

SAGT MRB-2A:

**Validation and regional application of
urbanization models in the South Atlantic-
Gulf & Tennessee River Basins**

Objectives:

- 1. Describe PCB conditions for urban sites in different environmental settings.**
- 2. Test and verify EUSE/ULUG response models (ATL, BIR, RAL).**
- 3. Place EUSE/ULUG models in regional context.**
- 4. Assess trends in urbanization over time and validity of substituting space for time.**

Background

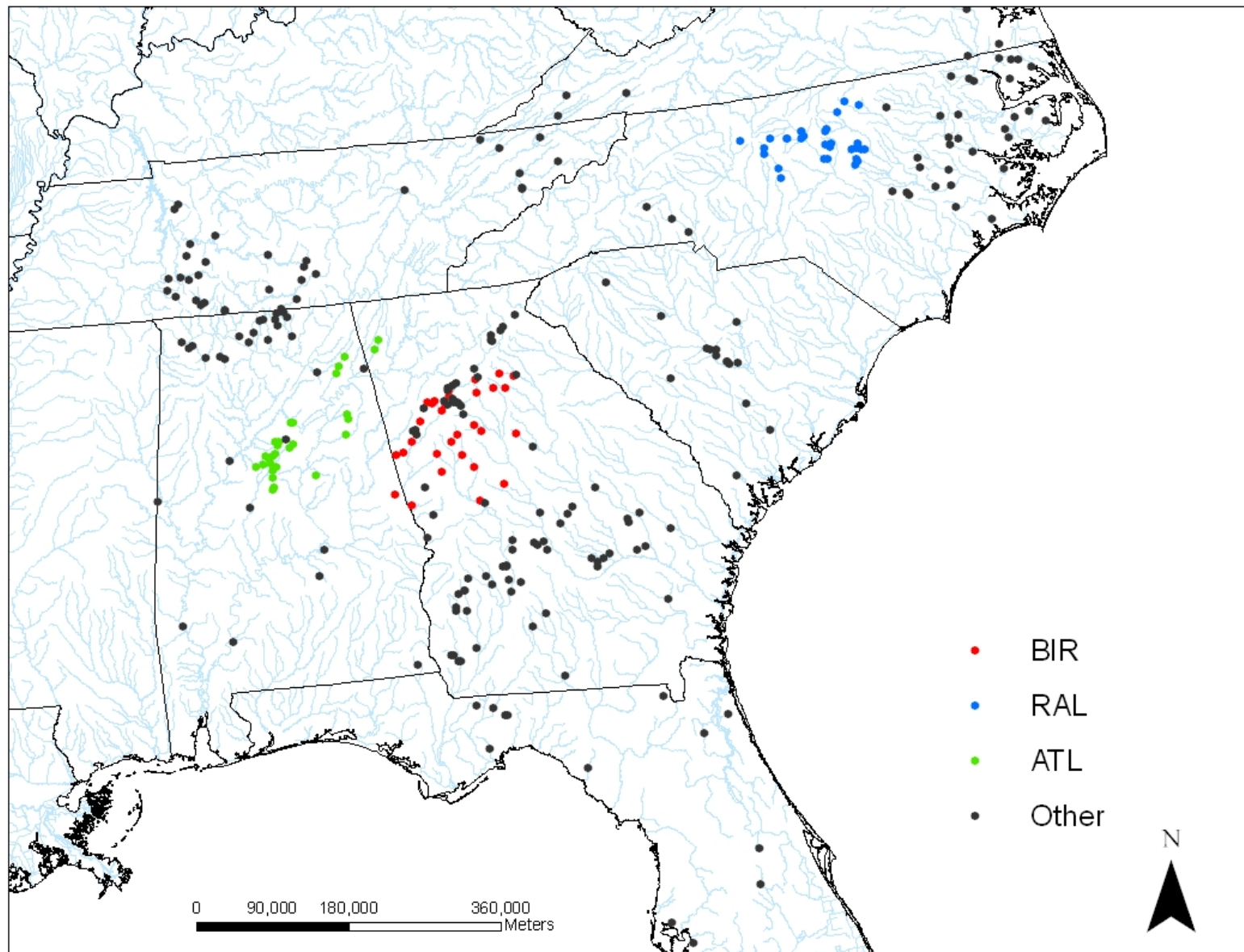
- **Three urban models in SAGT.**
 - **2 Piedmont (RAL riffle, ATL snag) , 1 Valley and Ridge (BIR riffle).**
- **How applicable are these urban models to other environmental settings?**
- **How well do these models predict changes over time (space for time)?**
- **Use Cycle I data to address these issues.**

Data scope

NAWQA Cycle I, ULUG, and EUSE sites with invertebrate data (ACFB, ALBE, GAFL, MOBL, SANT, TENN)

- 1. Invertebrate: RTH and/or QMH data**
- 2. Land-use data (basins and stream buffer)**
- 3. Population density**
- 4. Census data**
- 5. Infrastructure (road density)**

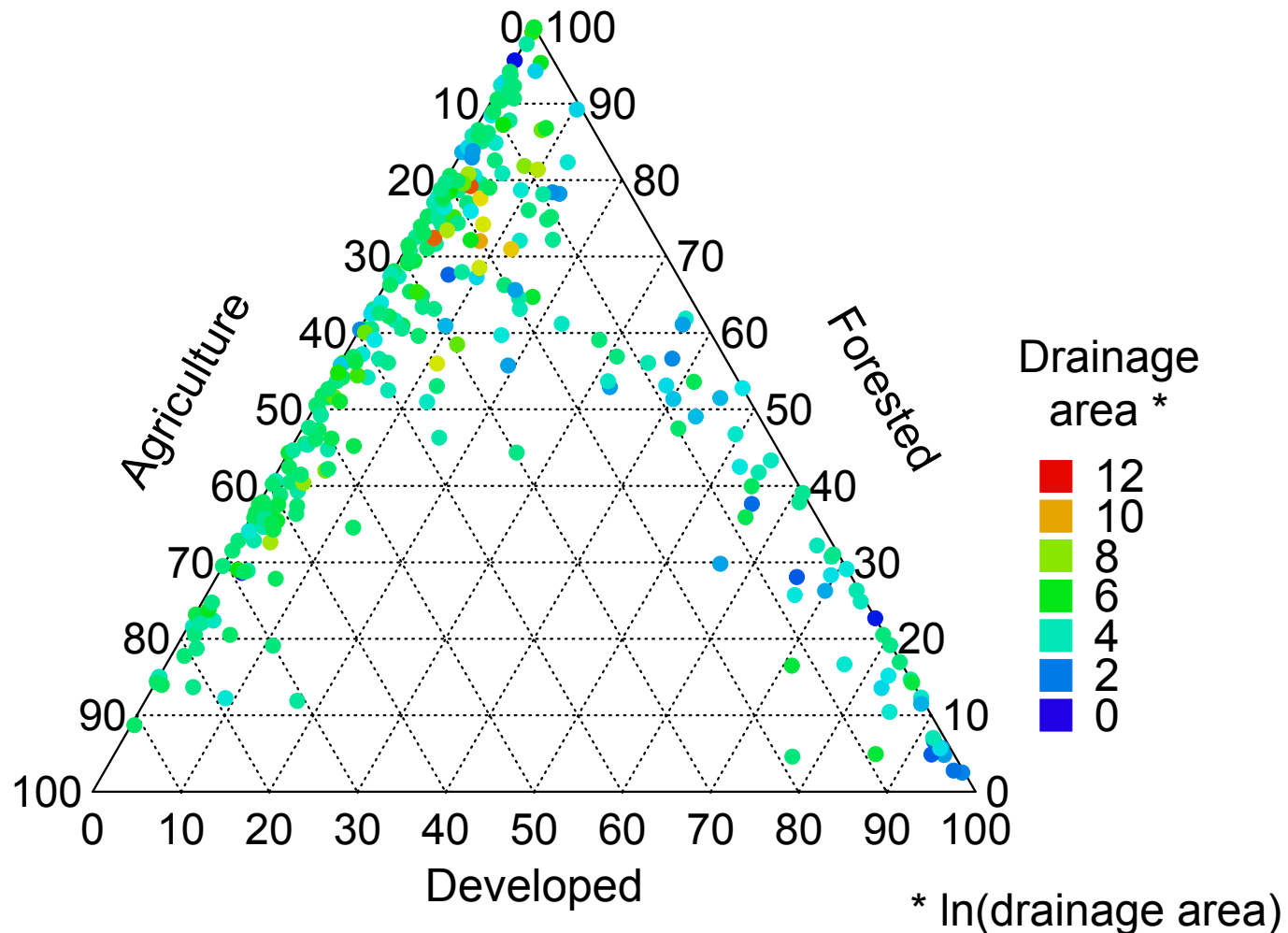
SAGT sites



No. of sites in SAGT urban MRB

SUID	Total	RTH	QMH
ACFB	95	67	69
ALBE	64	64	64
GAFL	25	13	25
MOBL	37	37	37
SANT	18	18	17
TENN	55	55	55
Total	294	254	267

SAGT MRB Land-use, Basin Area



No. of sites with multiple year data*

SUID	Total	RTH	QMH	No. Yrs
ACFB	95	15	32	2-9
ALBE	64	21	21	2-5
GAFL	25	11	11	2-6
MOBL	37	6	4	2-6
SANT	18	6	7	2-6
TENN	55	12	12	2-6
Total	294	71	87	2-9

*2 or more years of data from same reach

Data availability issues

- **Problems obtaining ancillary data:**
 - 2000 nlcd data (limited spatial availability).
 - 1990 census data (basin boundary problem from Geolytics).
- **Resolution to availability issues:**
 - Given up on 2000 nlcd data: using enhanced 1992 data.
 - Sylvia currently processing revised 1990 census data from Geolytics.

Staffing issues

- **Brian Gregory withdrew.**
- **Elise Giddings covering EUSE to free up Cuffney's time for MRB** (back for interpretation & writing, multiple year work).
- **Robin Brightbill**
 - **Replace Gregory**
 - **Clean-up data sets (mix of '90 and '00 data).**
 - **Revise models with cleaned-up data.**

Issues with objectives:

1. **Insufficient physical (e.g., flow, habitat) and chemical data to allow description of urbanization beyond the EUSE/ULUG models.**
2. **Probably(?) insufficient urban coverage to adequately address the concept of substituting space for time.**

Status of data analysis

1. Invertebrate data

- a. Data prep (IDAS): completed
- b. Metrics: completed
- c. CANOCO/Primer data sets completed

2. GIS data

- a. Census data: most 2000, little 1990
- b. NLCD data: 1990, 2000e
- c. Infrastructure: most 2000, 1990
- d. Envir. setting: completed (ecoregion and stream size)

Status of data analysis (cont'd)

3. Derivation of SAGT MRB UII (common UII)
 - a. Subset of UII variables (7) selected
 - b. SAGT MRB compared with EUSE/ULUG UII
 - i. 28-30 site standardization (EUSE/ULUG models)
 - ii. 294 site standardization
 - c. SAGT MRB 294 UII calculated for all sites*
4. Invertebrate response models
 - a. RTH, QMH, QQ Ordination (CA) models completed (RB)
 - b. Multiple regression models completed*
 - c. Model performance (slopes and R^2) compared for urban model data sets (ATL, RAL, BIR)*

*Preliminary calculation with mixed year data

Status of data analysis (cont'd)

4. Multiple year studies: pending
5. Presentations:
 - a. AGU: December '05, San Francisco
 - b. NWQMC: May '06, San Jose

Preliminary results

Derivation of a common Ull for the SAGT

**ACFB, ALBE, GAFL, MOBL, SANT,
TENN**

Ull varies by urban area

- Number of variables in Ull:

ACFB (ATL): 19

ALBE (RAL): 30 (7)

MOBL (BIR): 14

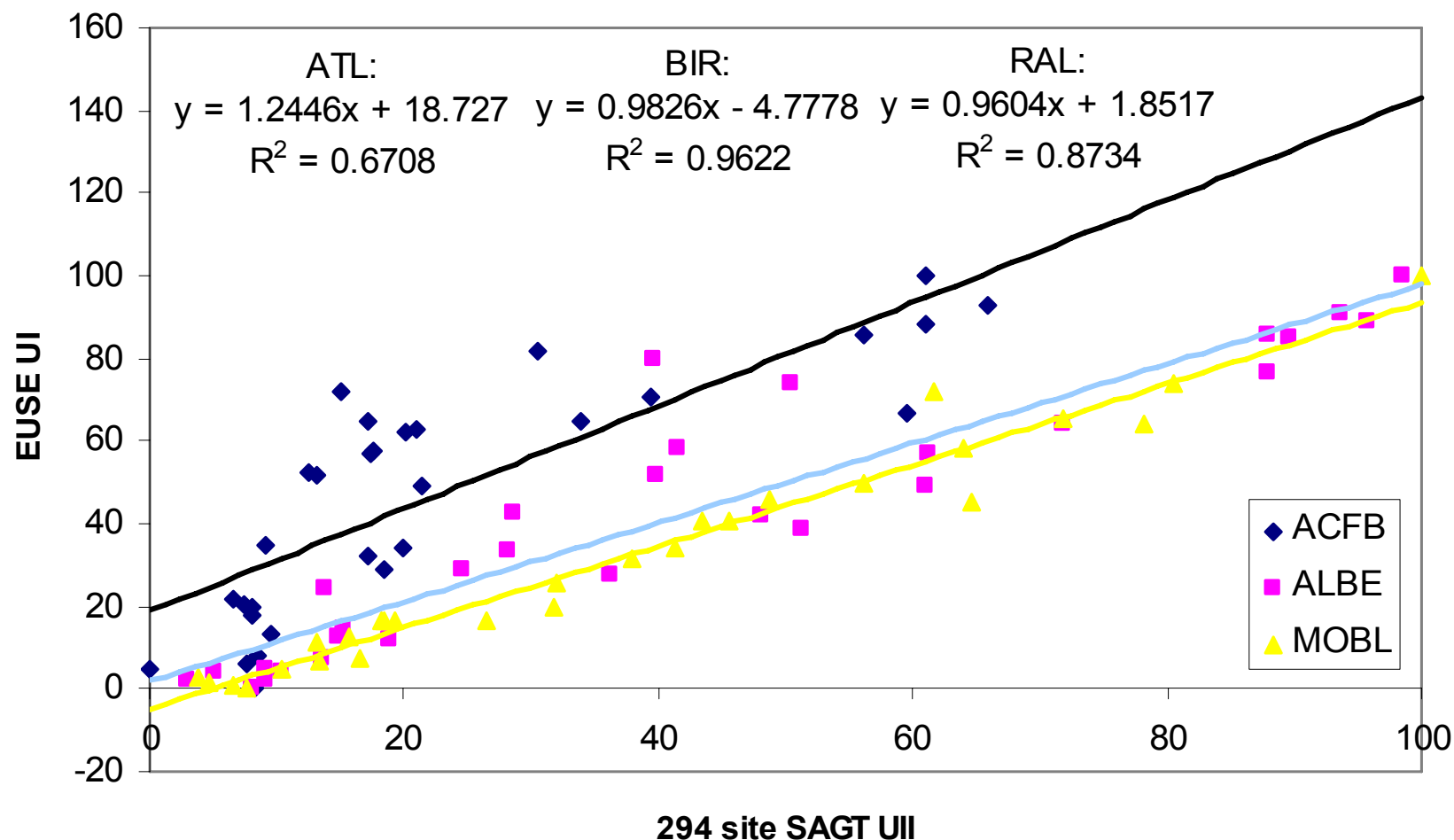
Need to derive a UII that is common to each urban area and to the 294 SAGT sampling sites.

Searched for a small combination of census, NLCD, and infrastructure variables that gave the same relationship among sites as the original UII (EUSE, ULUG)

Variables (7) in SAGT MRB UII

- **Census**
 - Population density
 - Household density
- **NLCD**
 - Basin data
 - % basin in developed land
 - % basin in forested land
 - Riparian buffers
 - % riparian buffer in developed land
 - % riparian buffer in forested land
- **Infrastructure**
 - Road density

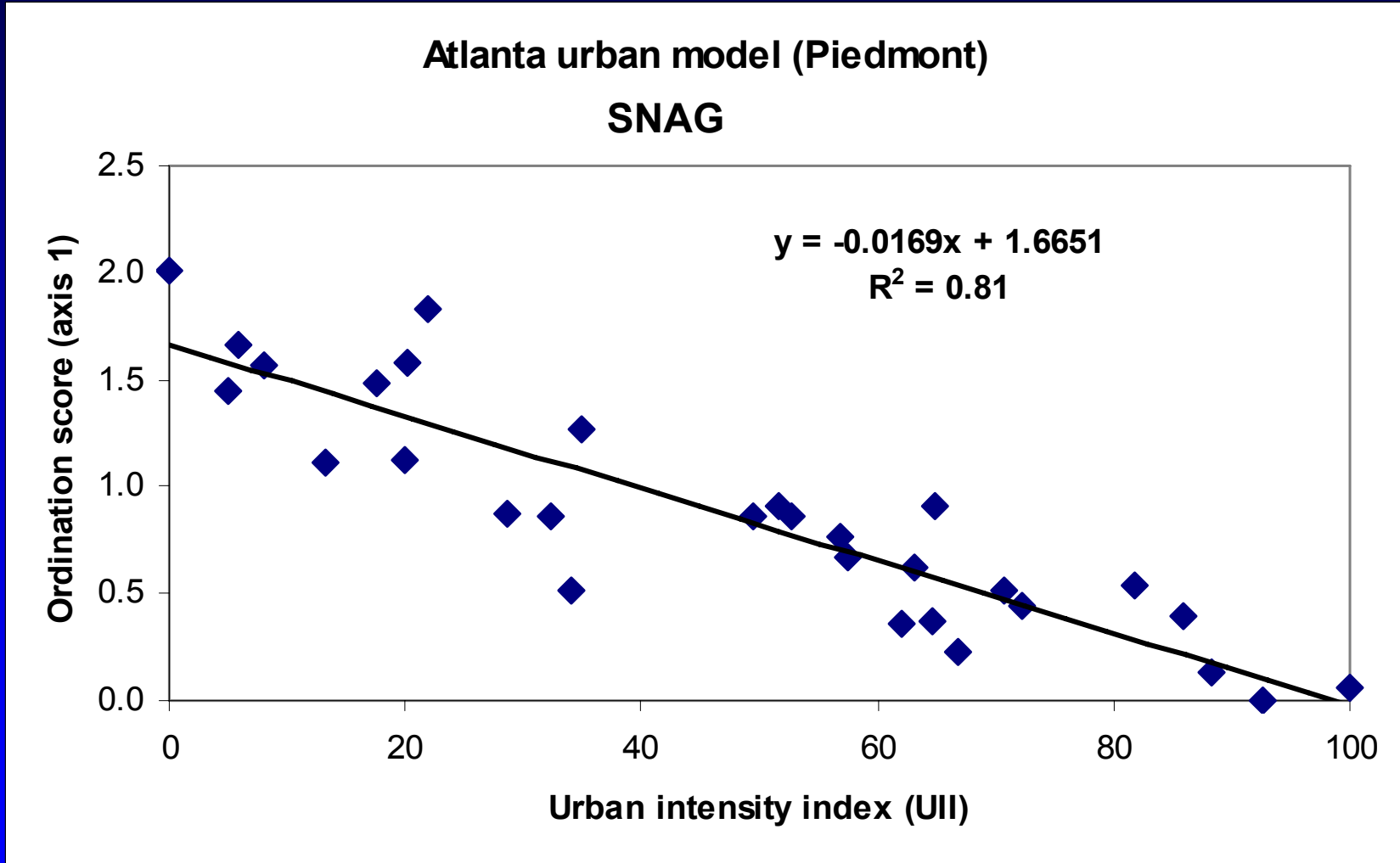
SAGT UII vs. EUSE



Urban models

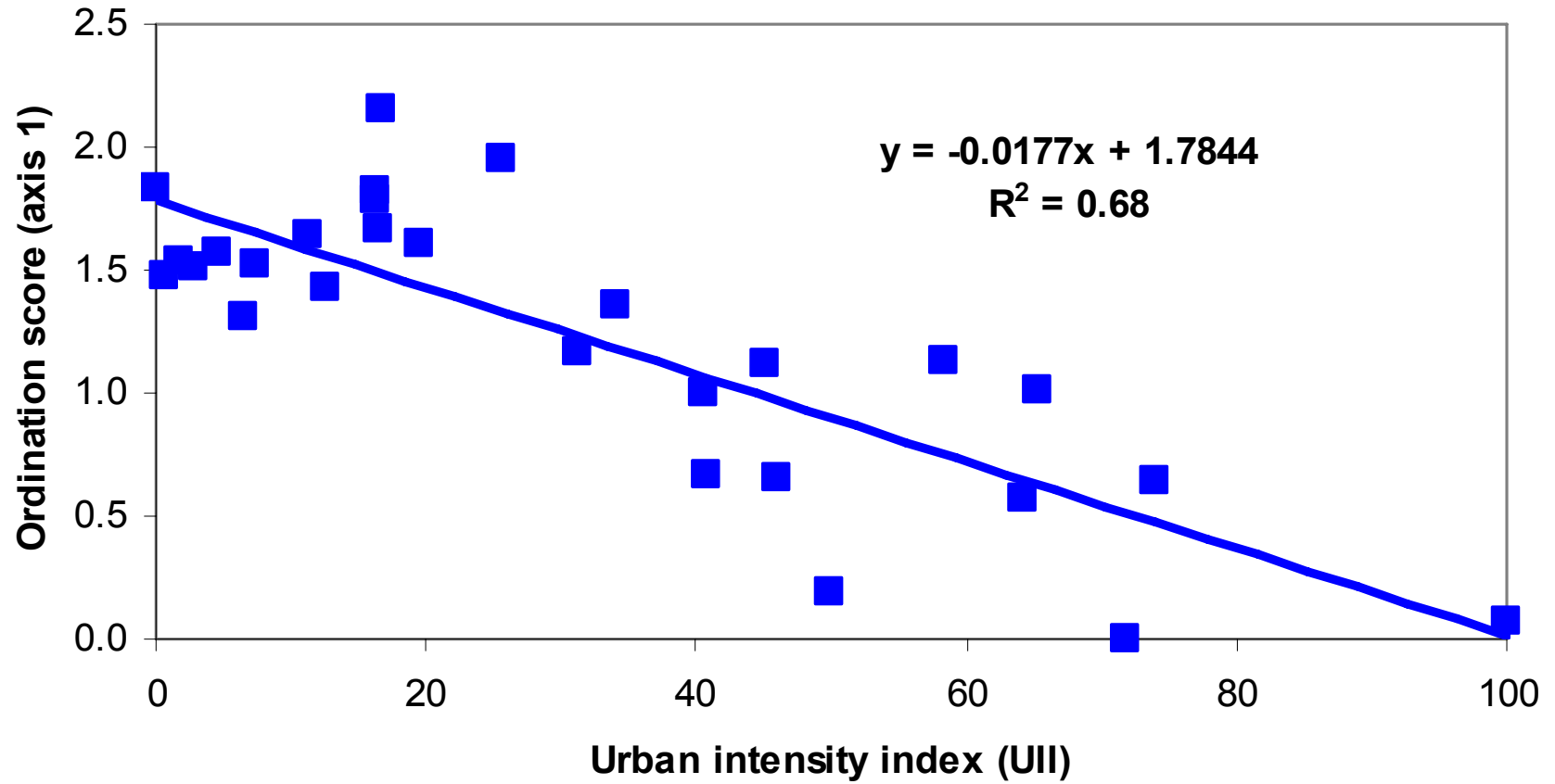
ATL, BIR, and RAL

RTH Inverts: EUSE model



RTH Inverts: EUSE model

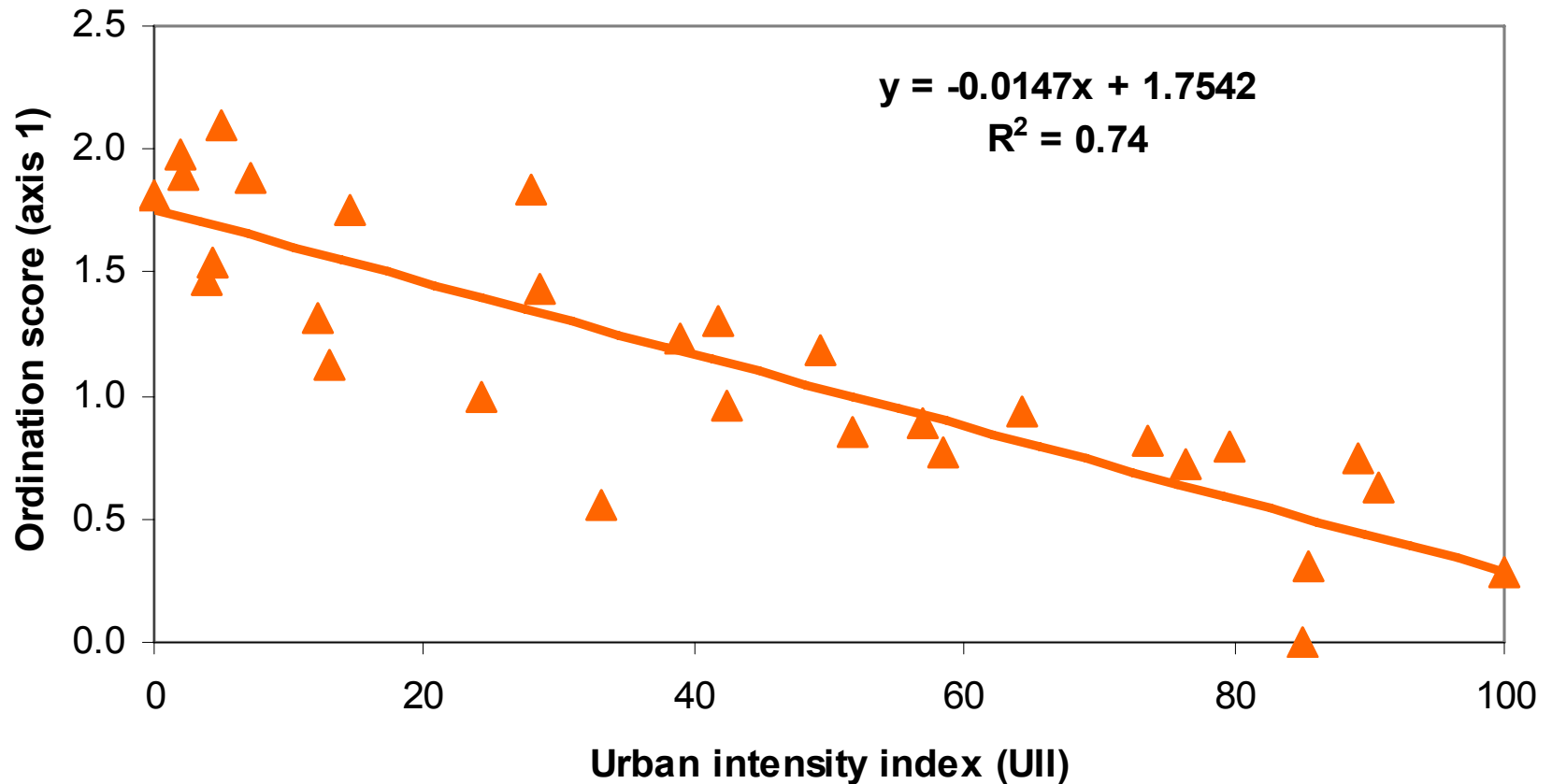
Birmingham urban model (Ridges & Valleys)
RIFFLE



RTH Inverts: EUSE model

Raleigh urban model (Piedmont)

RIFFLE



EUSE models: Site Characteristics

Abbrev	Core urban area	Dominant ecoregion	RTH habitat	Basin area (km ²)
ATL	Atlanta, GA	Piedmont	Woody snags	43-146
RAL	Raleigh, NC	Piedmont	Riffles	5-82
BIR	Birmingham, AL	Central Appalachian Ridges and Valleys	Riffles	5-66

Characteristics of sites not used to construct urban models

Dominant ecoregion	Small basins (< 160 km ²)		Large basins (≥ 160 km ²)	
	Sites	Ull range	Sites	Ull range
Piedmont	8	3-80	8	3-83
Mid Atlantic Coastal Plain	14	12-33	3	23-32
Southeastern Plains	48	4-81	21	11-51
Blue Ridge Mountains	0		3	9-13
Central Appalachian Ridges and Valleys	1	11	7	9-23
Southwestern Appalachians	2	9-16	4	3-12
Interior Plateau	35	3-32	11	8-66

Derivation of response model:

1. **Ordination model**
2. **Multiple regression model**

Predicting responses: EUSE

Assemblage ordination models

Ordination “model”

$$\text{Site score} = f(\text{taxa})$$



**Measure of dissimilarity
among assemblages**

**Know that the axis 1 CA site score is strongly
related to urban intensity (UII).**

Therefore, the site score is indicative of urban effect.

Passive samples

- Added to existing ordination by projection: regressing data on to the existing ordination axes.
- Calculated from the eigenvector scores:

passive sample score $\rightarrow x_i^* = \lambda^{\alpha-1} \sum_k w_k^* y_{ik} \mu_k / \underbrace{\sum_k w_k^* y_{ik}}_{\text{scaling factor}}$

Eigenvalue of axis $\rightarrow \lambda^{\alpha-1}$

weight assigned to taxon k $\rightarrow w_k^*$

abundance of taxon k in sample i $\rightarrow y_{ik}$

score for taxon K $\rightarrow \mu_k$

axis scaling (0, 1/2, 1) $\rightarrow \alpha$

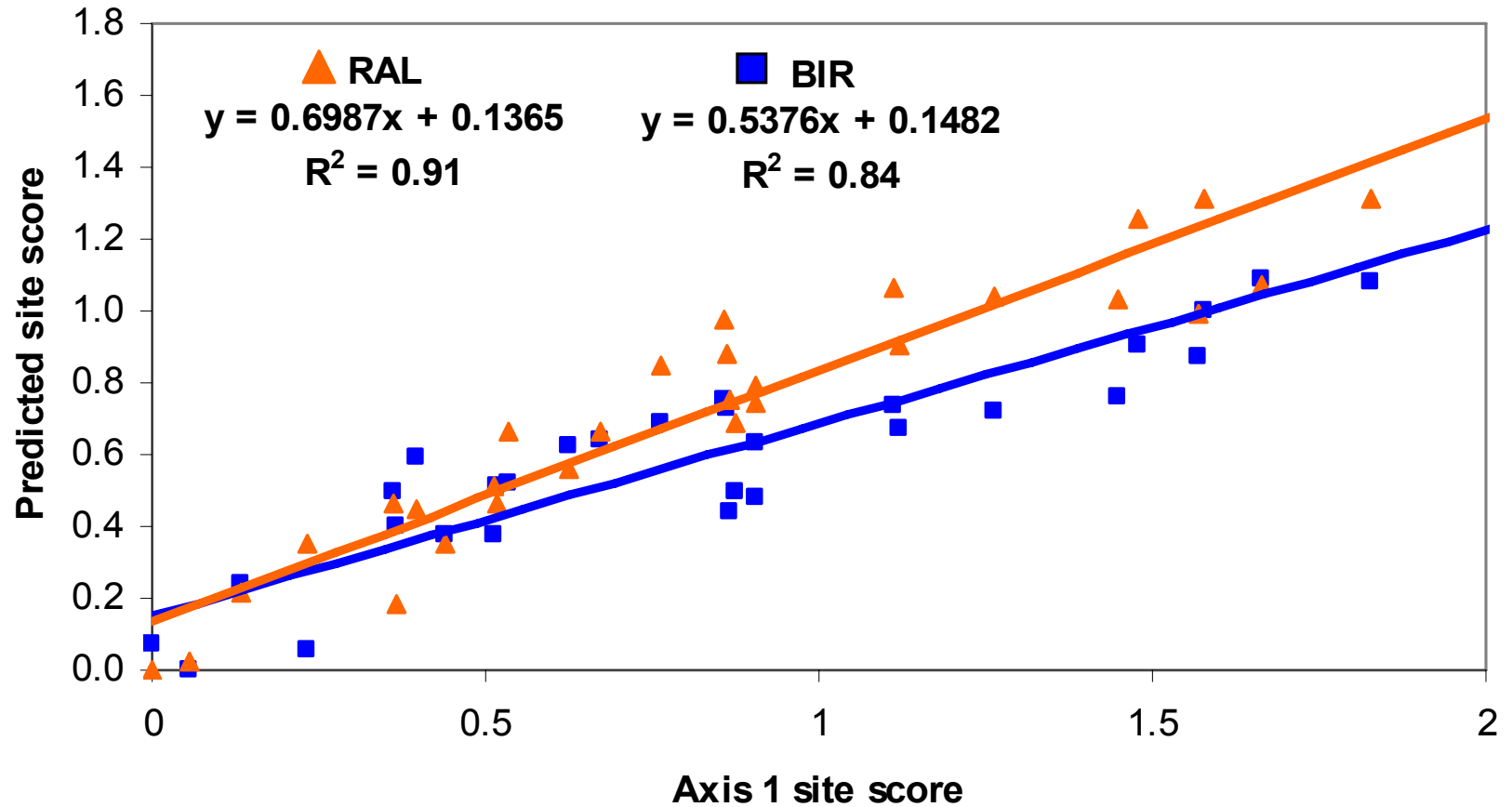
Ordinations only use the taxa that are part of the model data set

RAL	BIR model	RAL modeled using BIR
A		
B	B	B
C		
	D	
E	E	E
F		
G		
	H	

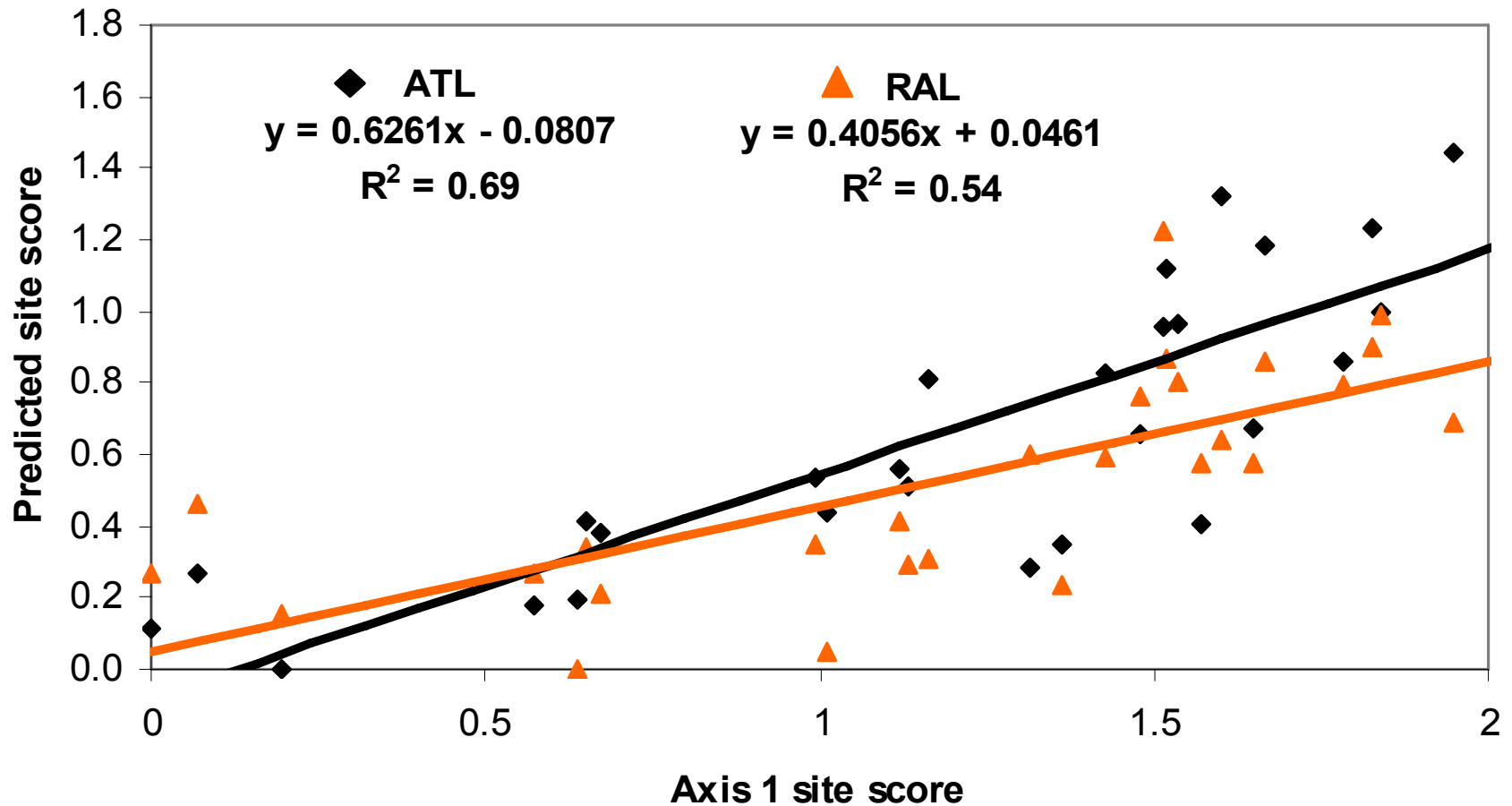
No. taxa used in ordination models

Urban Model	ATL		BIR		RAL	
	Common	Unique	Common	Unique	Common	Unique
ATL	155	0	69	58	83	47
BIR	69	86	127	0	69	61
RAL	83	72	69	58	130	0

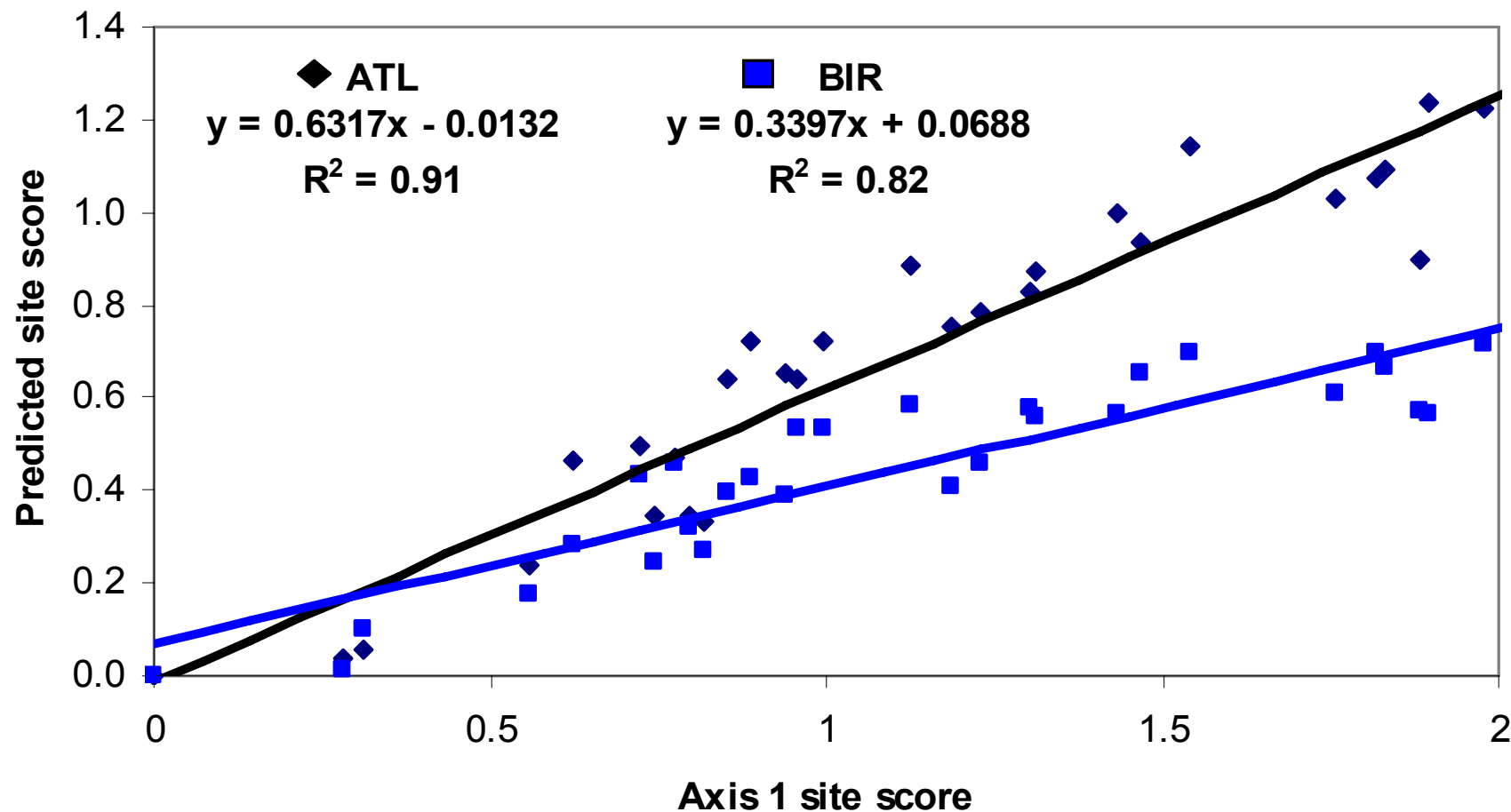
Atlanta predicted by Birmingham and Raleigh models RTH



Birmingham predicted by Atlanta and Raleigh models RTH



Raleigh predicted by Atlanta and Birmingham models RTH



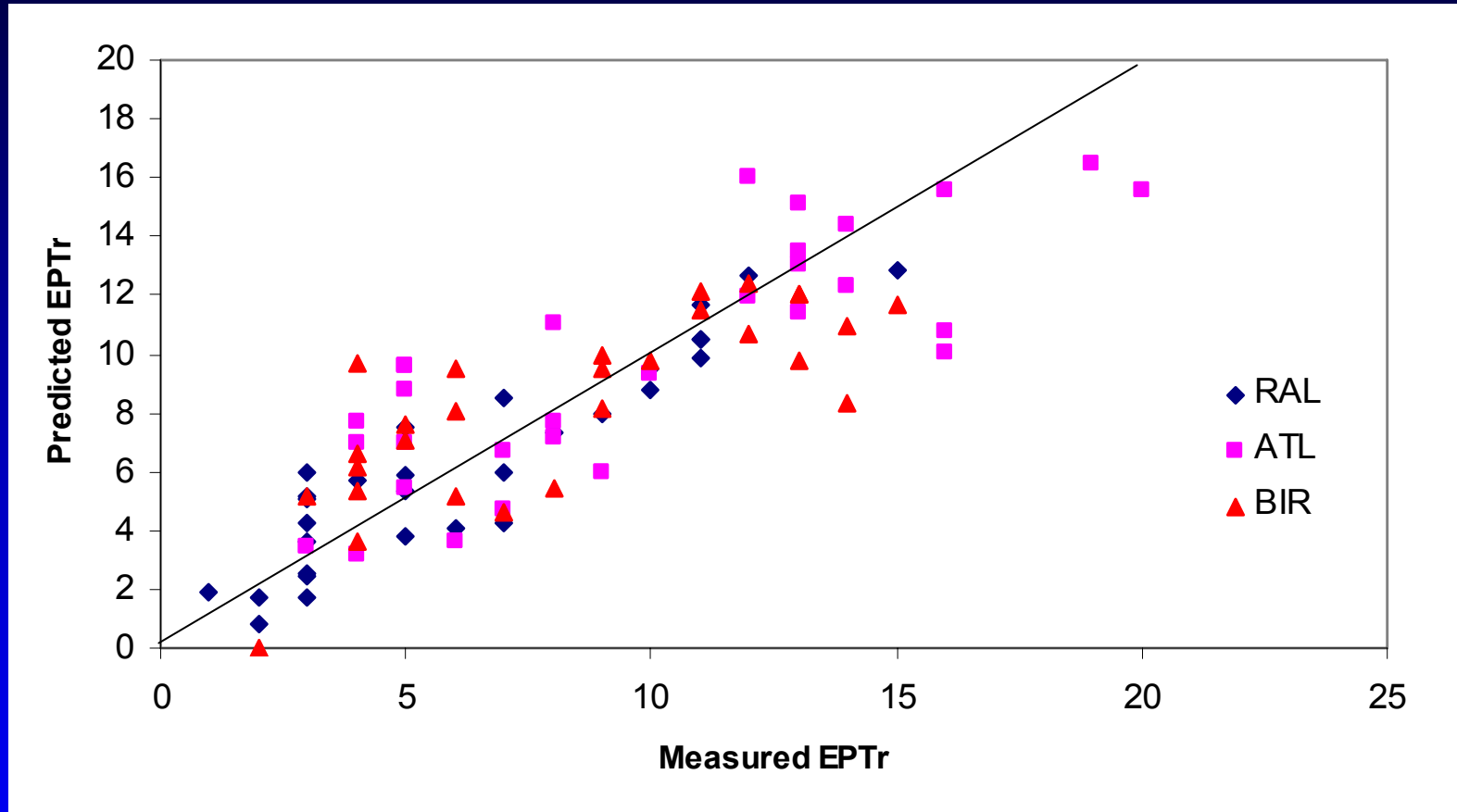
Multiple Regression Models: ATL, BIR, and RAL

- $EPT_r = f(X_i)$
- X_i :
 - Land-use: forest (4), developed (2), ag/urban grass (8)
 - % basin area
 - % stream buffer
 - Population density
 - Road density

EPT_r EUSE models

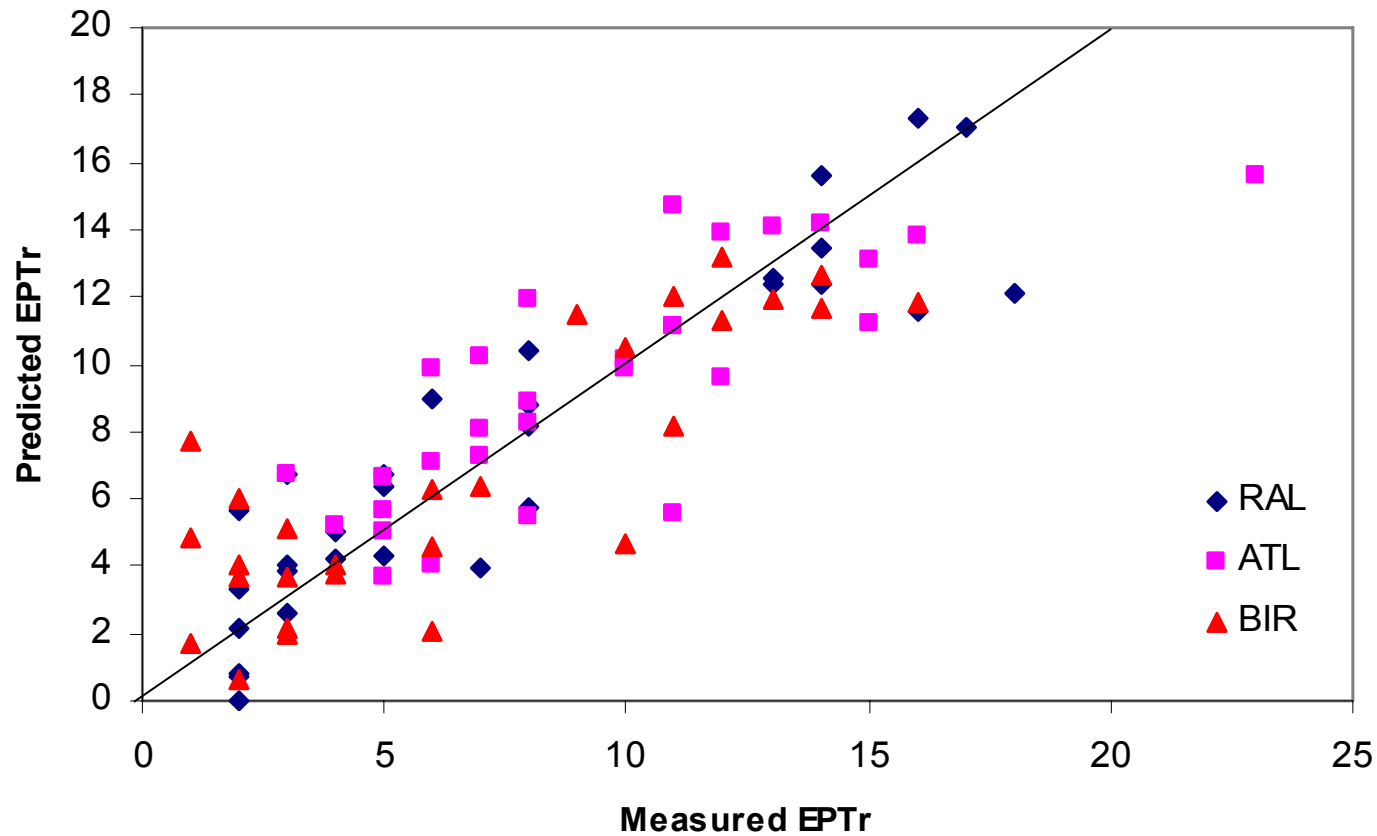
RTH	% basin area				% stream buffer			PopD	Road	R²	P
	Const	Devel	Forest	Grass	Devel	Forest	Grass				
RAL	-0.04		0.11		0.13		0.47	0.01	-1.44	0.81	<0.001
ATL	-8.09		0.27	0.33	0.45				-3.19	0.62	<0.001
BIR	13.75								-1.13	0.59	<0.001
QMH											
RAL	-25.24		0.37	0.17	0.27		0.70			0.81	<0.001
ATL	3.88			0.26	0.27	0.10			-2.76	0.79	<0.001
BIR	-41.27	1.22	1.25	0.84	-0.83	-0.73	-0.40			0.62	<0.001
QQ											
RAL	-11.33		0.26		0.21		0.82		-0.70	0.85	<0.001
ATL	15.95			0.23	0.23				-3.53	0.58	<0.001
BIR	15.86				0.19				-2.15	0.03	0.25

RTH EPTr: EUSE model



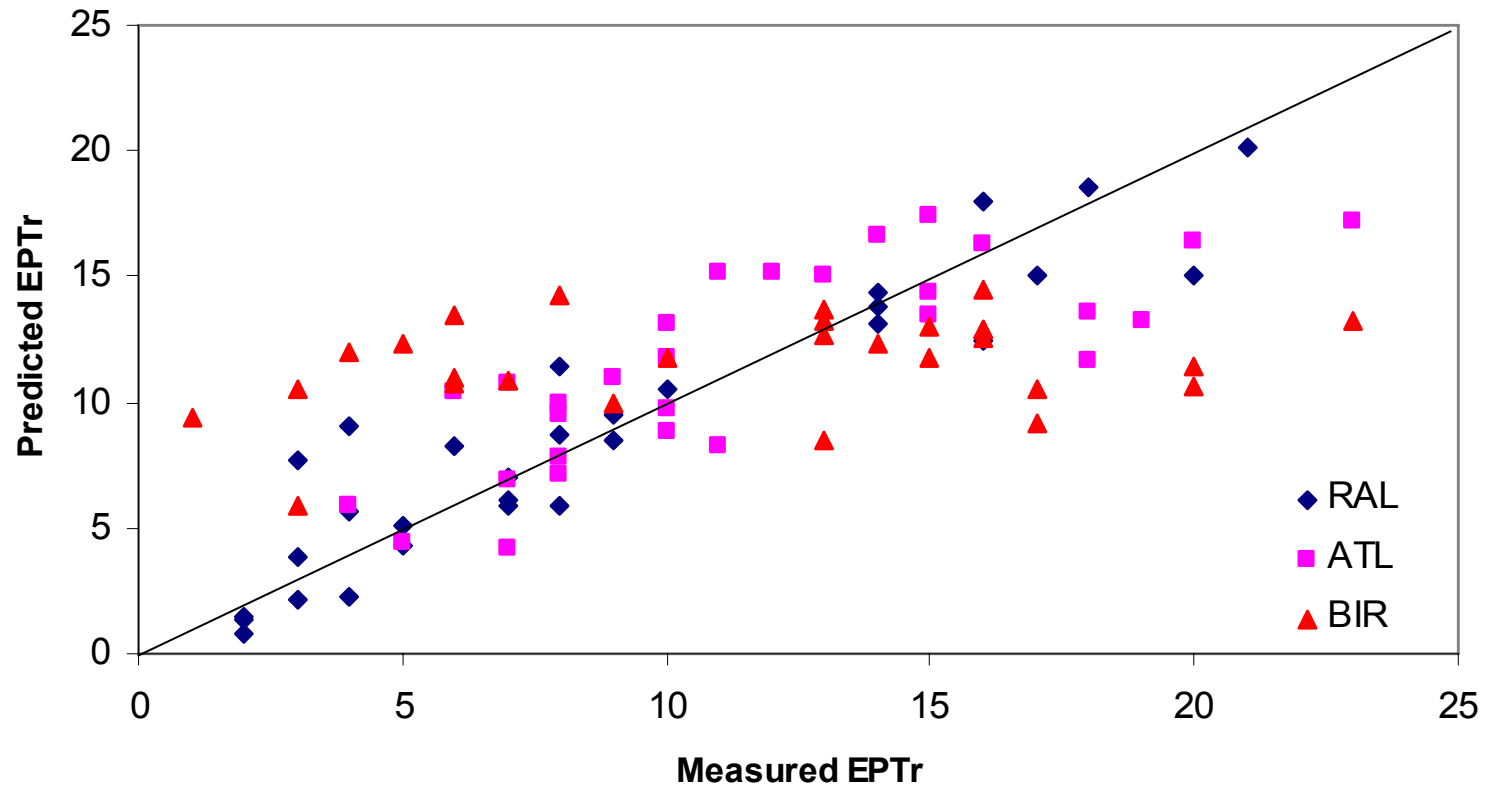
	RAL	BIR	ATL
a	0.953	3.255	3.247
b	0.848	0.609	0.669
R ²	0.85	0.61	0.67

QMH EPTr: EUSE model



	RAL	BIR	ATL
a	1.233	2.000	3.465
b	0.837	0.702	0.631
R ²	0.84	0.70	0.63

QQ EPTr: EUSE model



	RAL	BIR	ATL
a	1.178	10.392	4.240
b	0.867	0.105	0.627
R ²	0.87	0.10	0.63

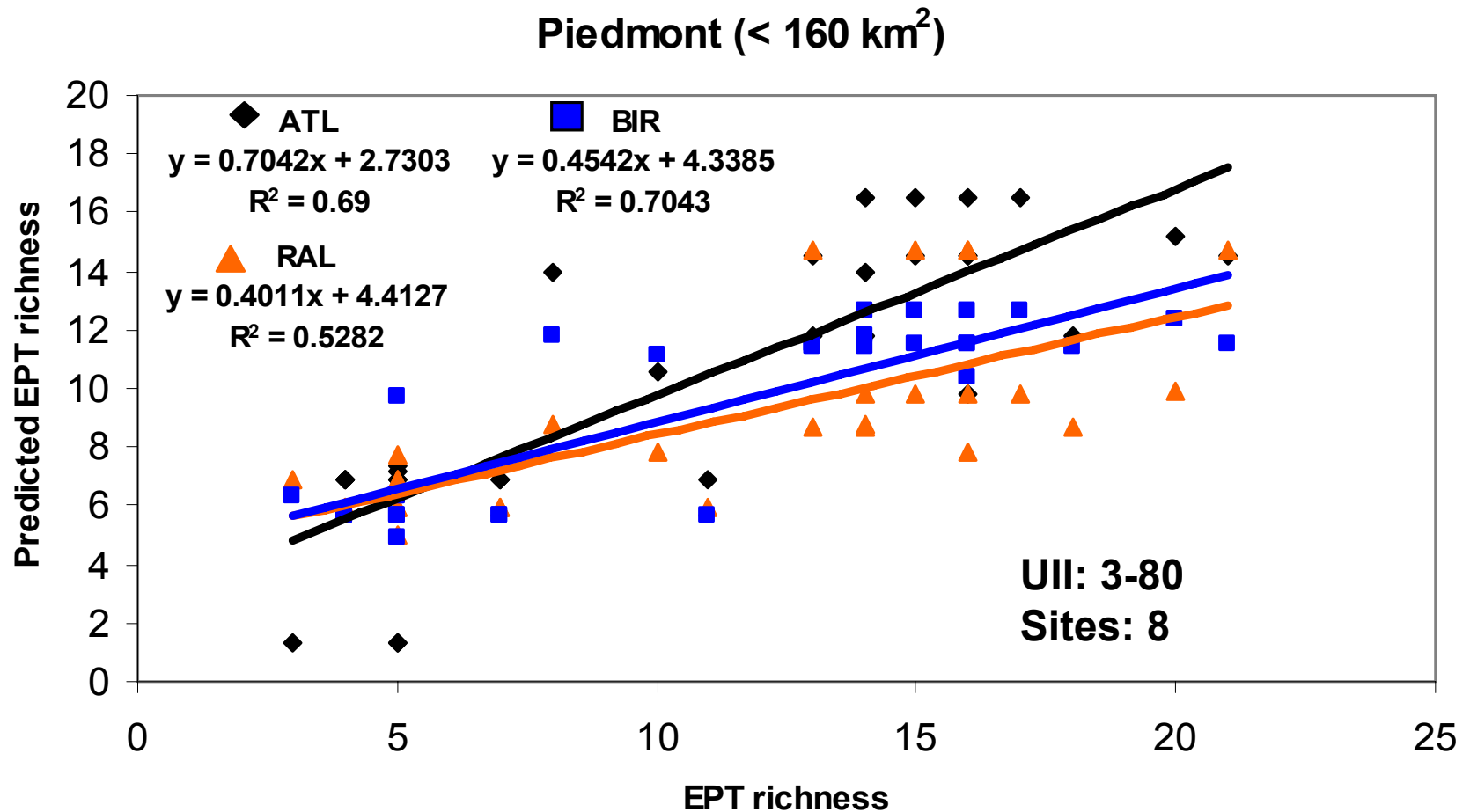
Modeling at non-EUSE sites

EPTTr multiple regression models

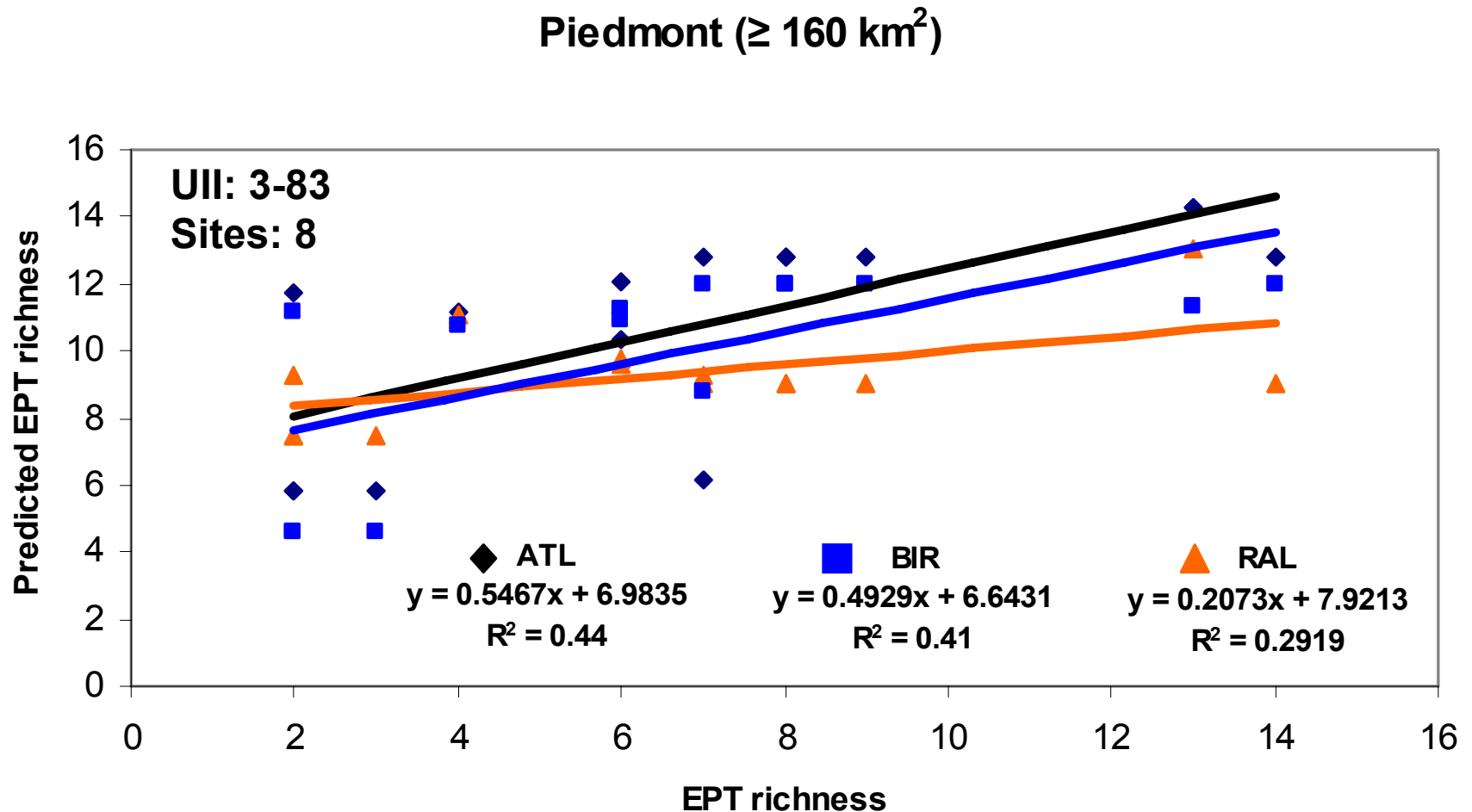
Multiple Regression Models

- $EPT_r = f(X_i)$
- X_i :
 - Land-use: forest (4), developed (2), ag/urban grass (8)
 - % basin area
 - % stream buffer
 - Population density
 - Road density

EPT_r: small streams, RTH

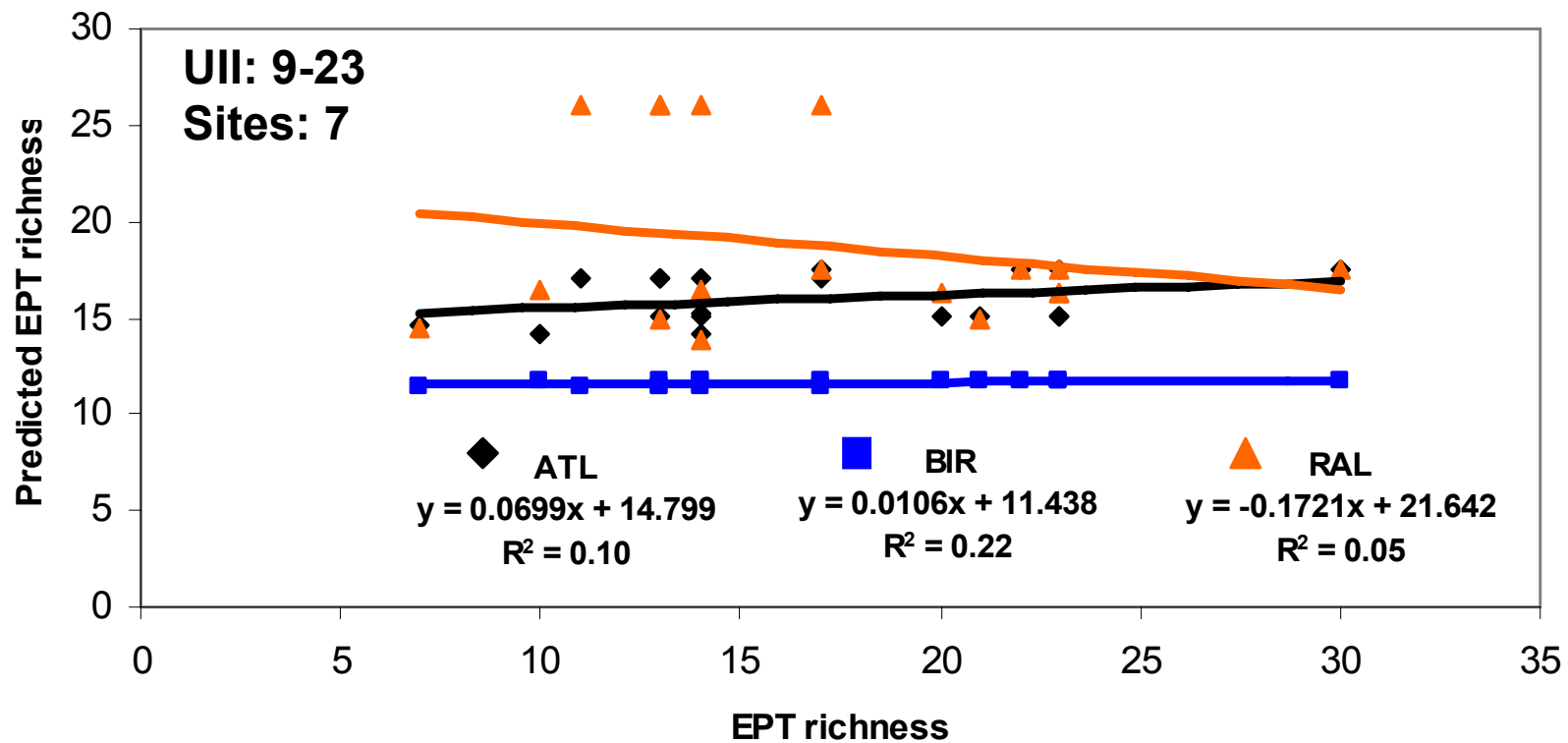


EPT_r: Large streams, RTH

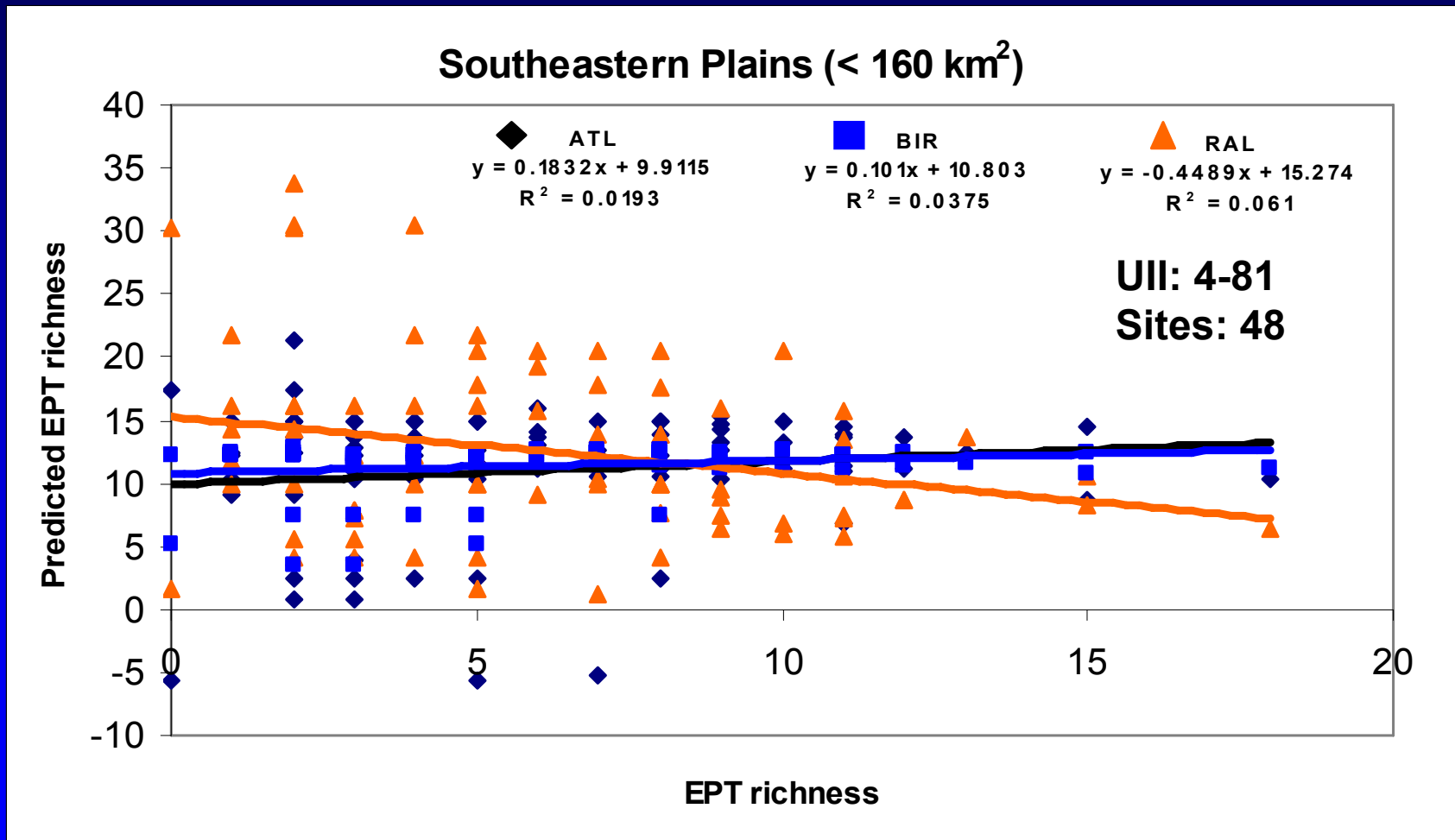


EPTr: Large streams, RTH

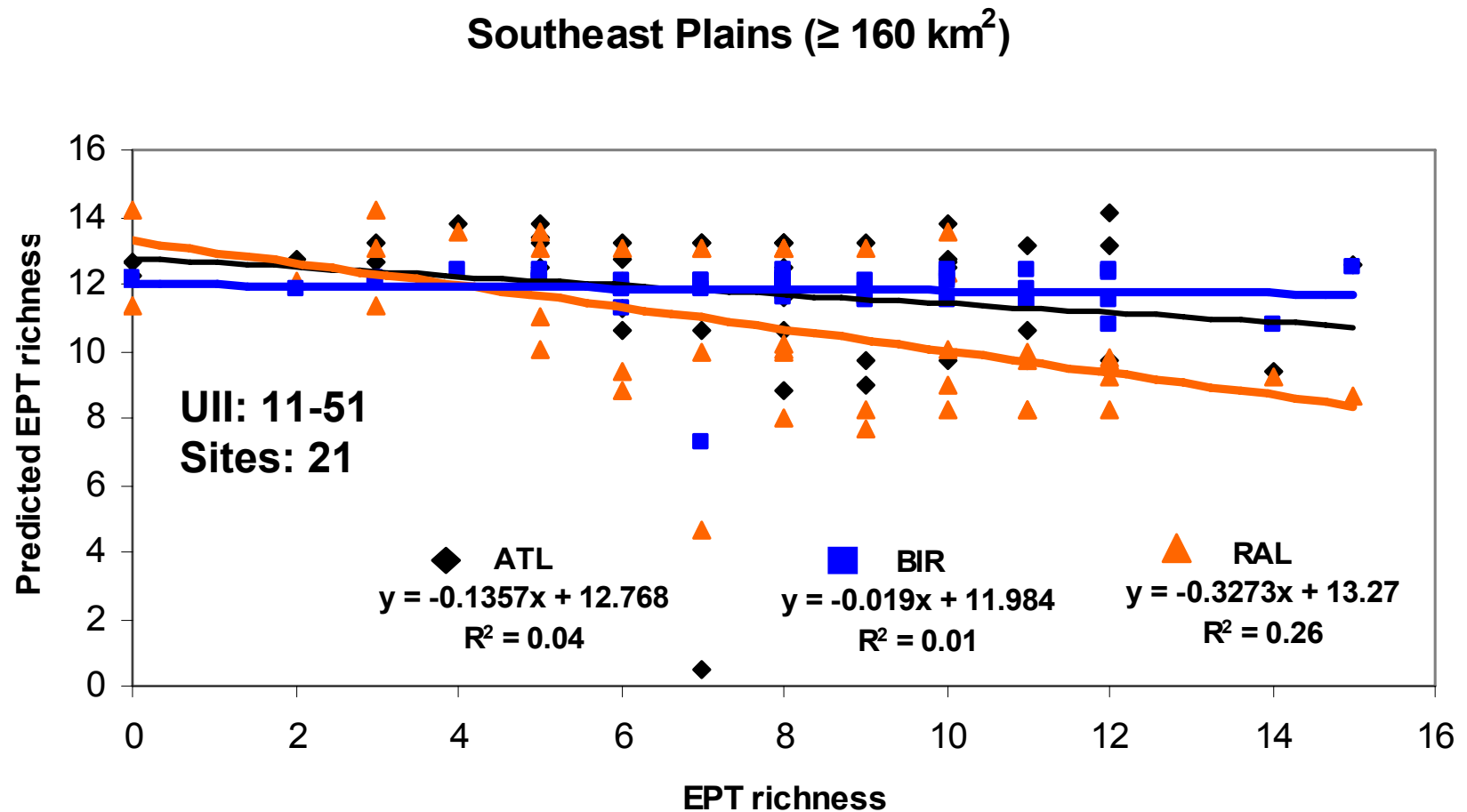
Ridges and Valleys ($\geq 160 \text{ km}^2$)



EPTTr: small streams, RTH

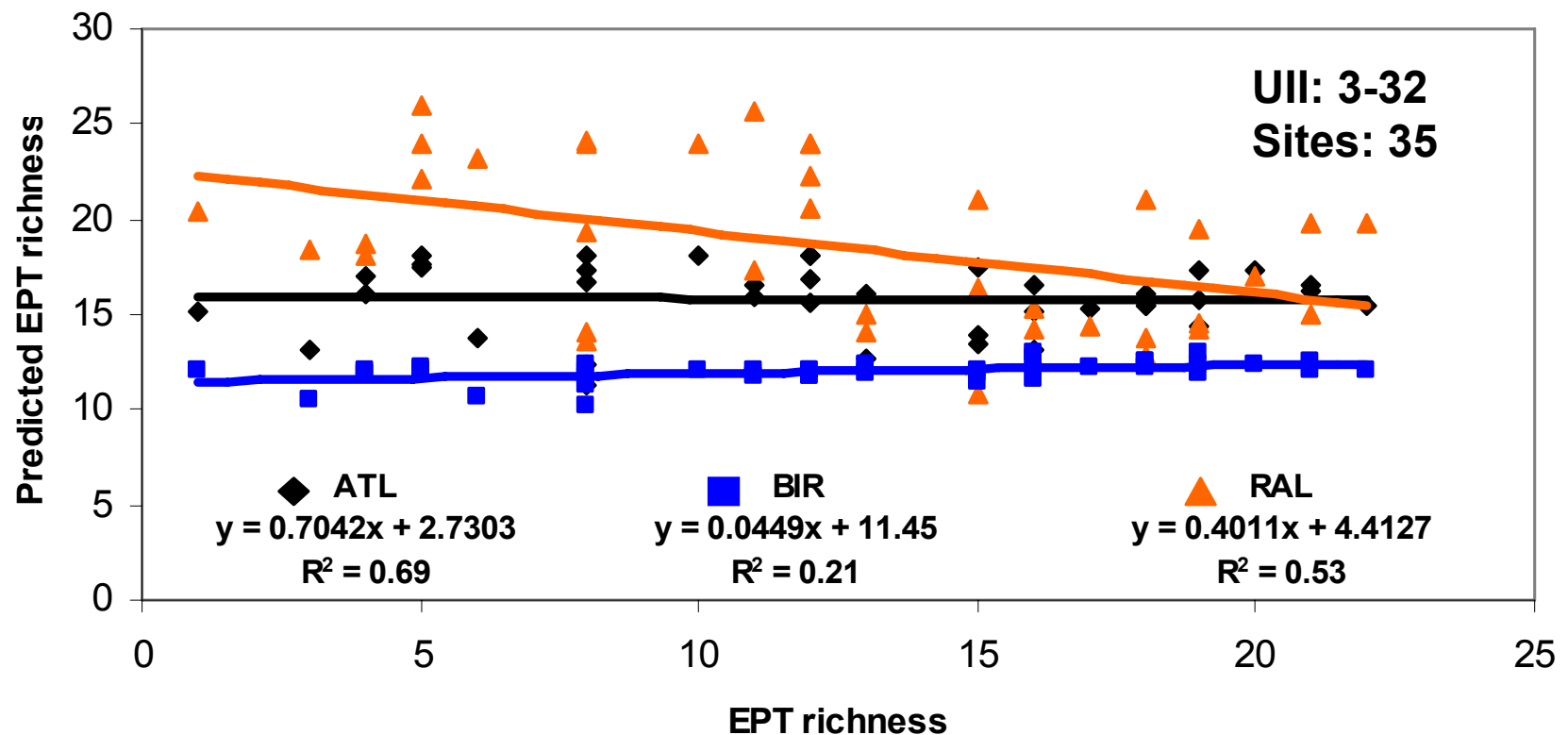


EPTr: larger streams, RTH



EPT_r: small streams, RTH

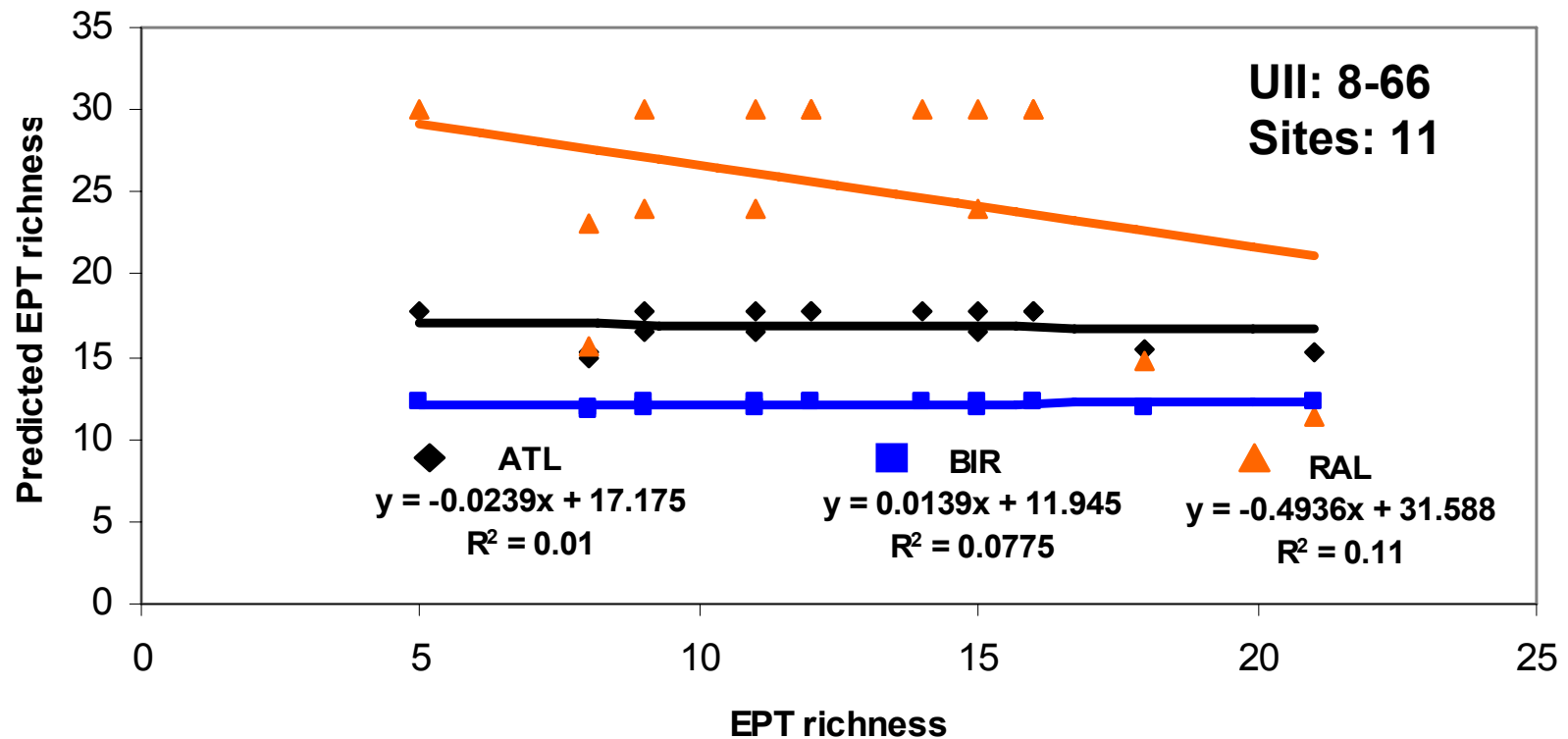
Interior Plateau (< 160 km²)



EPTTr: large streams, RTH

Interior Plateau ($\geq 160 \text{ km}^2$)

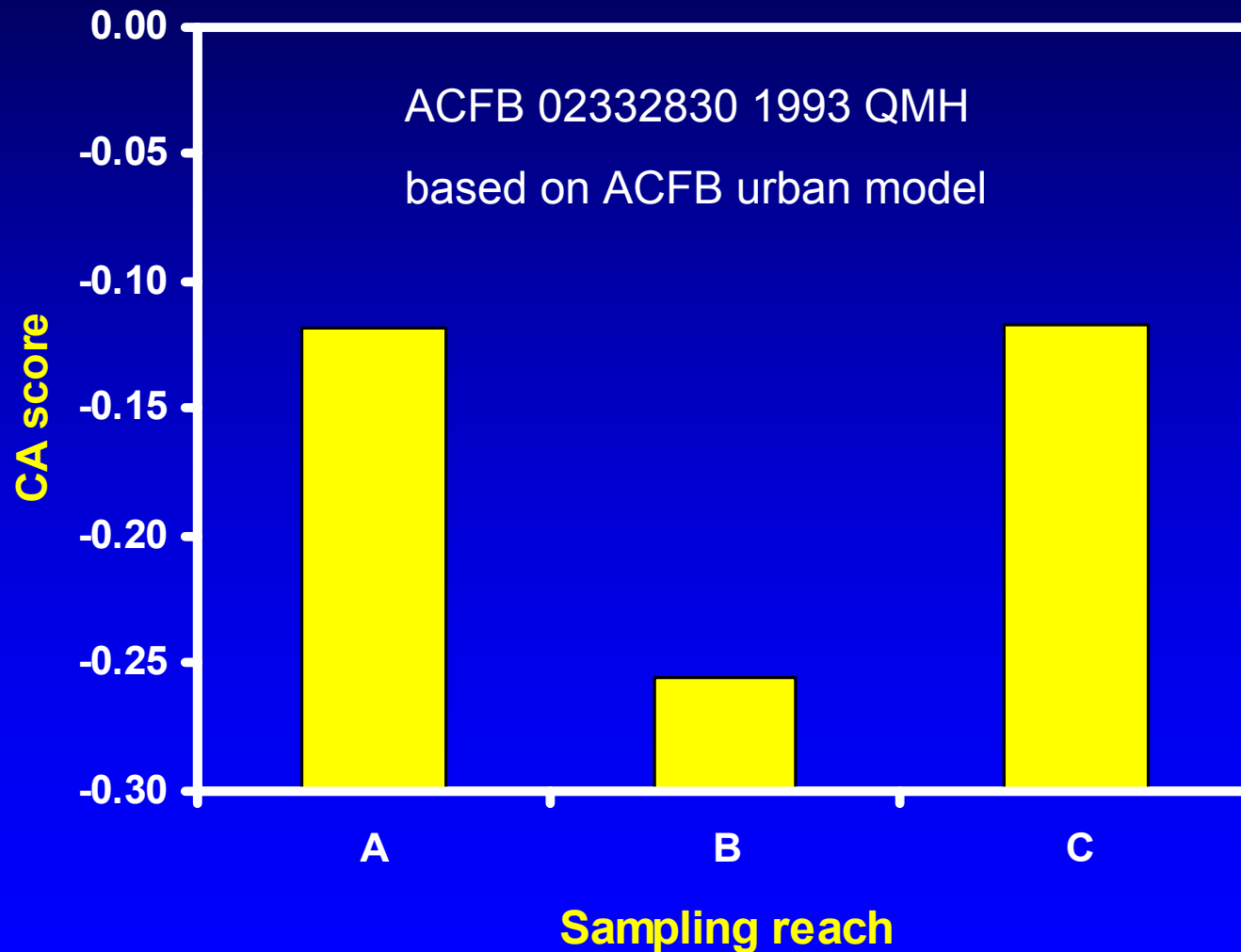
Ull: 8-66
Sites: 11



Other sources of variability

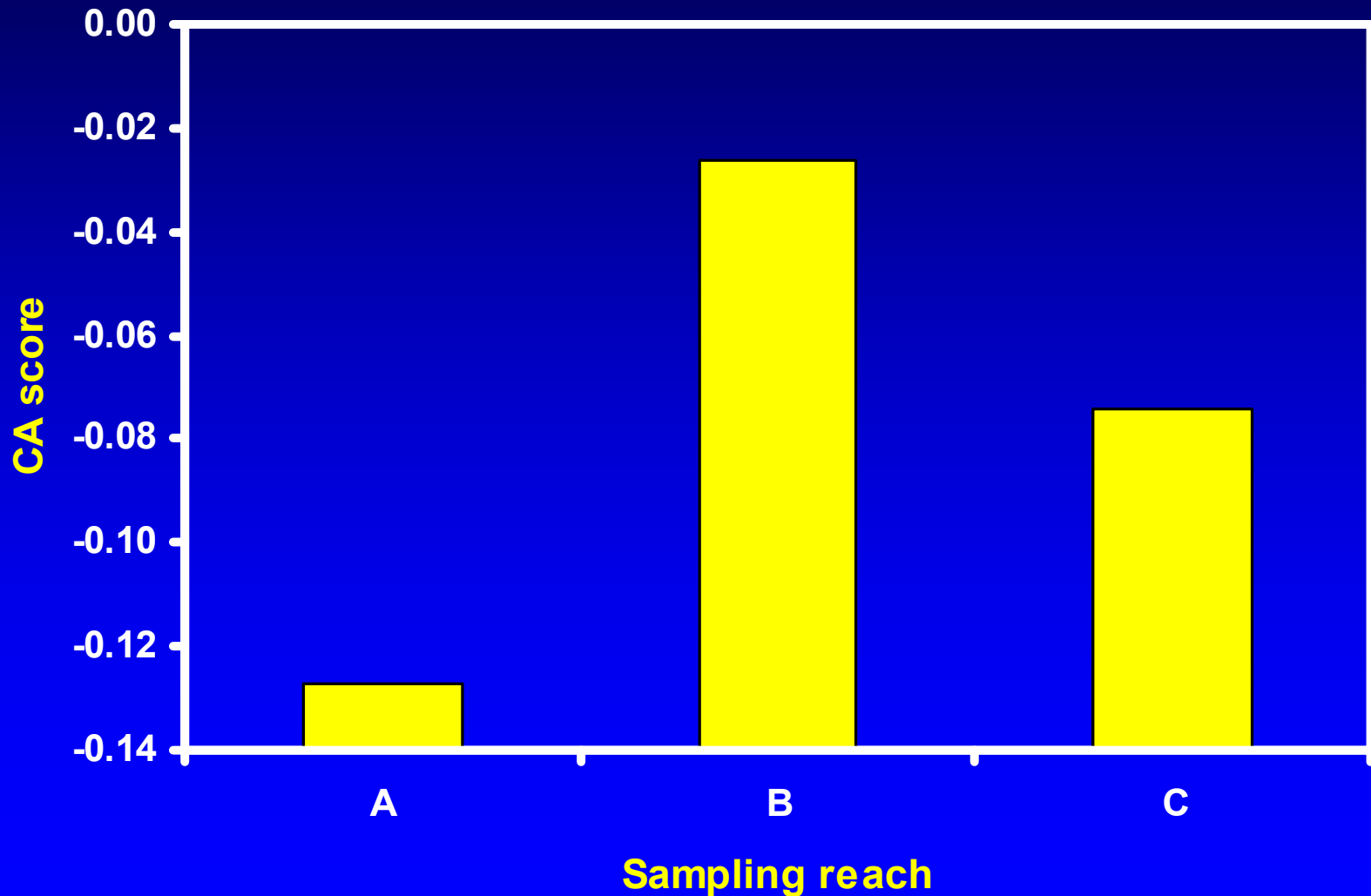
Multiple year and multiple reach

Variability among reaches



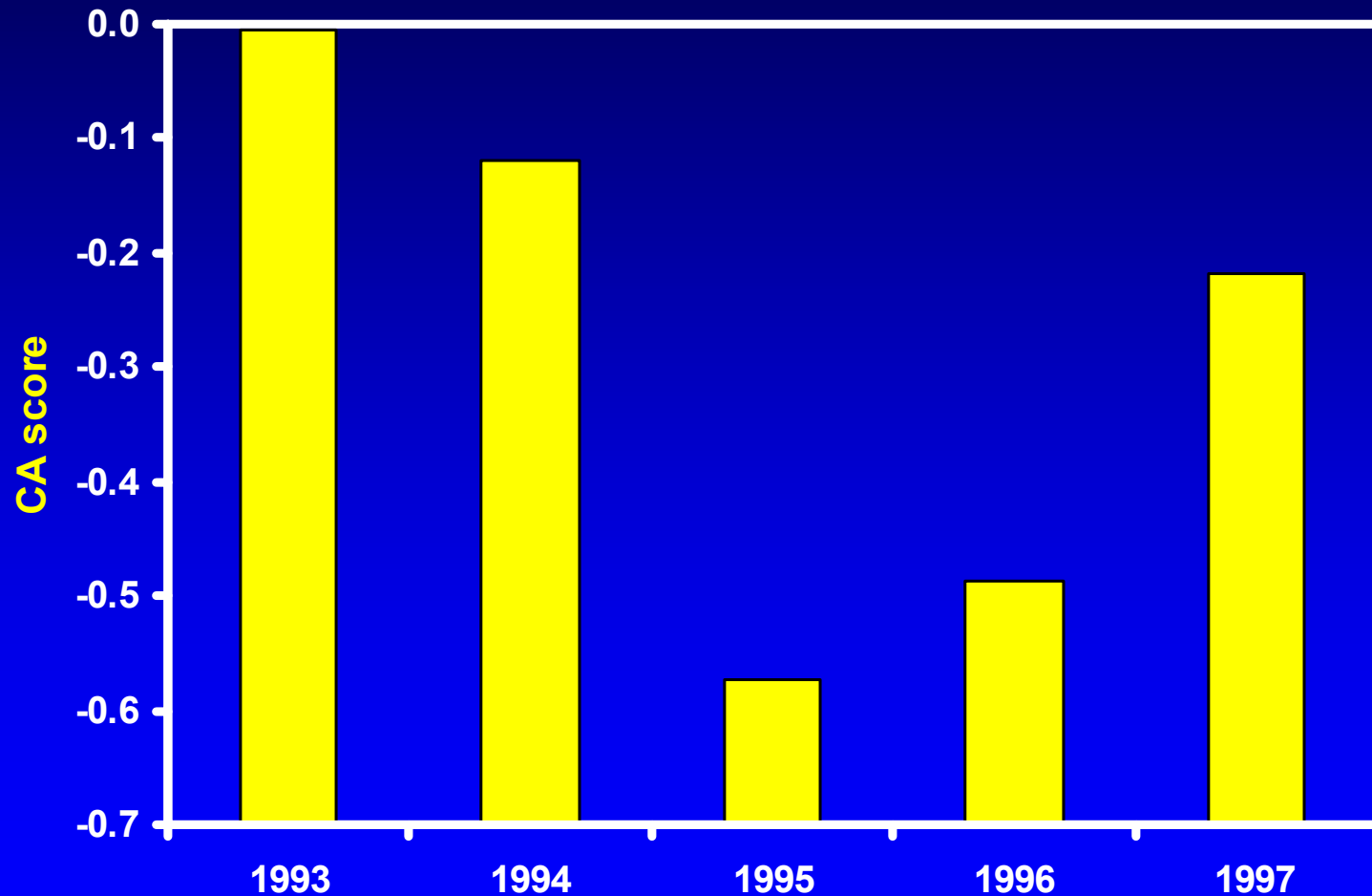
Variability among reaches

ALBE 02087580 2002 QMH based on ALBE urban model



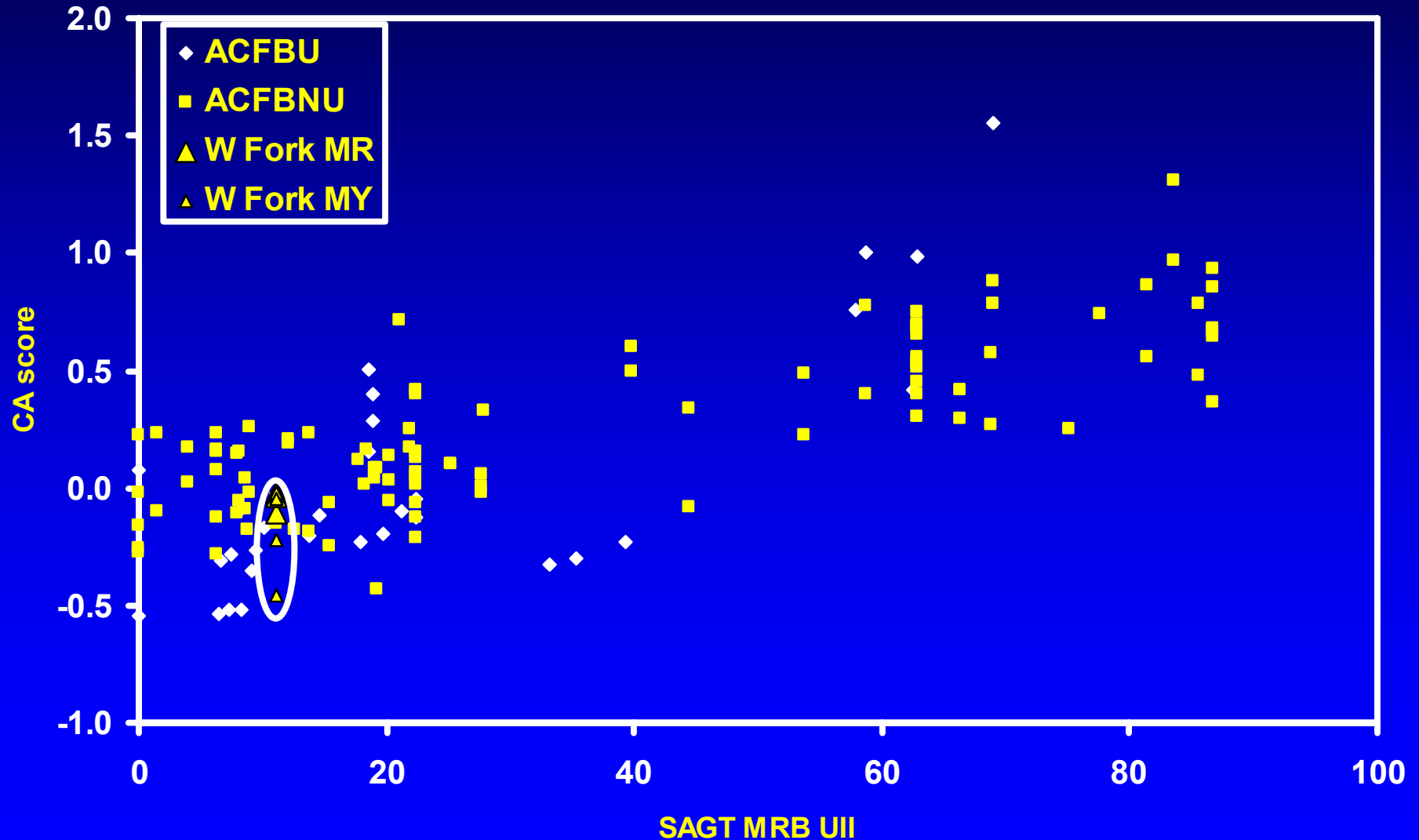
Variability among years

ACFB 02332830 QMH based on ACFB urban model



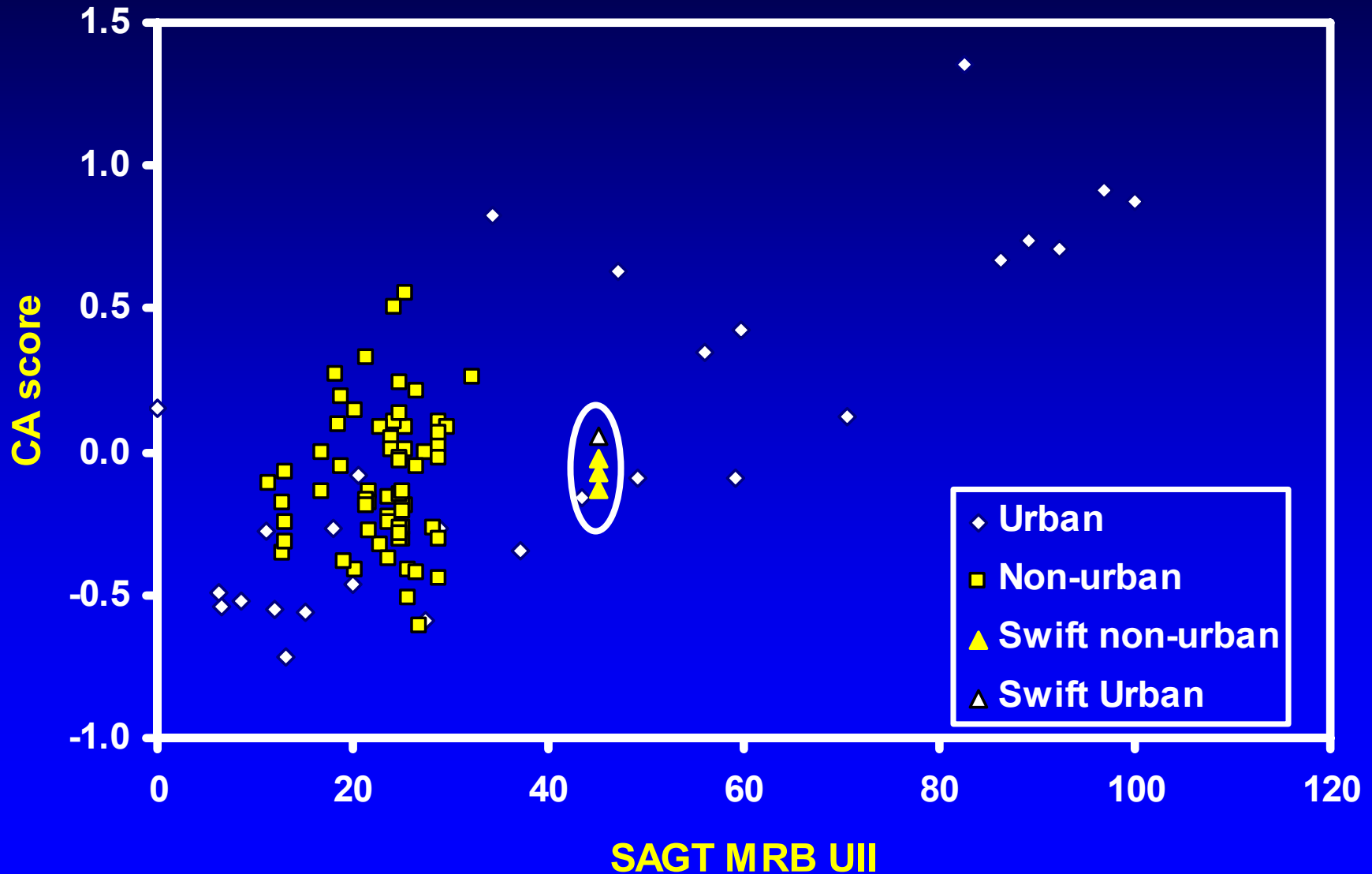
Multiple Reach, Year: W. Fork

ACFB 02332830 (W. Fork Little R.) QMH based on ACFB urban model



Multiple Reach: Swift Creek

ALBE 02087580 2002 QMH based on ALBE urban model



Multiple year & Multiple reach
work is ongoing

Expected major findings

1. **Ordination and multiple regression models are effective predictors of response** (ordination models limited by comparability of modeled fauna).
2. **Models are applicable to other urban areas with similar environmental settings and stream sizes.**
3. **Models are not applicable in areas with widely different environmental settings** (ecoregions and/or stream size – Piedmont vs. Coastal Plain).
4. **Snag vs. Riffle sampling is not as big a problem as expected.**
5. **Model validation and expansion is “best” accomplished by directed *a priori* studies rather than *a posteriori* (i.e., Cycle I data poorly represent the extent of urban areas in SAGT).**

Major tasks

- **Revision of ancillary data sets:**
 - 1990 popln, census, nlcd, infrastructure
 - 2000 popln, census, nlcde, infrastructure
- **Revision of models with corrected ancillary data.**
- **Analysis of multiple year/reach data.**

Timeline

- **Depends on...**
 - **Staffing: Robin and Elise's availability**
 - **EUSE demands**
- **Goal is to have a journal article in review by end of FY.**