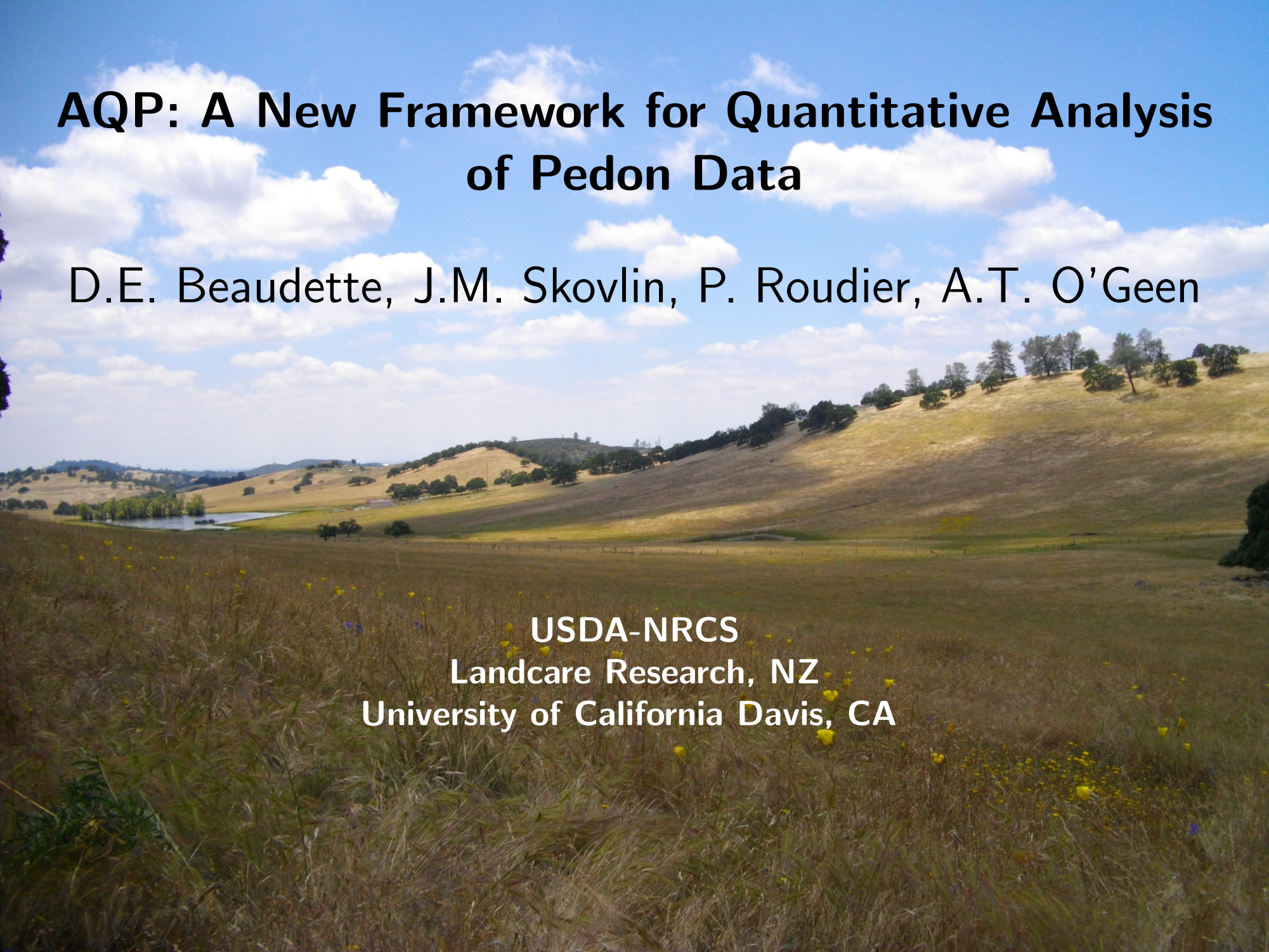
AQP: A New Framework for Quantitative Analysis of Pedon Data

D.E. Beaudette, J.M. Skovlin, P. Roudier, A.T. O'Geen

USDA-NRCS
Landcare Research, NZ
University of California Davis, CA



What is **R**? (if you use NASIS, you already have R!)

R is a drop-in replacement for 99% of tasks done via spreadsheet, and more

- univariate / multivariate summaries
- graphical representations of complex data
- extensible via "packages"
- 2800+ packages on CRAN: 100+ packages on GIS, ecology, and **soils**!
- $\rightarrow$ repeatable, self-documenting work

R is a convenient environment testing ideas

- testing by eye: simple graphical "grammar" used to plot data
- formalized testing: correlation, regression, classification, ordination, ...
- $\rightarrow$ algorithm development by experts, application by trained users

R is good for rapid development of new software

- I/O capabilities: file, URL, SOAP, SQL, ODBC, PDF, PNG, SHP, KML, ...
- optimizers, matrix operations, custom data structures, ...
- $\rightarrow$ well-tested, huge support base, mailing lists, **free**

The Algorithms for Quantitative Pedology (aqp) Package

Soil Profile Visualization

- sketches depicting horizon sequence, depth, color, ...
- soil colors translated from Munsell system to RGB triplets
- ordering of sketches along environmental gradients

Soil Profile Aggregation

- summary of soil property or morphology along regular depth slices
- transcends typical problems with soil profile collections:
  $\rightarrow$ high variability in horizon description style and horizon types / depths

Soil Profile Classification

- similar to above: distance metric computed along regular depth slices
- can assist with similar/dissimilar evaluation

Soil-Specific Classes/Methods

- data structures & vocabulary for working with collections of soil profiles

The Soil Database Interface (`soilDB`) Package

PedonPC

- get all pedons from a 'pedon.mdb' file: `fetchPedonPC(dsn)`

Local NASIS, after querying national database

- get all associated pedons: `fetchNASIS()`
- get all assocaited DMU data: `fetchNASIS_component_data()`

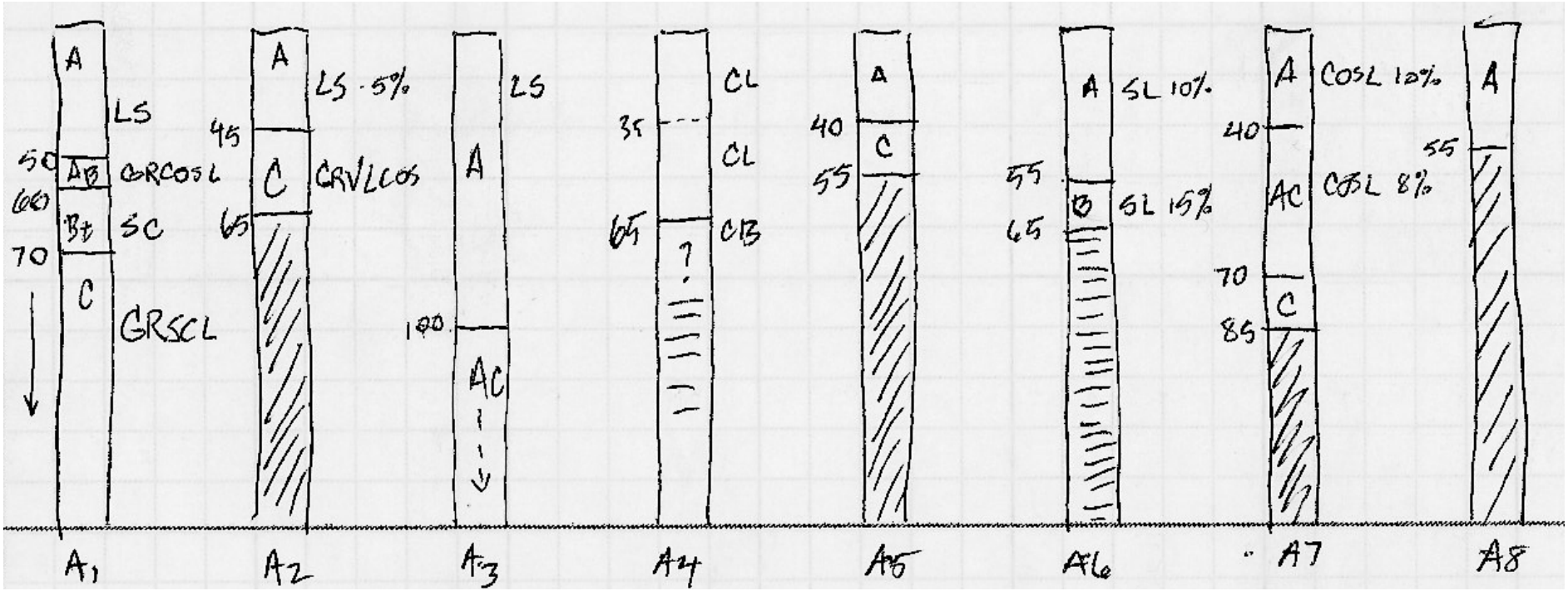
SSURGO Data - SDA

- get SSURGO (tabular data) via SQL: `SDA_query(query)`
- get SSURGO linework via BBOX: `mapunit_geom_by_ll_bbox(BBOX)`
- get SSURGO mukey(s) via BBOX: `MUKEYS_by_ll_bbox(BBOX)`

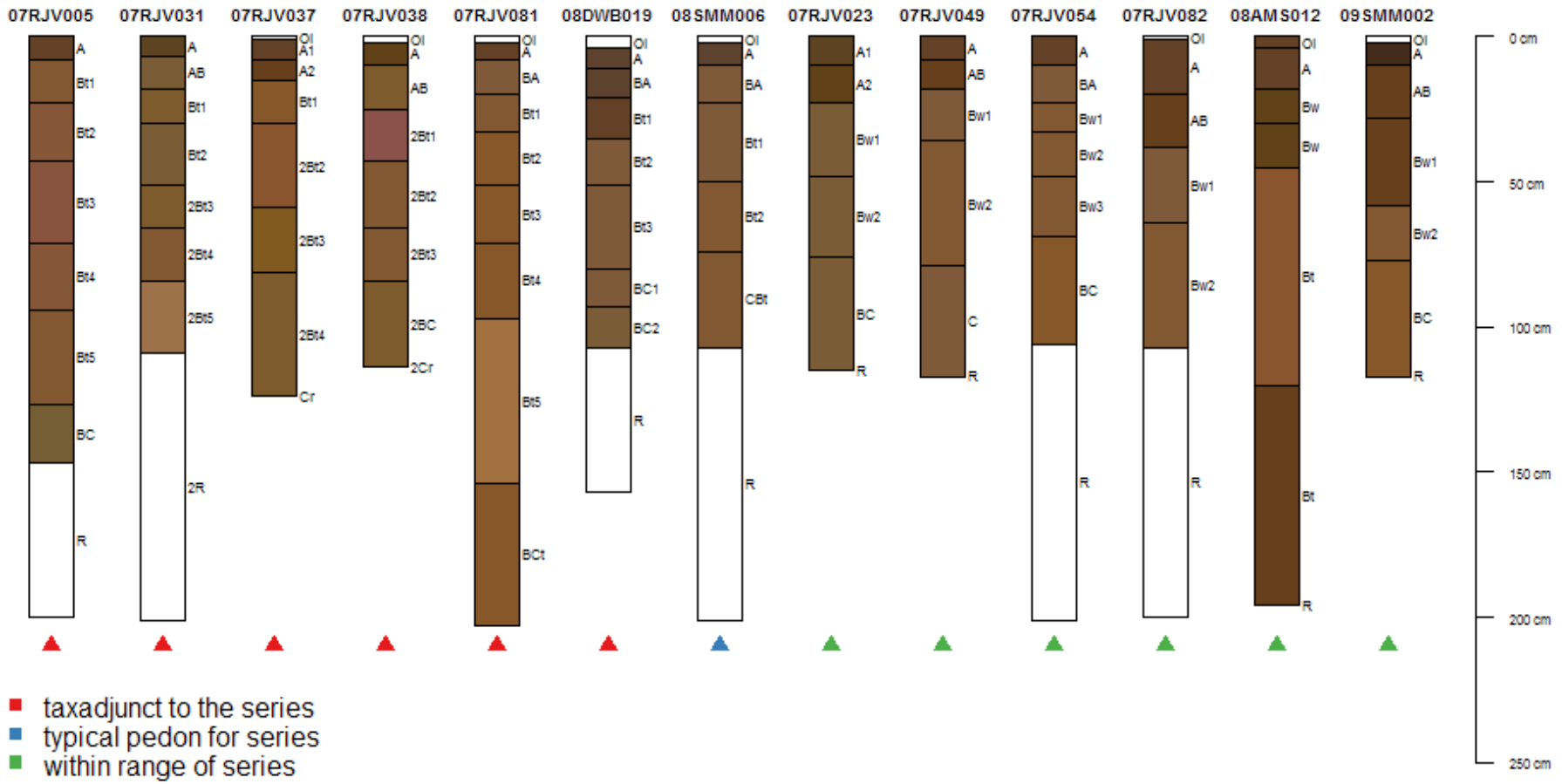
Results

- `SoilProfileCollection` object $\rightarrow$ spatial-site-diagnostic-pedon-hz
- colors converted from Munsell to RGB
- hz-errors flagged

Soil Profile Sketches

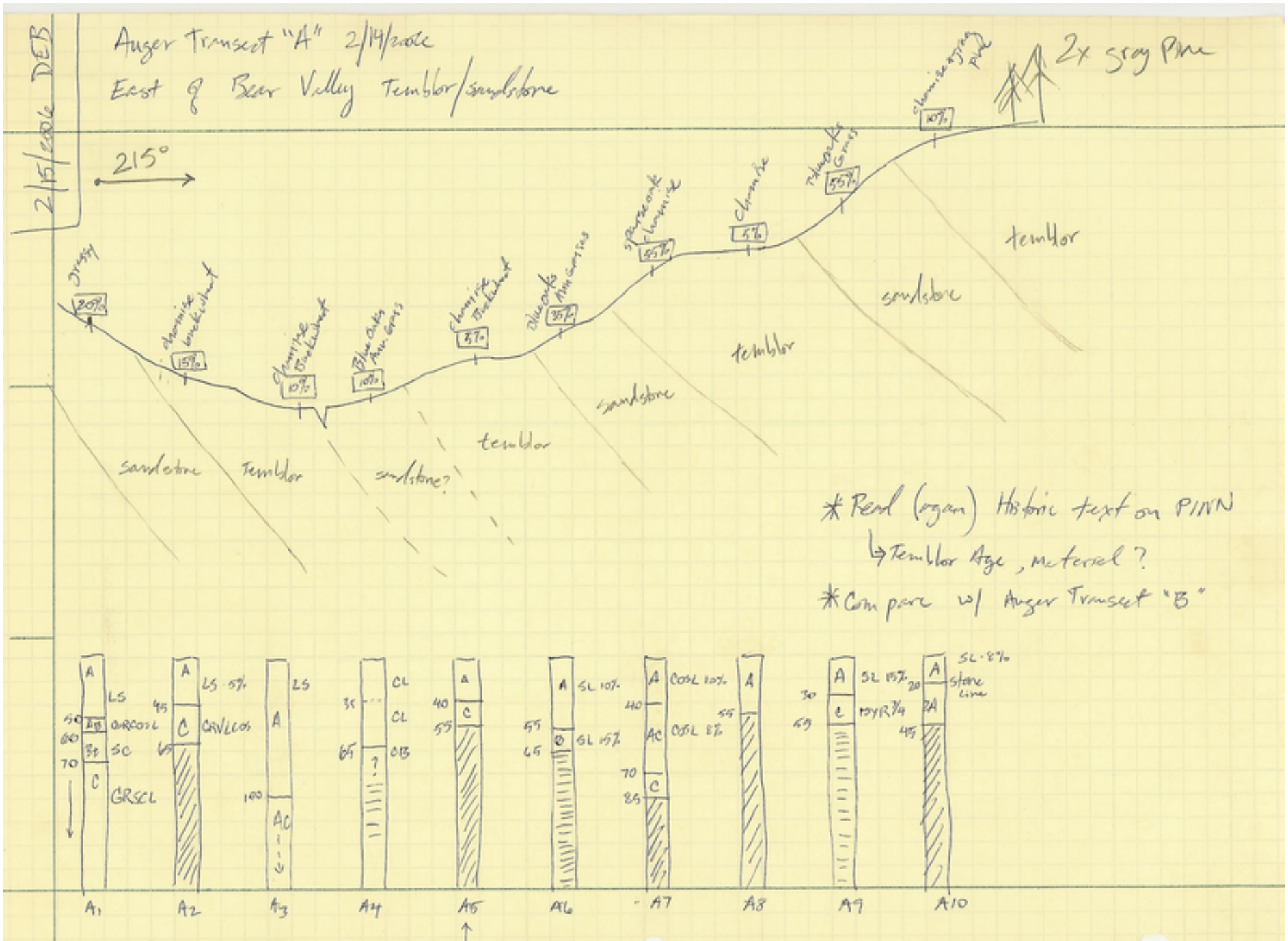


(Digital) Soil Profile Sketches: Pedon Kind

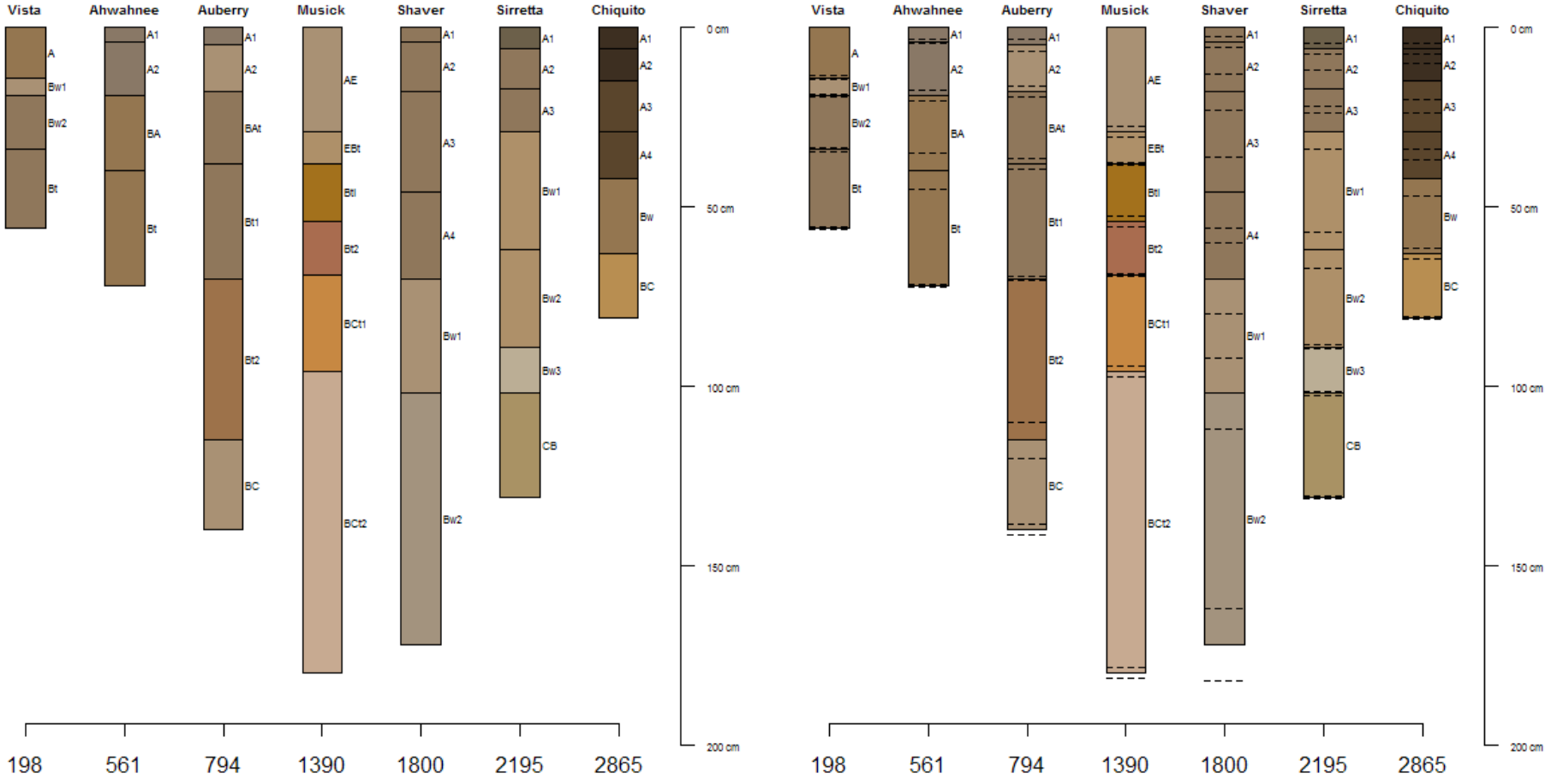



```
library(soilDB) # load package
f <- fetchNASIS() # load pedons from NASIS
new.order <- order(f$pedon_kind) # new ordering vector
plot(f, name='hzname', plot.order=new.order) # plot pedons as sketchs
legend(...) # make a legend (details ommited)
```

Soil Profile Sketches & Landscape Gradients

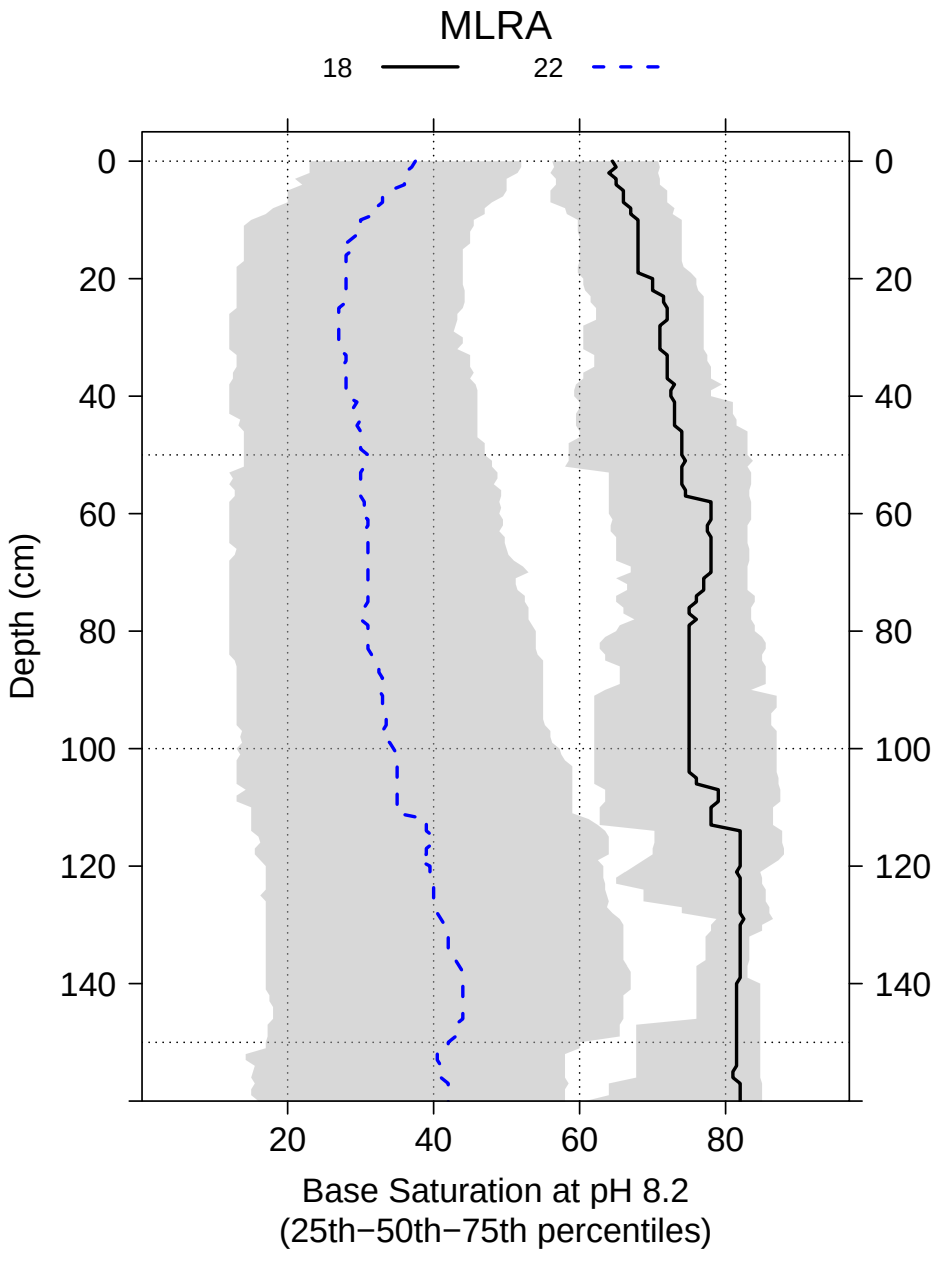


(Digital) Soil Profile Sketches & Climatic Gradients




```
x <- read.csv('dahlgren-granitics.csv') # load data from CSV file
x$soil_color <- with(x, munsell2rgb(hue, value, chroma)) # convert Munsell to RGB
x$HzD <- hzDistinctnessCodeToOffset(substr(x$hz_boundary, 0, 1)) # convert hz-distinct code

depths(x) <- id ~ top + bottom # init SoilProfileCollection object
site(x) <- ~ elev + MAAT + MAP + geo # split site vs. hz-level data

g.new.order <- order(x$elev) # generate plotting order via elevation
plot(x, name='name', plot.order=g.new.order, hz.distinctness.offset='HzD')
axis(1, at=1:length(x), labels=x$elev[g.new.order], line=-2)
```

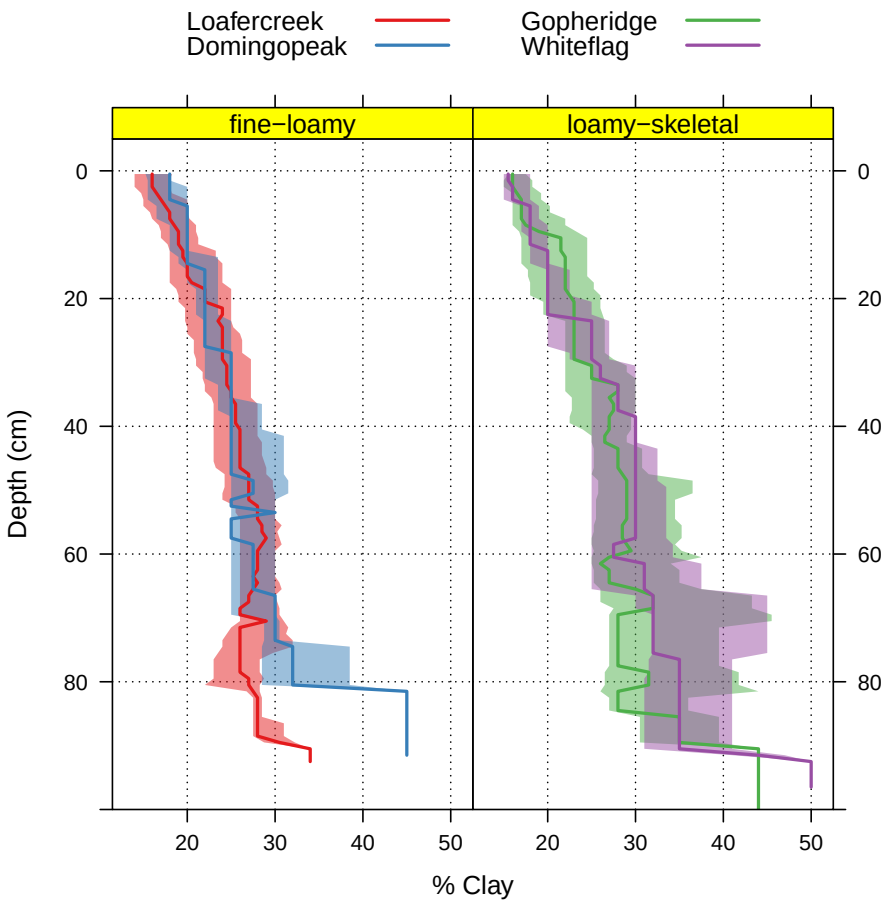
Aggregate Representation of Soil Profile Collections

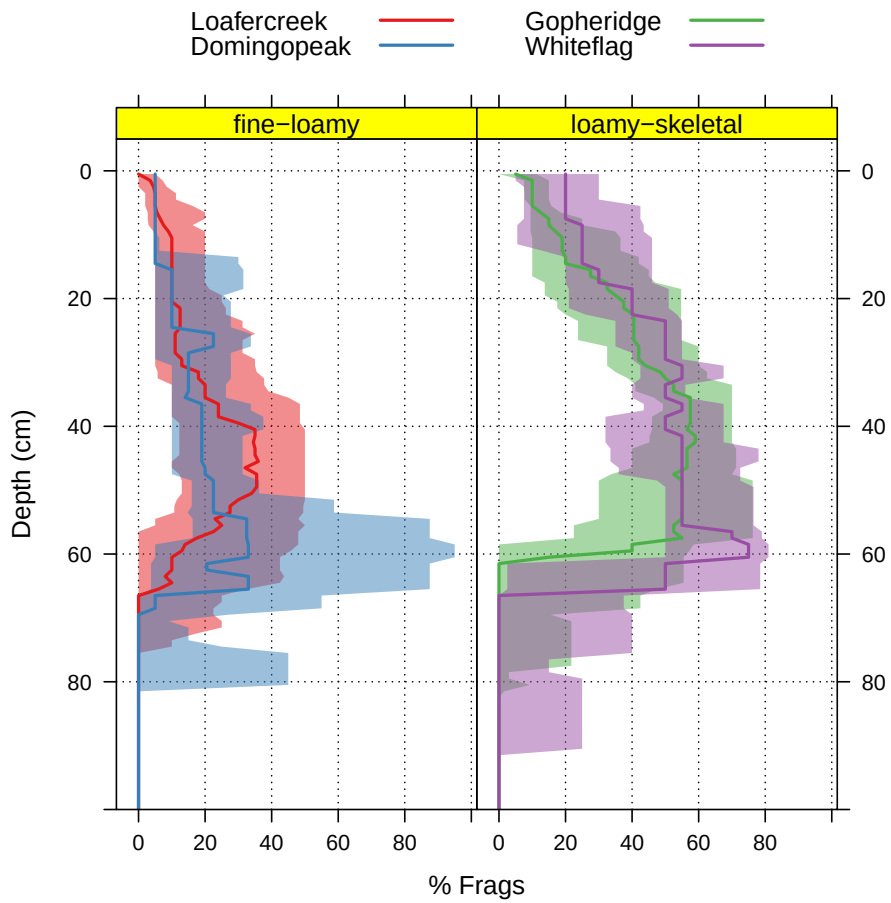


Slice-wise aggregation of soil properties

```
library(aqp)
library(lattice)
# load lab data from CSV file
x <- read.csv('MLRA_18_22-lab_data.csv')

# init SoilProfileCollection object
depths(x) <- pedon_id ~ hztop + hzbot
site(x) <- ~ site_id + lon + lat + mlra

# slice-wise aggregation by MLRA of \%BS at pH 8.2
a <- slab(x, mlra ~ bs82)

# plot
xyplot(
top ~ p.q50, groups=mlra, data=a, lower=a$p.q25, upper=a$p.q75,
ylim=c(160,-5), alpha=0.5, scales=list(cex=1.25,
y=list(tick.num=7, alternating=3), x=list(alternating=1)),
panel=panel.depth_function, prepanel=prepanel.depth_function,
ylab=list('Depth (cm)', cex=1.25),
xlab=list('Base Saturation at pH 8.2', cex=1.25), par.settings=list(
superpose.line=list(col=c('black','blue'), lty=c(1,2), lwd=2)),
auto.key=list(columns=2, title='MLRA', points=FALSE, lines=TRUE)
)
```

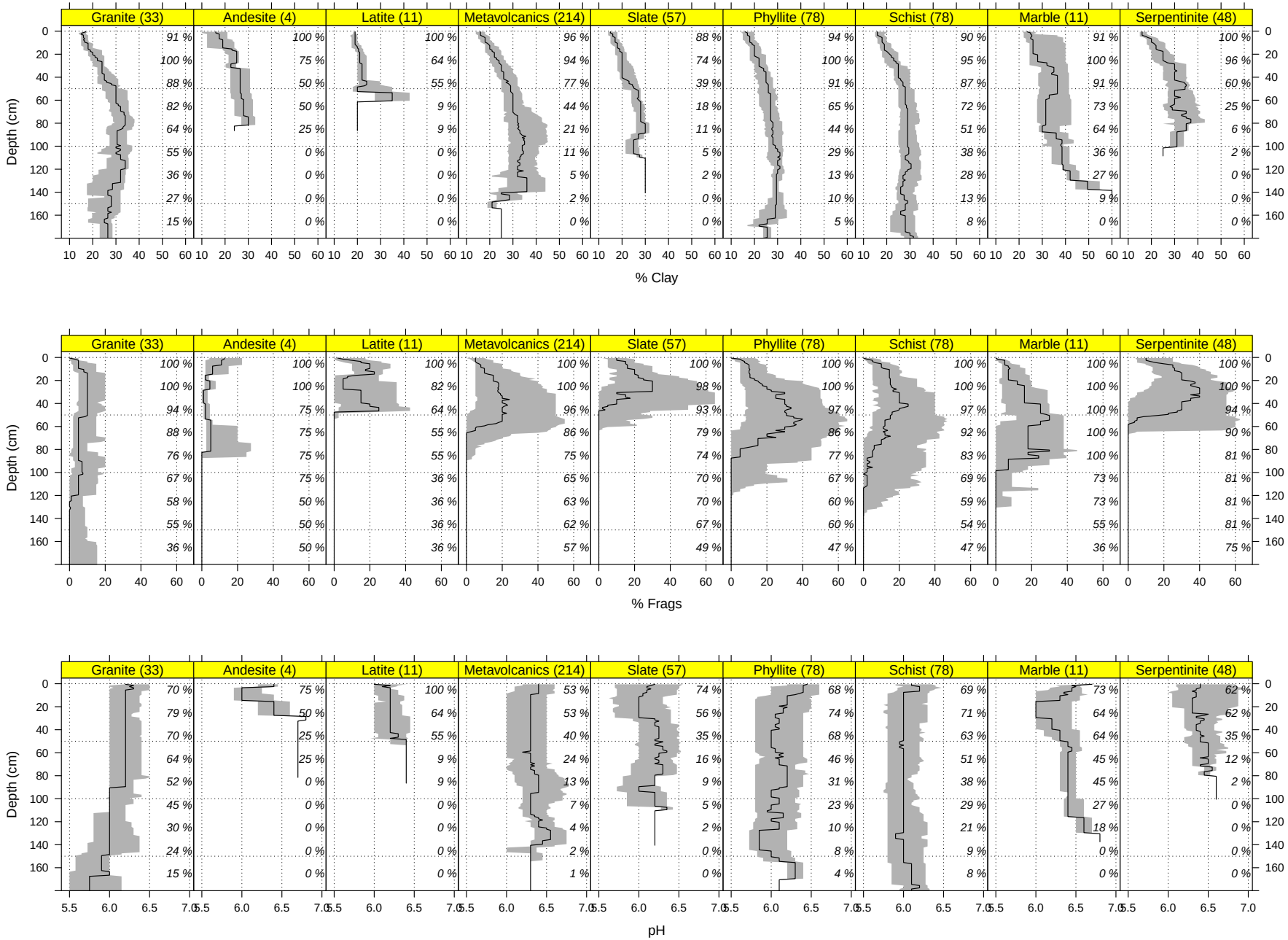
Aggregate Representation of Soil Profile Collections



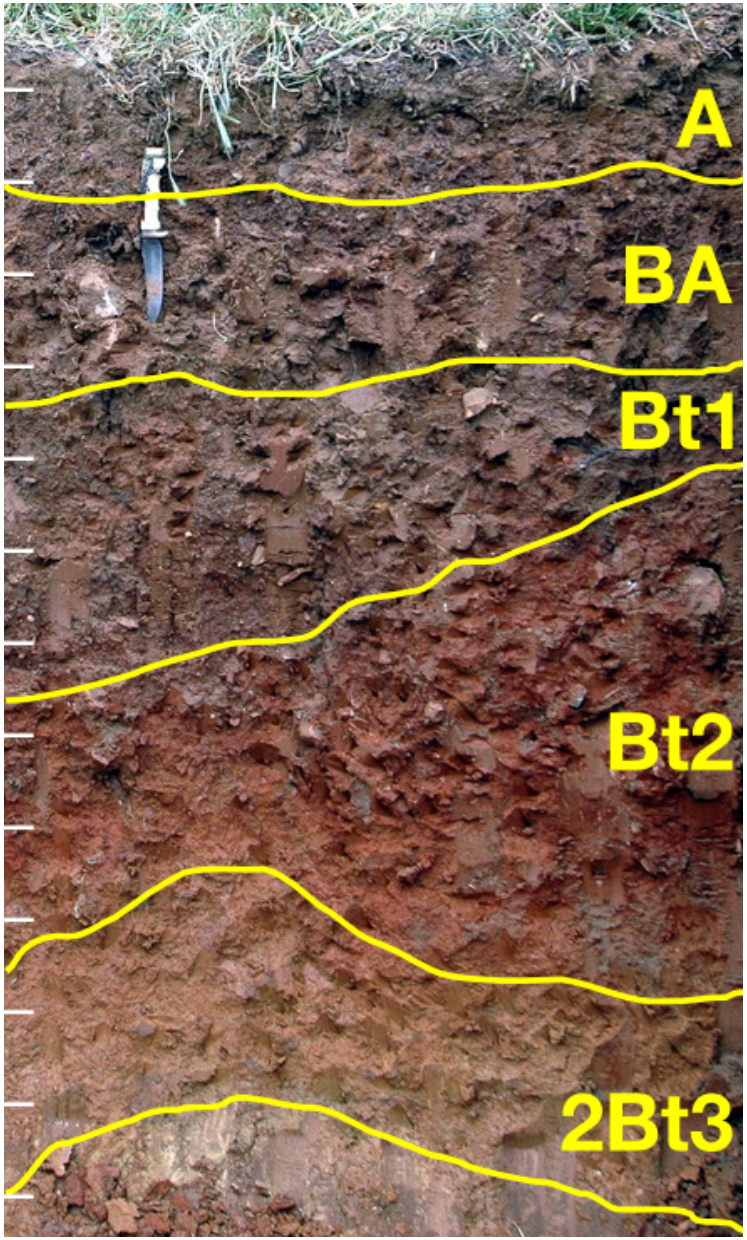


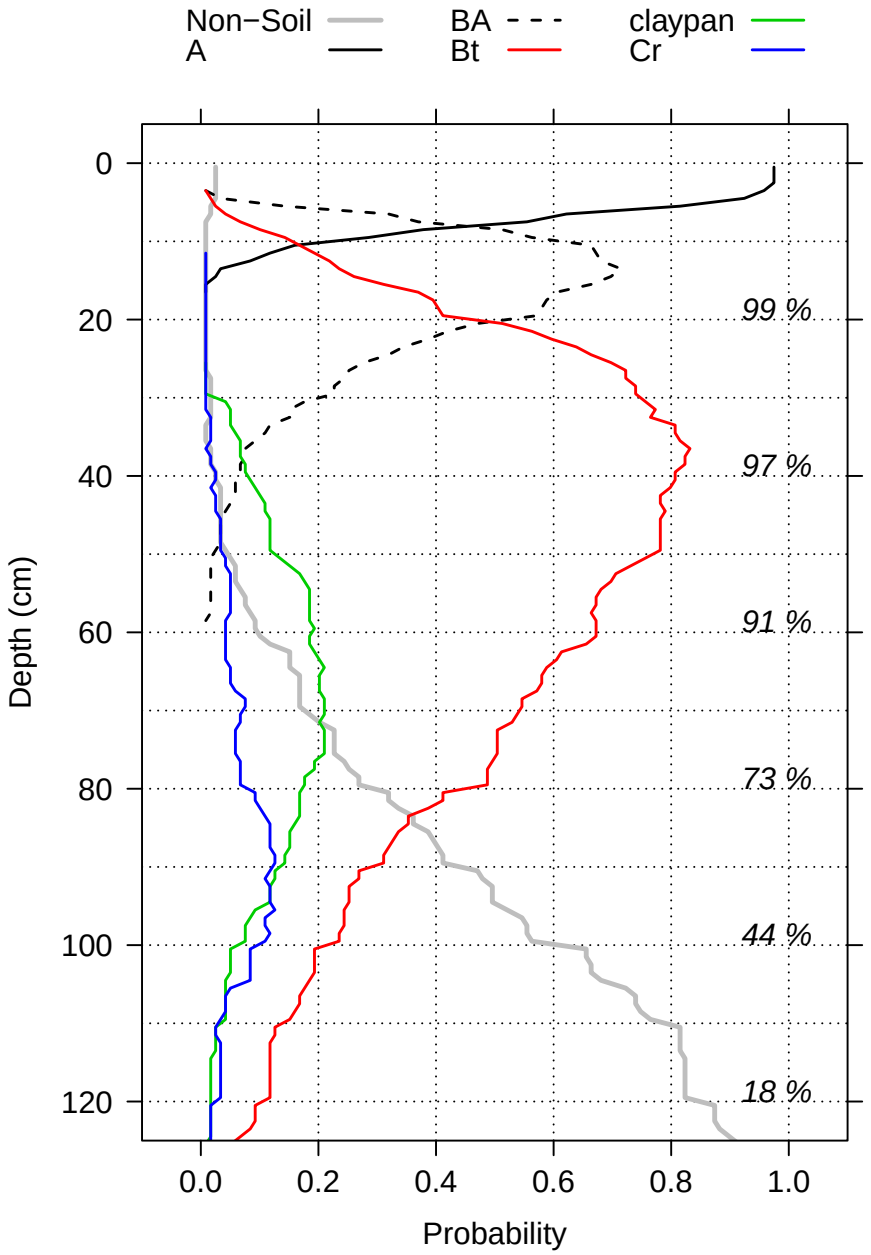

```
# load package
library(soilDB)
# load MVO soils from local NASIS (query national first)
f <- fetchNASIS()
# slice-wise aggregation of select properties
a <- slab(f, sampled_as ~ clay + total_frags_pct + phfield)
# plotting details ommitted...
```

Aggregate Representation of Soil Profile Collections



Aggregate Representation of Soil Profile Collections





Quantitative (pair-wise) Comparison of Soils

(a)

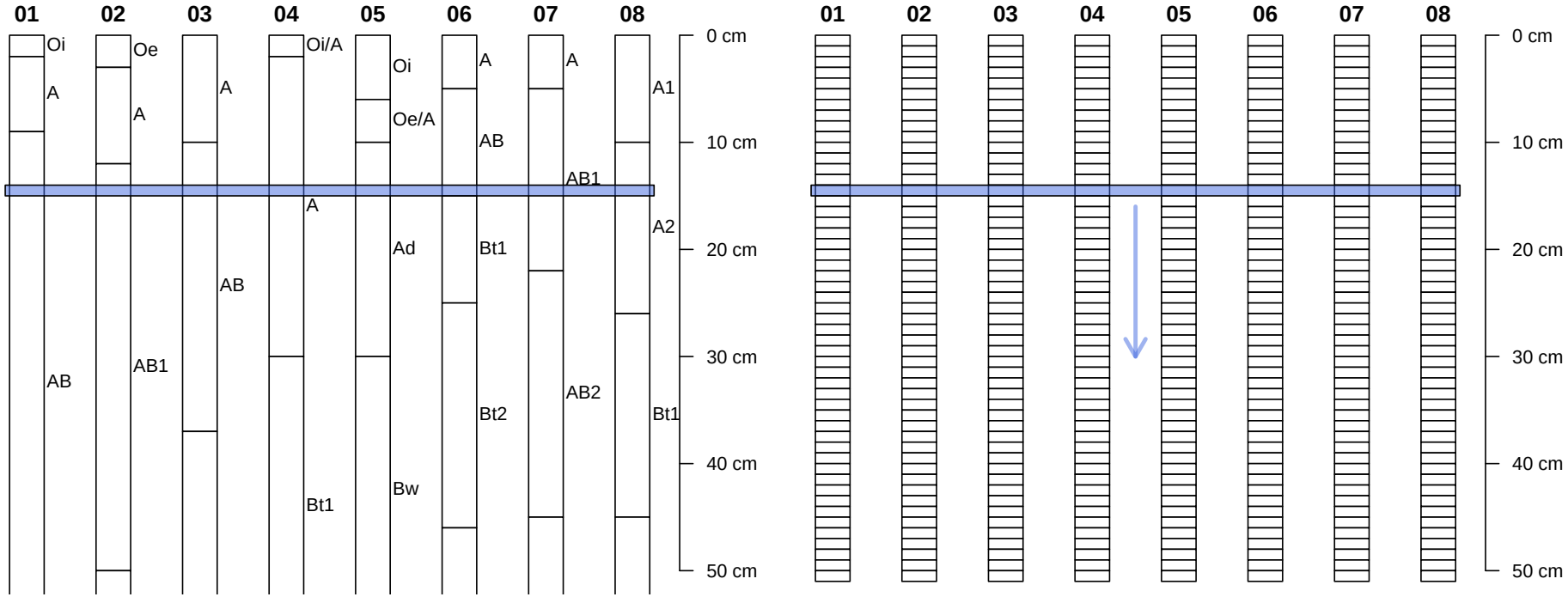
(b)

(c)

"is **a** more like **b**, as compared to **c**?"

ideally transcending horizonation and description style

Pair-wise dissimilarity along depth-slices (Moore et al, 1972)

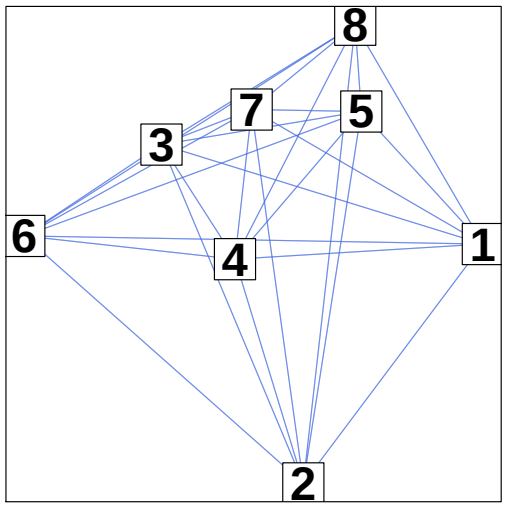


soil properties at slice 15

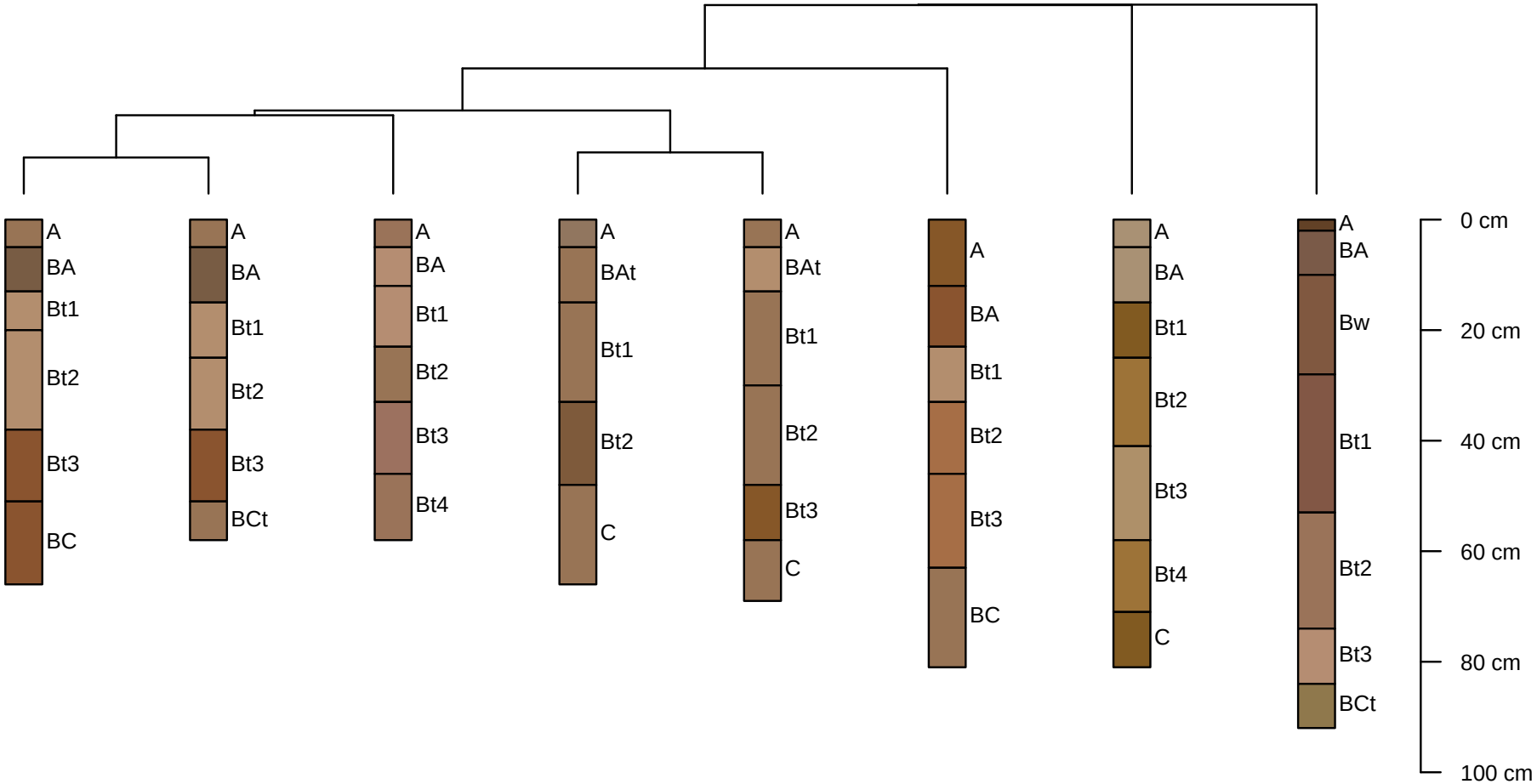
	clay	vcs	ln_tc	cec	A
1	6.4	21.5	-1.5	3.8	8.1
2	10.6	17.6	-1.5	5.9	4.9
3	8.8	10.5	-0.2	7.5	6.0
4	9.1	12.6	-0.8	5.6	5.9
5	6.8	16.2	-0.5	5.0	8.2
6	17.9	11.0	-0.1	9.4	5.2
7	7.0	11.7	-0.4	4.7	6.3
8	6.3	17.0	-0.1	4.7	7.9

pair-wise dissimilarity at slice 15

	1	2	3	4	5	6	7
2	42						
3	68	46					
4	51	29	19				
5	30	46	40	30			
6	96	59	30	45	68		
7	48	45	20	16	22	48	
8	32	50	39	37	11	64	24

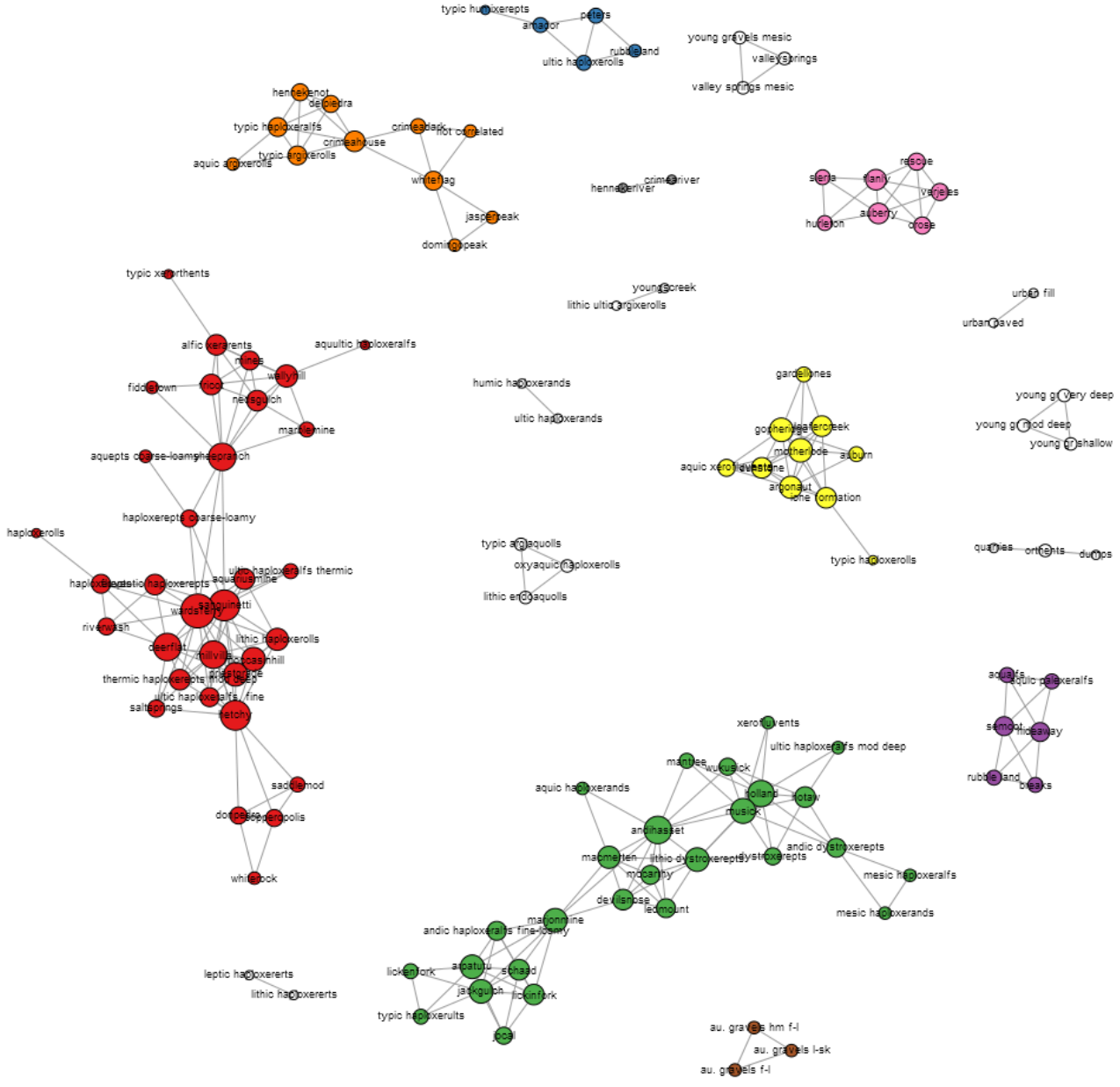


Results can support simmilar/dissimilar evaluation

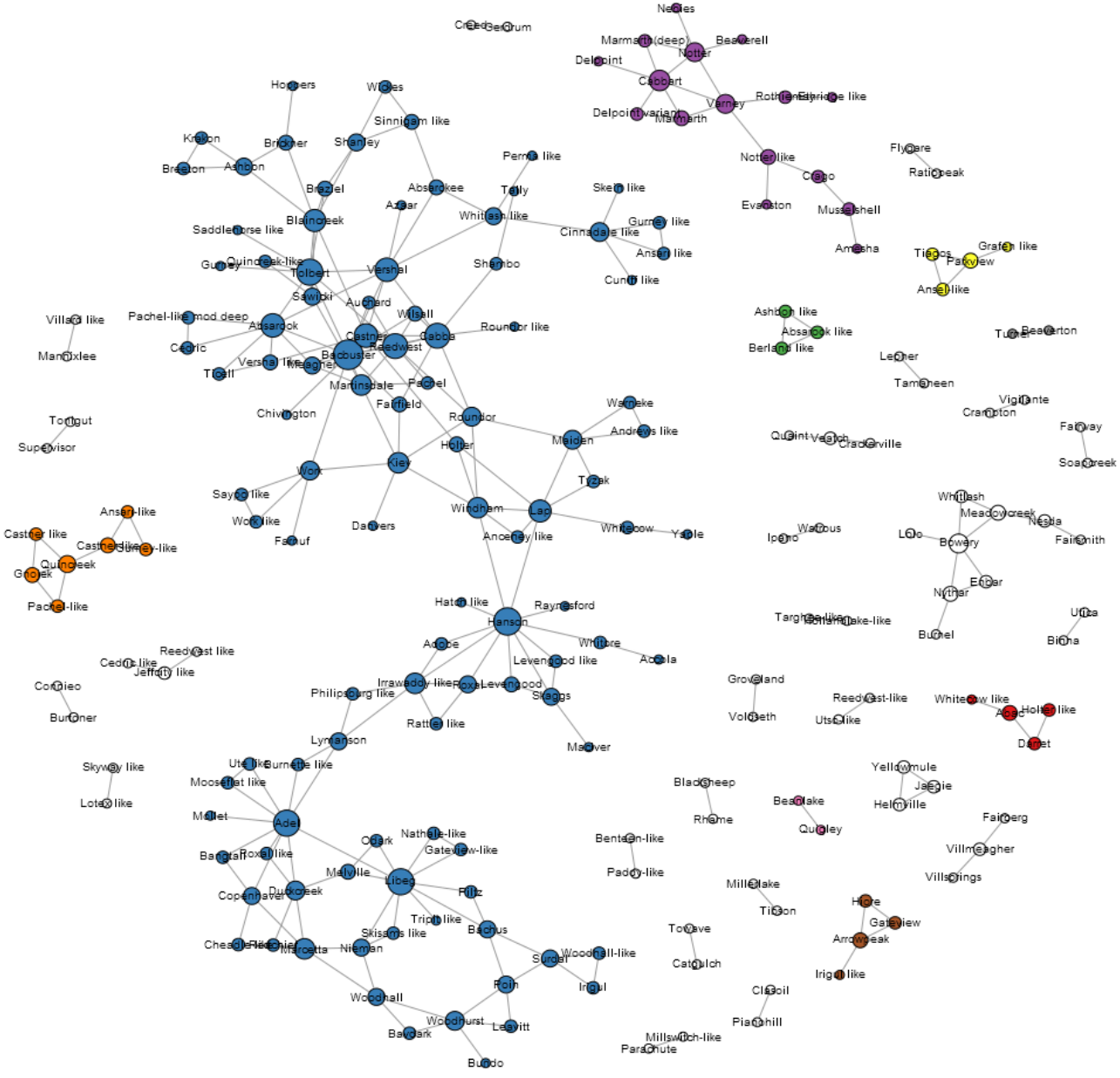


Domingopeak

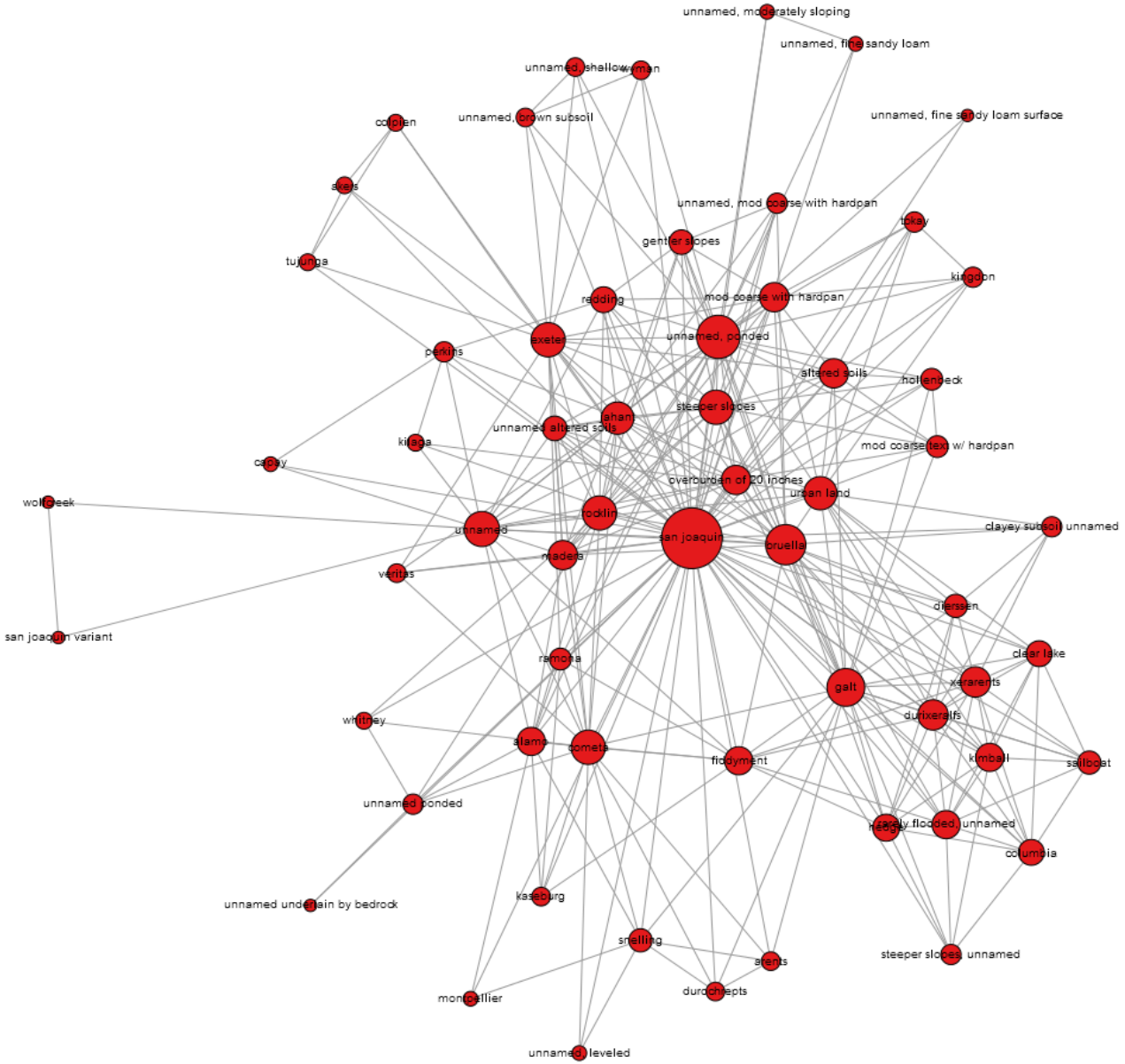
Component Relationship Graphs: CA630



Component Relationship Graphs: MT637



Component Relationship Graphs: San Joaquin Series



Possible Application and Relevance to SDJR

Visualization: rapid evaluation of complex data

- outliers / mistakes tend to stick-out (profile sketches / aggregate data)
- dendrograms depicting similarity (similar / dissimilar eval.)
- graph-networks depicting connectivity (associated / related soils)

Aggregation: data-driven synthesis of a "representative soil profiles"

- depth-slices can transcend variations in description style (with expert review!)
- universally meaningful Low–RV–High values: 5th–50th–95th pctiles
- OSDs / DMUs via direct aggregation: morphology+properties (data permitting!)

Slice-Wise Similarity: property-driven eval. of "similarity"

- rapid ID of mistakes / outliers / missing data (!)
- similar / dissimilar discussion for map unit design / re-correlation / joins
- (semi-)automated re-allocation of profiles to new revised legend or series

Thank You

Online Resources:

- R Manuals: http://cran.r-project.org/manuals.html
- Some Favorites: http://casoilresource.lawr.ucdavis.edu/drupal/node/100
- R-Inferno: http://www.burns-stat.com/pages/Tutor/R_inferno.pdf
- Patrick Burns' Tutorials: http://www.burns-stat.com/pages/tutorials.html
- R-Forge Project Page: http://aqp.r-forge.r-project.org
- In-Depth Examples: http://casoilresource.lawr.ucdavis.edu/drupal/taxonomy/term/56
- Packages on CRAN (http://cran.r-project.org)
  - http://cran.r-project.org/web/packages/aqp
  - http://cran.r-project.org/web/packages/soilDB