*:**
- The seven-class SIE model is generally a meaningful representation of the soil-landscape relationships at the watershed scale.
  * DEM anomalies adversely affect local area modeling results.
  * Stratigraphic rules would improve the model (e.g., loess always upslope from till).
  * "Minor components" of soil survey map units are revealed by the SIE model.
- The continuous-surface PI models are difficult to evaluate because of the weak soil-yield relationships, but the technique appears promising and worthy of continued testing.

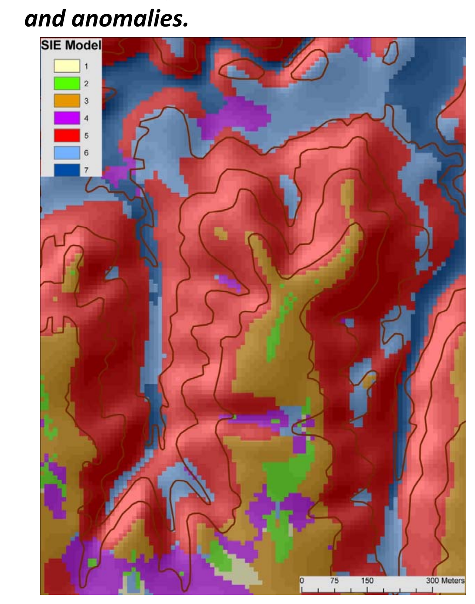
***Soil – Yield Relationships*:**
- Yield variance is high.
- Soil class alone (either via SSURGO or SIE models) is not a useful predictor.
- Year is a strong predictor:
  * Within good years, yields are high on all soils.
  * Within poor years, yields are low on all soils.
  * Within some years, soils affect yields.
- Field appears to be a strong predictor (not fully tested).
- Soil-yield relationships vary with year, and with individual field.
  * Soils are not consistently ranked year to year, or even field to field within a year.

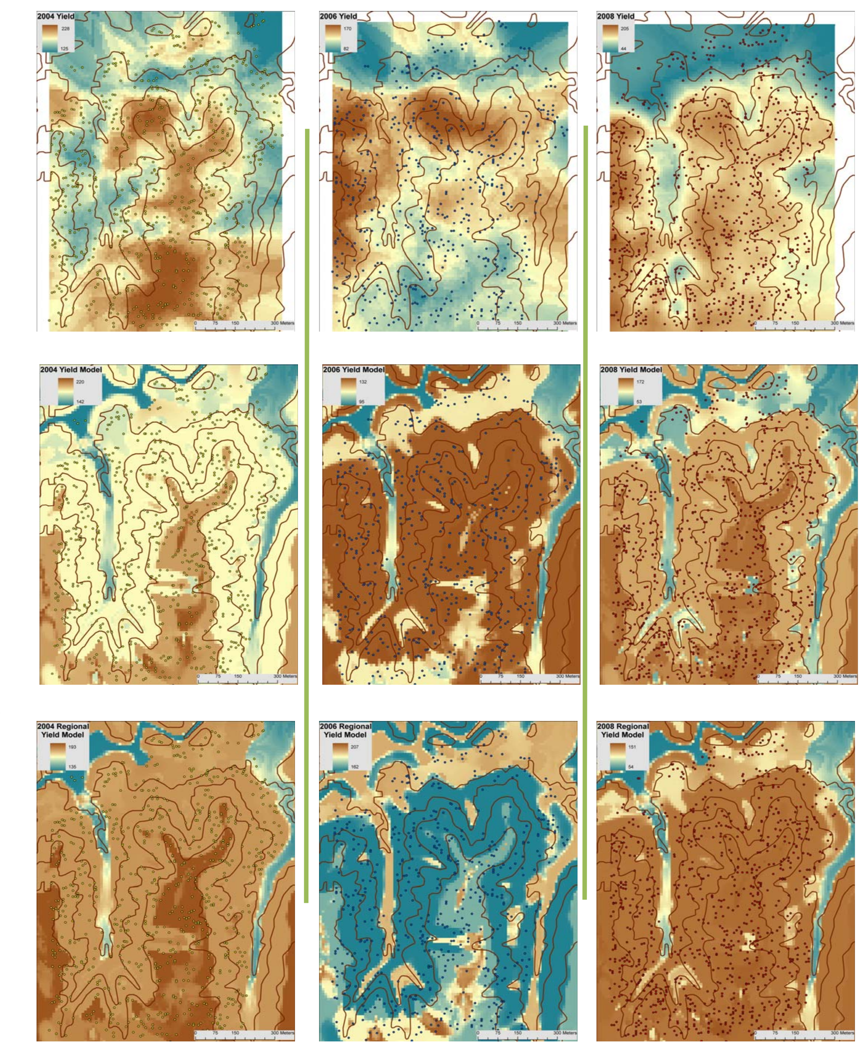
Future Investigations
- Continue to refine regional SIE modeling techniques.
  * Use of LiDAR, additional environmental covariates, etc.
- Select fields that span major soil-landscape units.
- Work with farmers to better understand sources of within-field yield variance.
- Investigate regional yield-weather relationships, and incorporate into a regional Productivity Index.

Within-field Soil – Yield Relationships



SSURGO soil map for the example field, compared to the SIE model "hardened" map. The SIE model shows some meaningful detail not shown in the SSURGO map (i.e., "minor components"), but also contains some artifacts and anomalies.

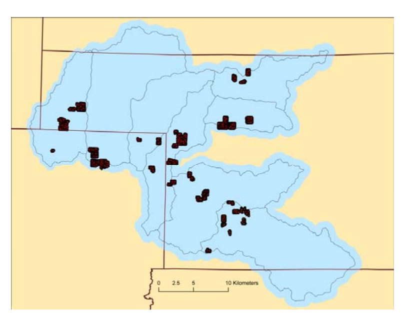




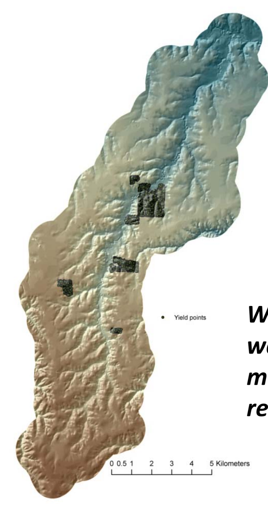
Three years of yield data for a single field.
- Top row is actual yield data (interpolated from points).
- Middle row is SIE continuous-surface modeled yield, using the mean yield for each soil within the field, by year.
- Bottom row is SIE continuous-surface modeled yield, using mean yield for each soil for the entire eight-watershed region, by year.
Yield is somewhat but not entirely related to soil class. Soil-yield relationships vary by year. Regional soil-yield relationships differ from the field-specific relationships (see especially 2006).

Study Area

Location of the eight, 12-digit watershed study area in Missouri.



Location of corn yield data within the study area. Each watershed was buffered by 1km prior to modeling.

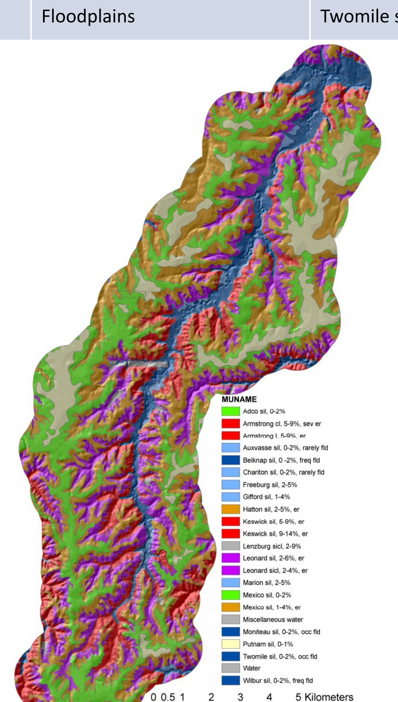


Wabash Lake – Youngs Creek watershed, used to illustrate modeling results and yield relationships.

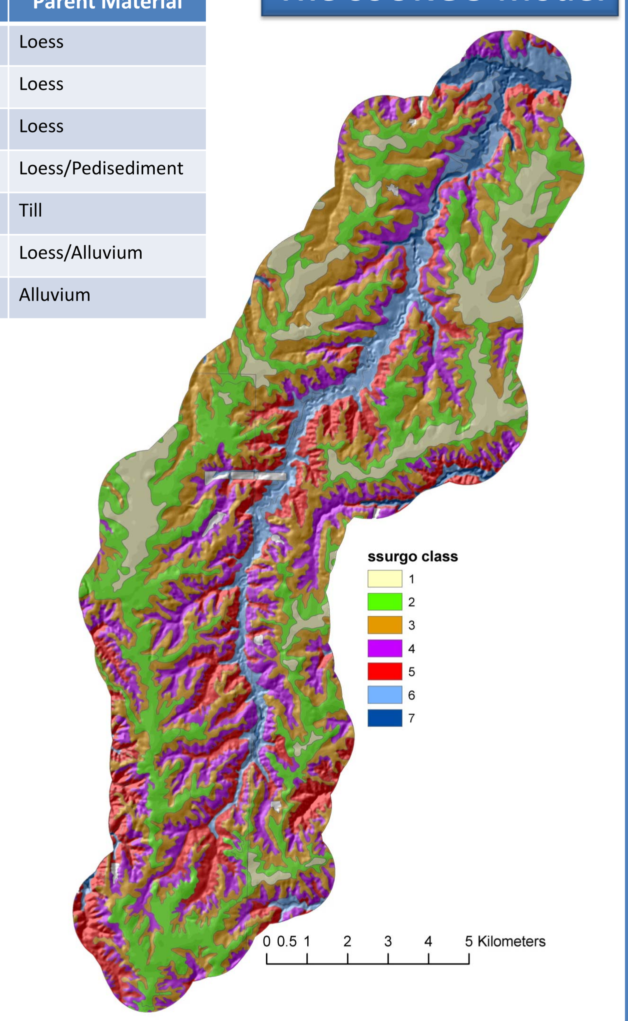
A single field within the watershed, used to illustrate field-specific results for three years of yield data.

The SSURGO Model

Class	Landscape Position	Target Map Unit	Tax. Class	Parent Material
1	Level Interfluve Divides	Putnam sil, 0-1%	F. Vertic Albaqualfs	Loess
2	Broadly Convex Interfluves	Mexico sil, 0-2%	F. Vertic Epiaqualfs	Loess
3	Convex Interfluves & Shoulders	Mexico sil, 1-4%, eroded	F. Vertic Epiaqualfs	Loess
4	Concave Upper Headslopes	Leonard sicl, 2-4%, eroded	F. Vertic Epiaqualfs	Loess/Pedisediment
5	Backslopes	Armstrong l, 5-9%, eroded	F. Aquertic Hapludalfs	Till
6	Footslopes & Terraces	Gifford sil, 1-4%	F. Vertic Epiaqualfs	Loess/Alluvium
7	Floodplains	Twomile sil, 0-2%, occ flood	Fi-si Typic Albaqualfs	Alluvium



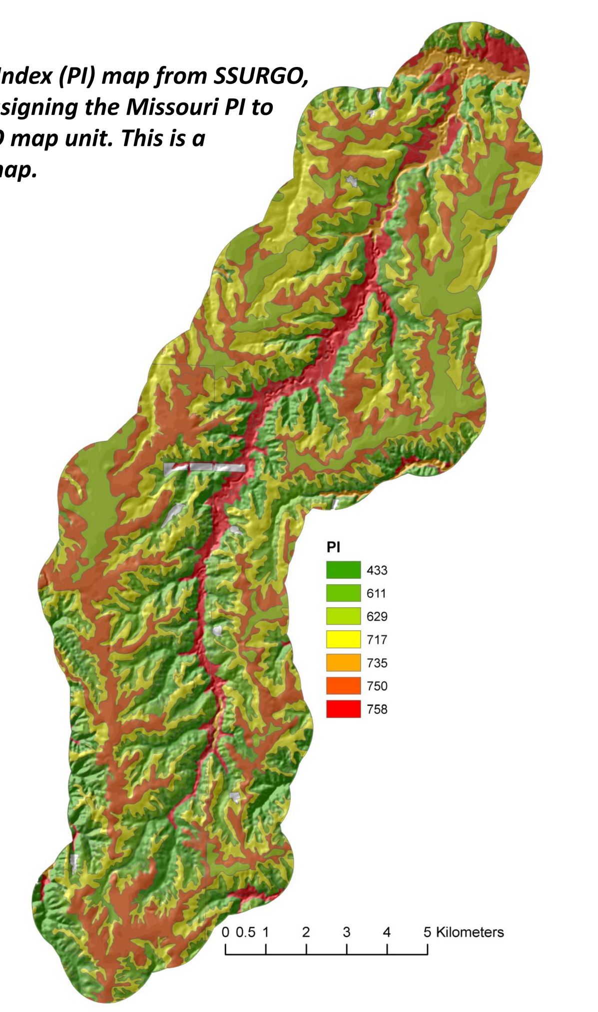
SSURGO map units, color-coded by targeted model class.



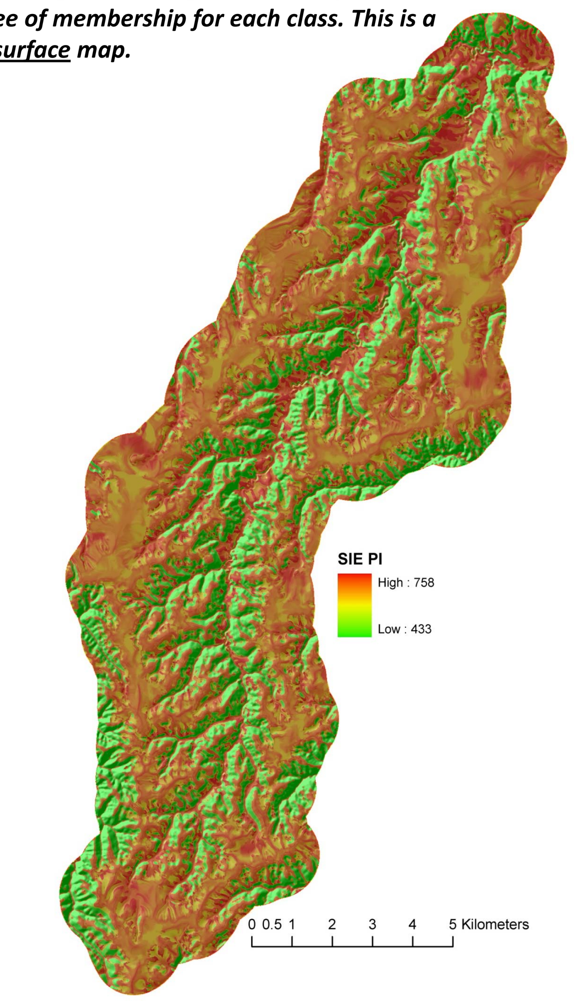
Seven-class map from grouping of SSURGO map units.

The Productivity Index Model

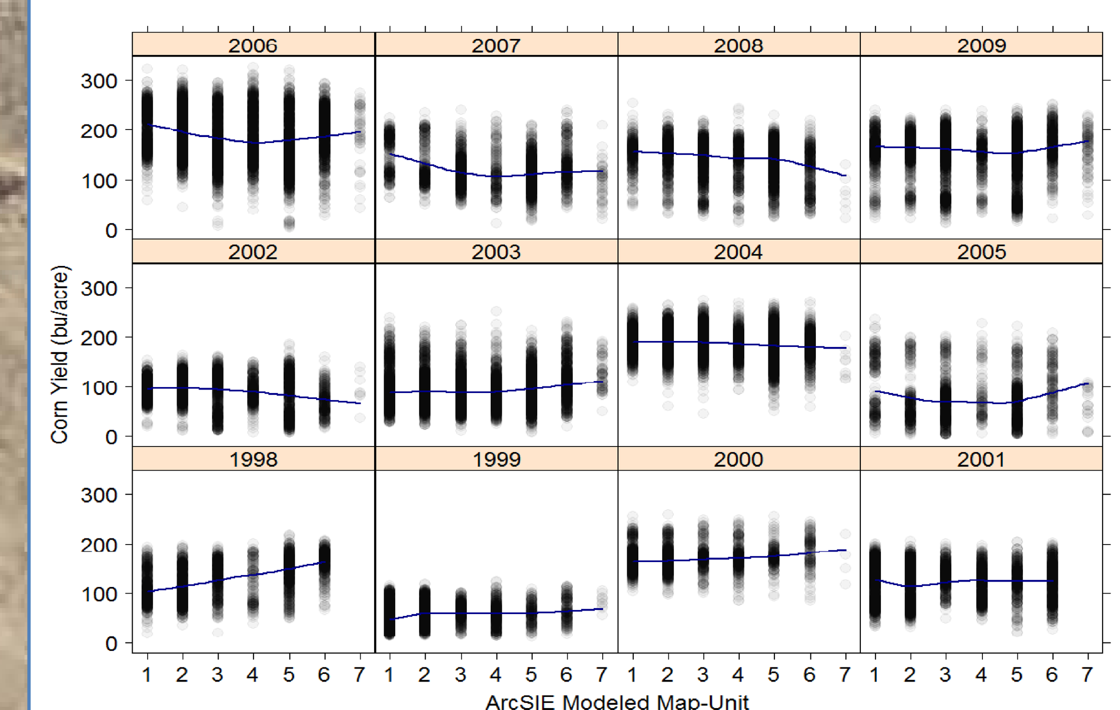
Productivity Index (PI) map from SSURGO, derived by assigning the Missouri PI to each SSURGO map unit. This is a categorical map.



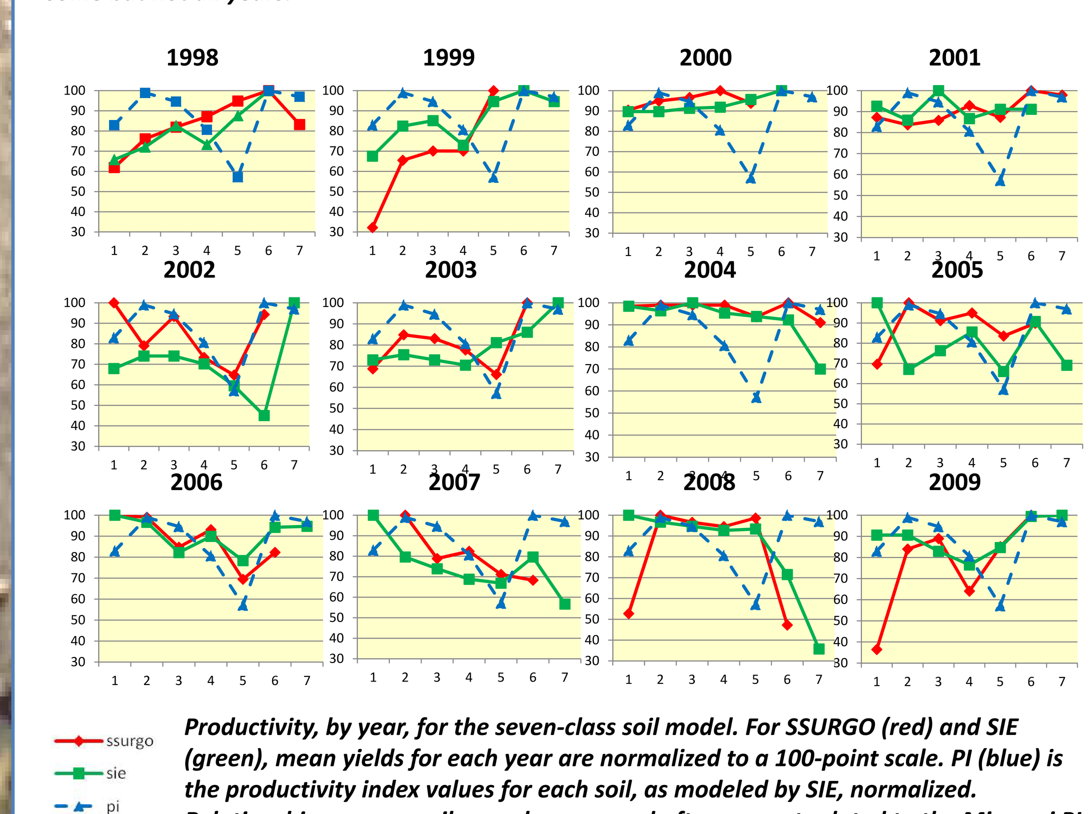
Productivity Index (PI) map from SIE, derived by assigning the weighted average PI to each pixel based on the degree of membership for each class. This is a continuous surface map.



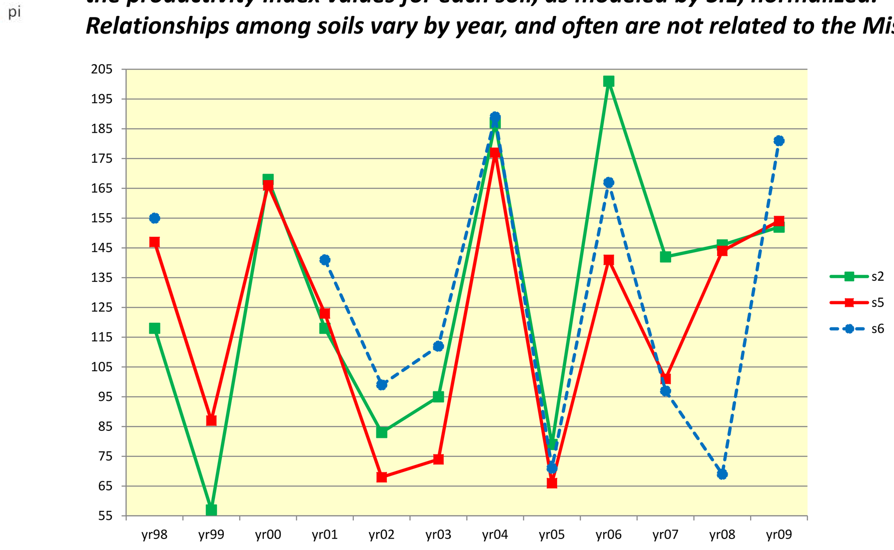
Regional Soil – Yield Relationships



Corn yield, by year, for the SIE model soils. Yield varies mainly by year, and is related to soil class in some but not all years.



Productivity, by year, for the seven-class soil model. For SSURGO (red) and SIE (green), mean yields for each year are normalized to a 100-point scale. PI (blue) is the productivity index values for each soil, as modeled by SIE, normalized. Relationships among soils vary by year, and often are not related to the Missouri PI.



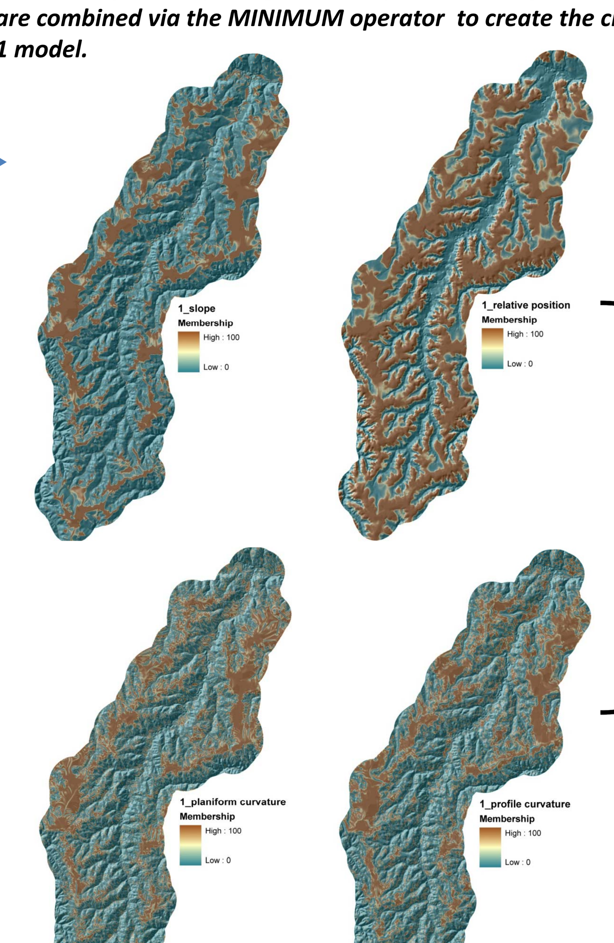
Mean yield by year for three contrasting SIE model soils (s2 = broad loess summits, s5 = till backslopes, s6 = stream terraces). Yield relationships among soils vary by year.

The SIE Model

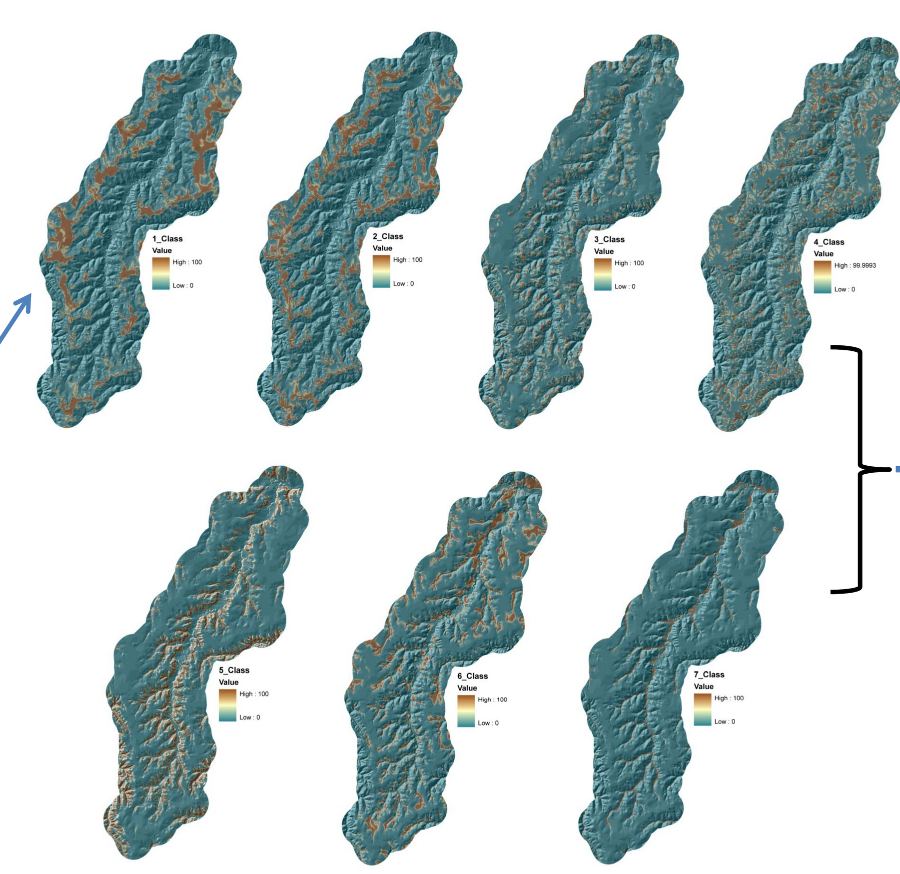
class	env_cov	curve	v1	v2	w1	w2
1	relpos	s	0.75		0.12	
1	slope	z		0.4		0.3
1	plan	bell	0	0	-0.005	0.005
1	prof	bell	0	0	-0.005	0.005
1	wetnes					
2	relpos	s	0.75		0.12	
2	slope	z		0.6		0.2
2	plan					
2	prof	bell	-0.004	-0.004	-0.005	0.005
2	wetnes	z		5		0.5
3	relpos	bell	0.65	0.65	0.15	0.15
3	slope	bell	1	1.5	0.5	0.5
3	plan					
3	prof					
3	wetnes	z		2		0.5
4	relpos	bell	0.6	0.6	0.2	0.2
4	slope	bell	1.5	1.5	0.5	0.5
4	plan	s	0		-0.001	
4	prof					
4	wetnes	s	5		1	
5	relpos	bell	0.5	0.5	0.2	0.2
5	slope	s	3		0.25	
5	plan					
5	prof					
5	wetnes					
6	relpos	bell	0.4	0.4	0.1	0.1
6	slope	z		2		0.5
6	plan					
6	prof					
6	wetnes					
7	relpos	z		0.15		0.1
7	slope	z		2		0.5
7	plan					
7	prof	s	0.02		0.05	
7	wetnes	s	8		0.5	

1) Rules for the five environmental covariates used in SIE to create the seven soil classes. Parameters specify the type of curve, the value(s) at which membership = 1, and the value(s) at which membership = 0.5.

2) Individual rule outputs for class 1 soil model, illustrating the rules of low slope, high relative position, and flat. These are combined via the MINIMUM operator to create the class 1 model.



3) Individual membership maps derived from the rules. Values are degree of membership.



4) Seven-class "hardened" model created from the individual membership maps, using the class with the highest membership value (MAXIMUM operator) at each pixel location.