A (very) Short Introduction to R, AQP, and soilDB

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**?

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**!
- → 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, ...
- → 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, ...
- → 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 → spatial-site-diagnostic-pedon-hz
- colors converted from Munsell to RGB
- hz-errors flagged

Online Resources

Getting Started with R

- Offical Manuals: http://cran.r-project.org/manuals.html
- CASRL Listing: 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

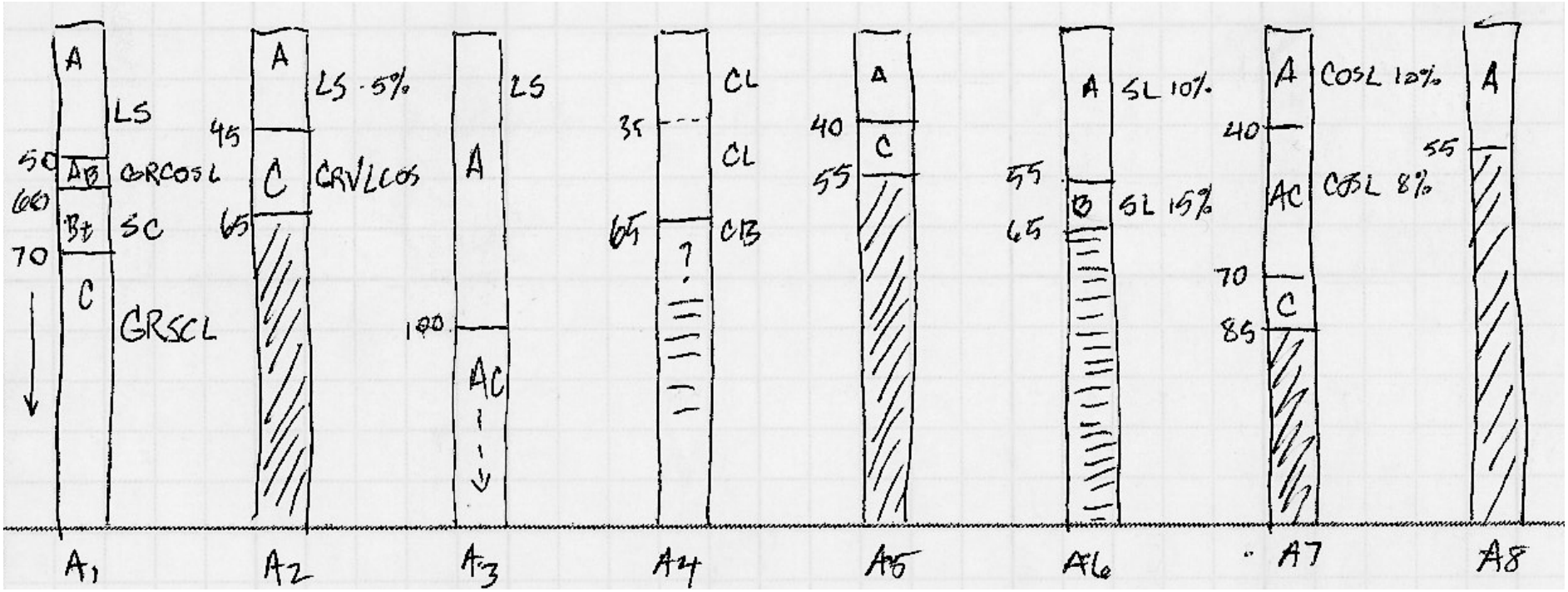
aqp / soilDB Examples

- R-Forge Project Page: http://aqp.r-forge.r-project.org
- In-Depth Examples: http://casoilresource.lawr.ucdavis.edu/drupal/taxonomy/term/56

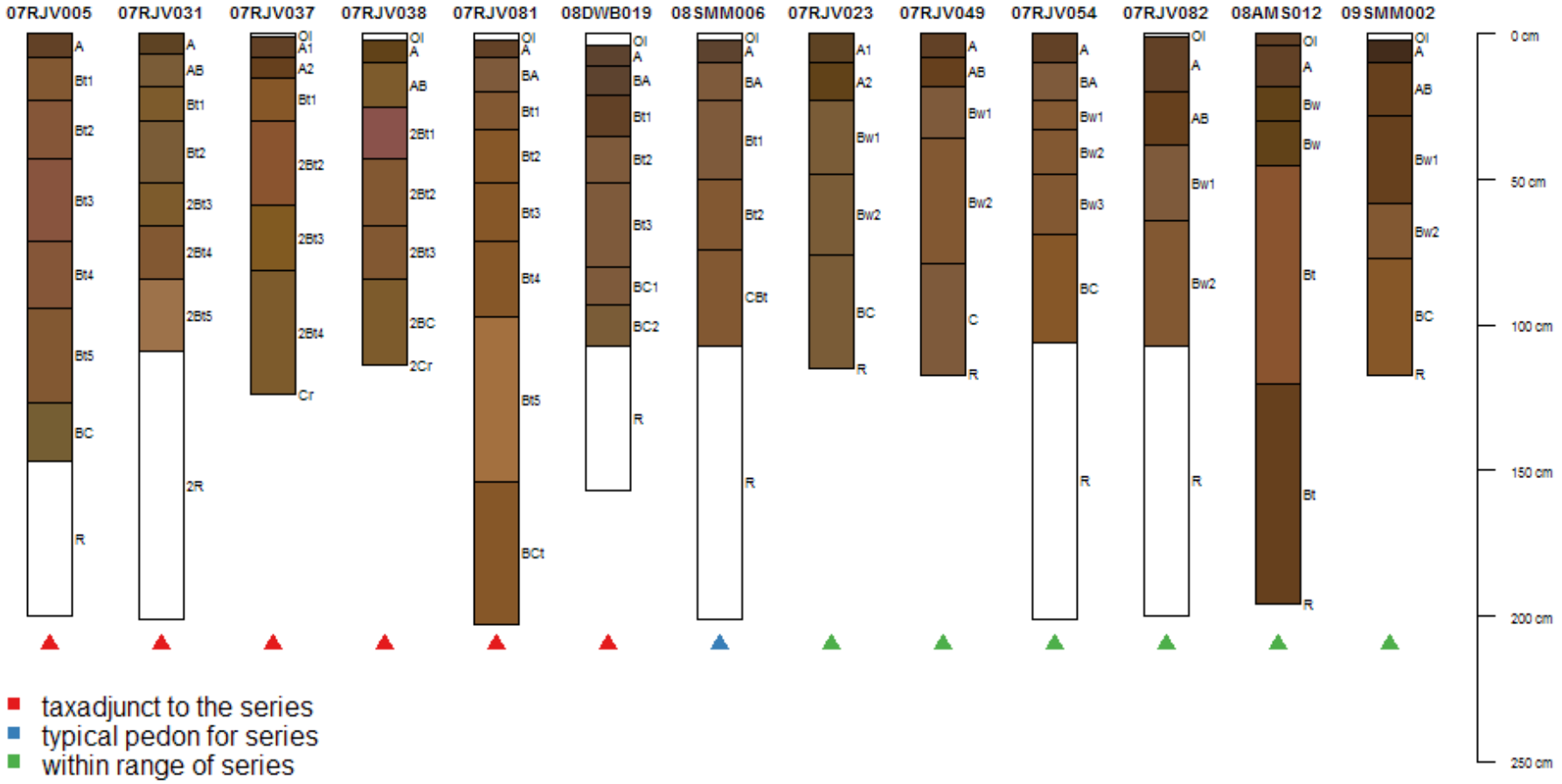
aqp / soilDB Software

- "stable" versions on CRAN (http://cran.r-project.org)
  - http://cran.r-project.org/web/packages/aqp
  - http://cran.r-project.org/web/packages/soilDB
- "development" versions on R-Forge: http://r-forge.r-project.org/R/?group_id=590
- "stable" versions on CRAN (http://cran.r-project.org)

Soil Profile Sketches

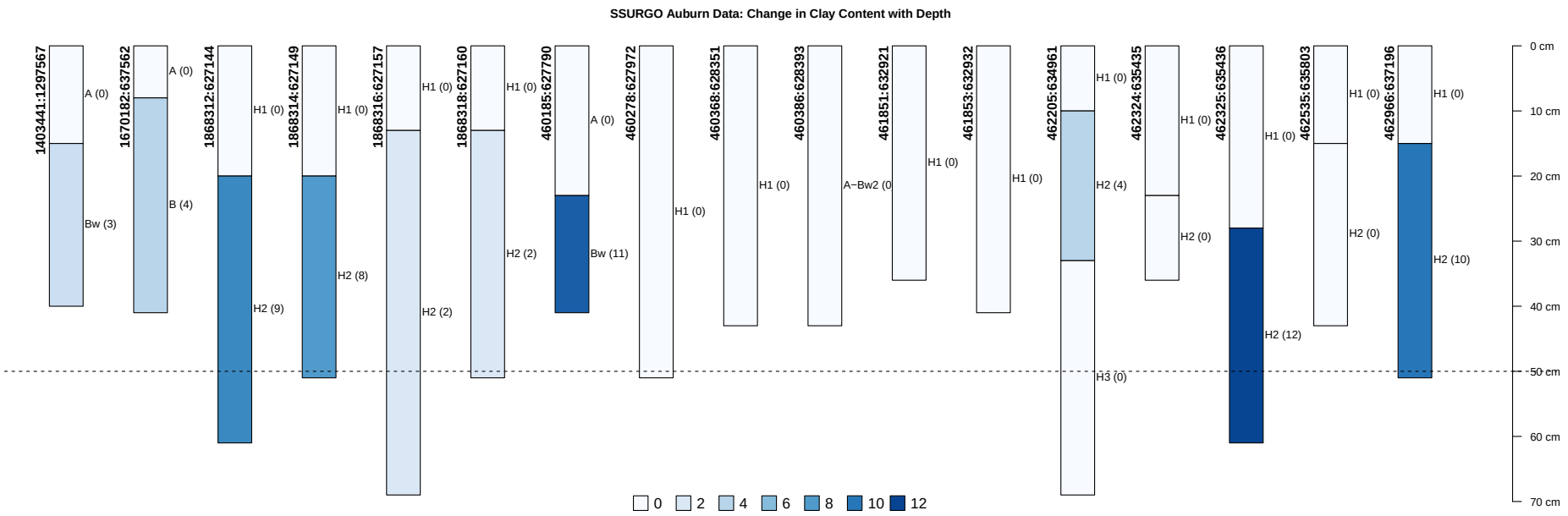


(Digital) Soil Profile Sketches: Pedon Kind

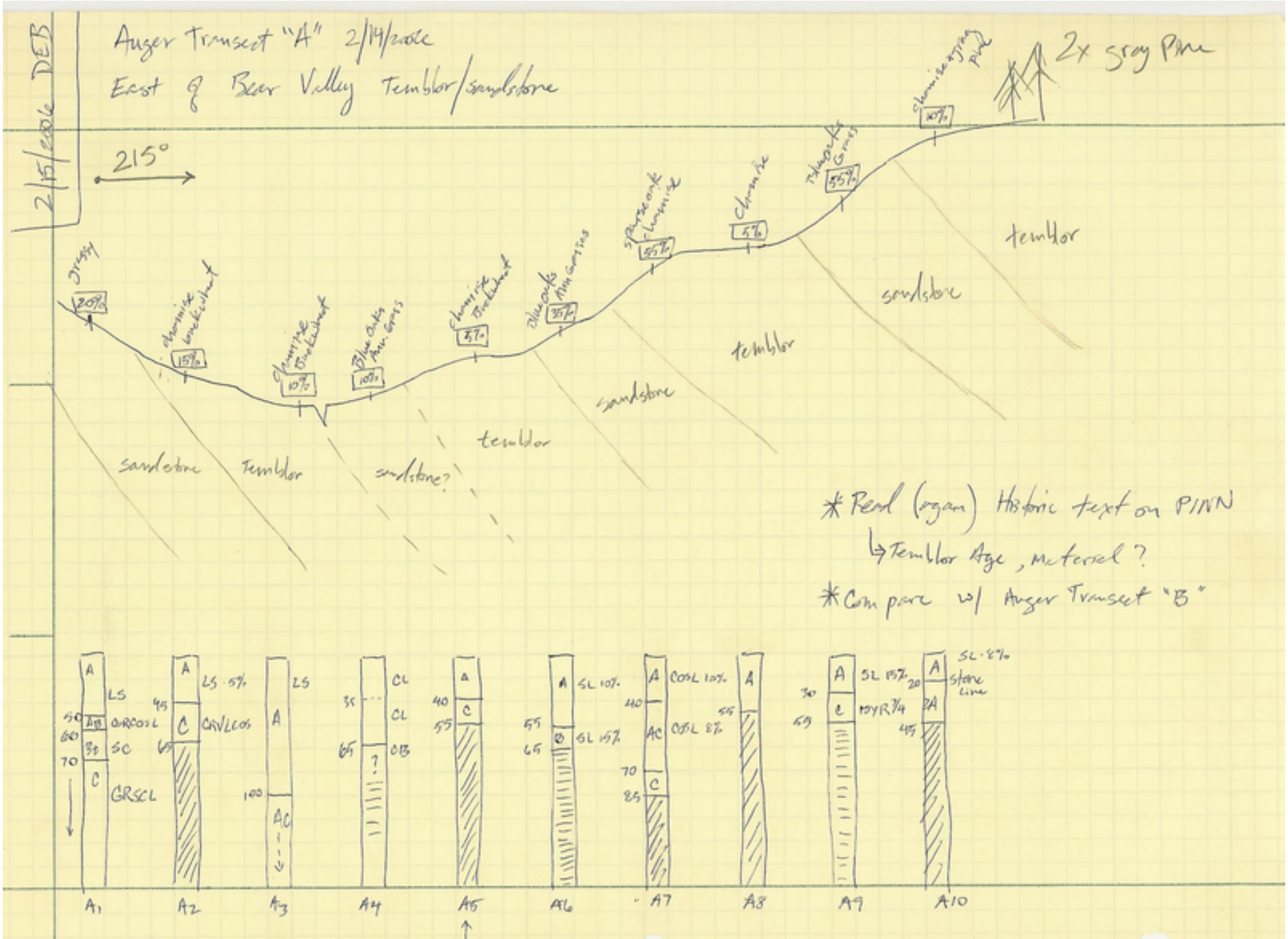


- taxadjunct to the series
- typical pedon for series
- within range of series

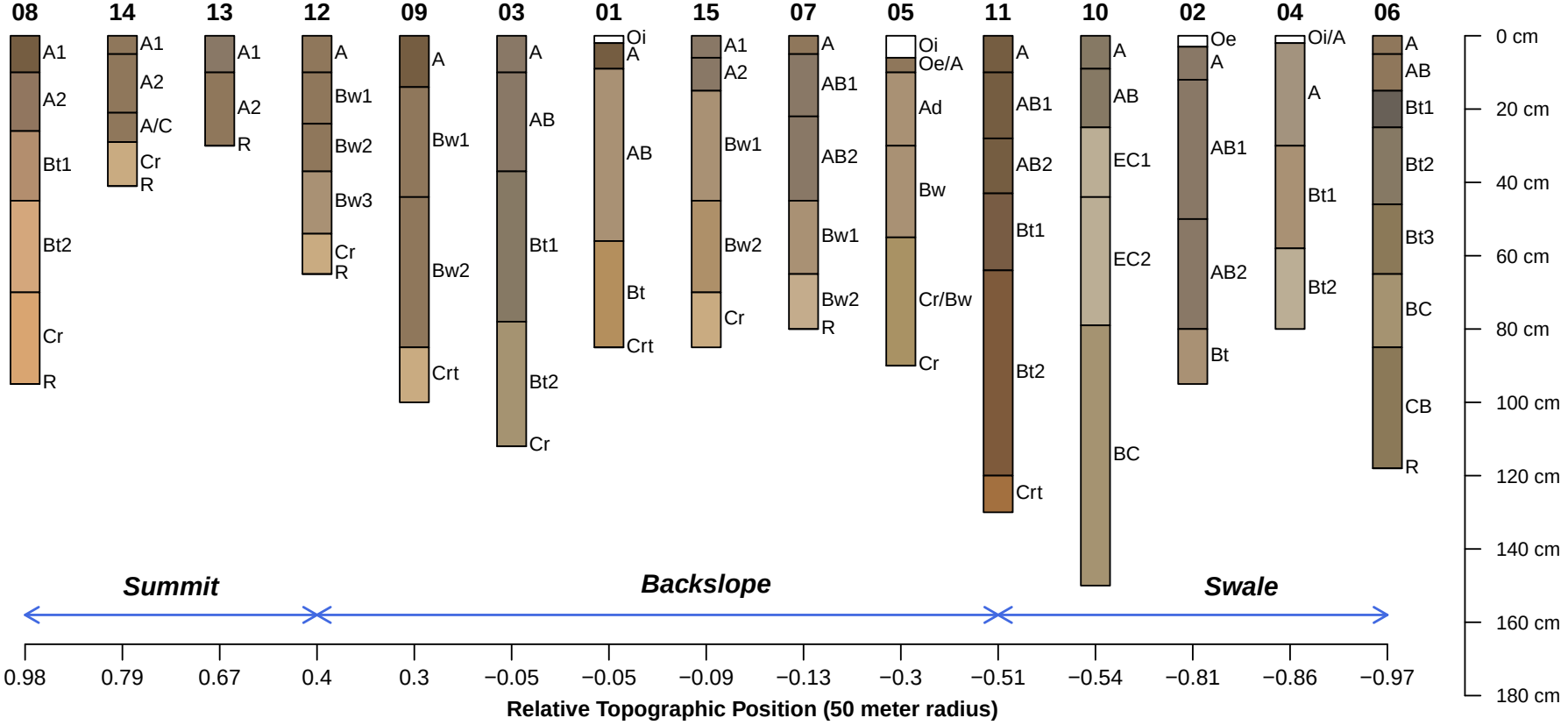
(Digital) Soil Profile Sketches: SSURGO Evaluation



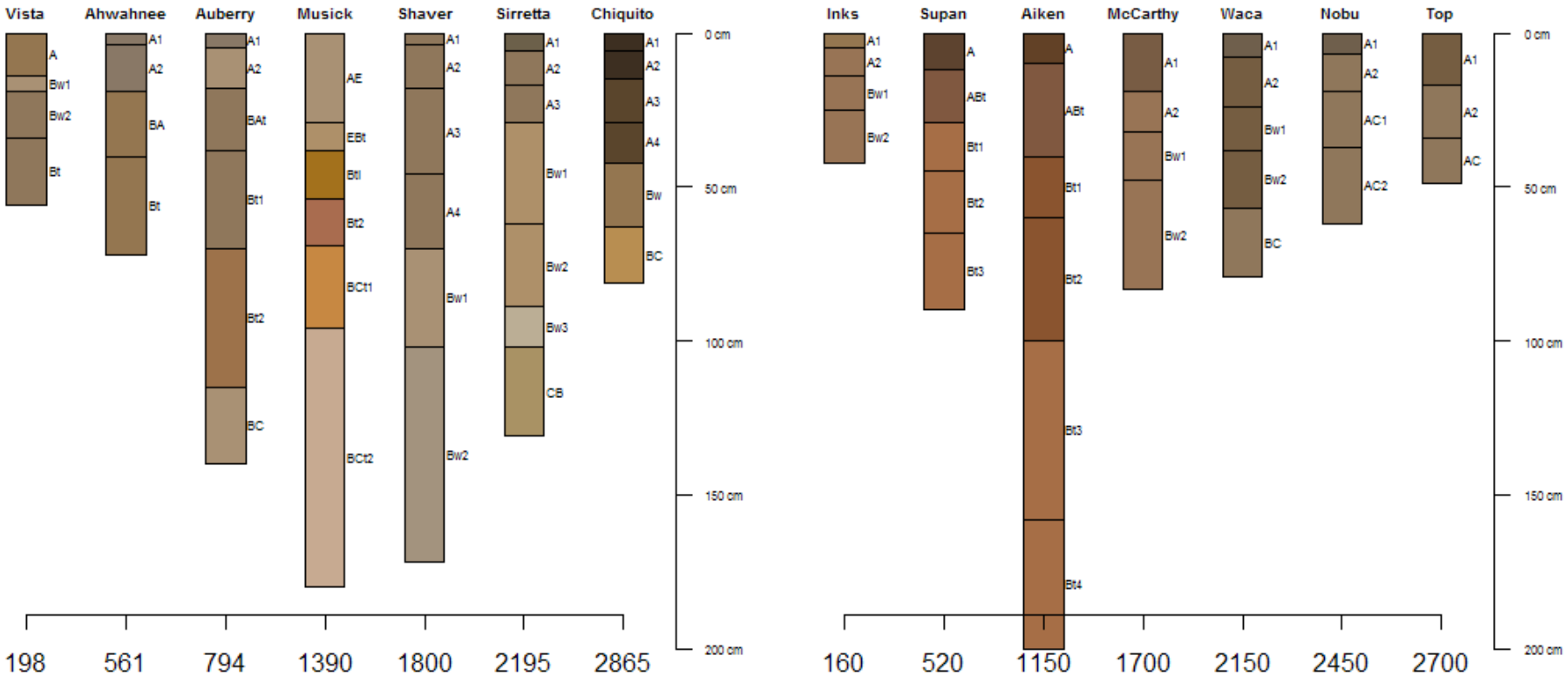
Soil Profile Sketches & Landscape Gradients



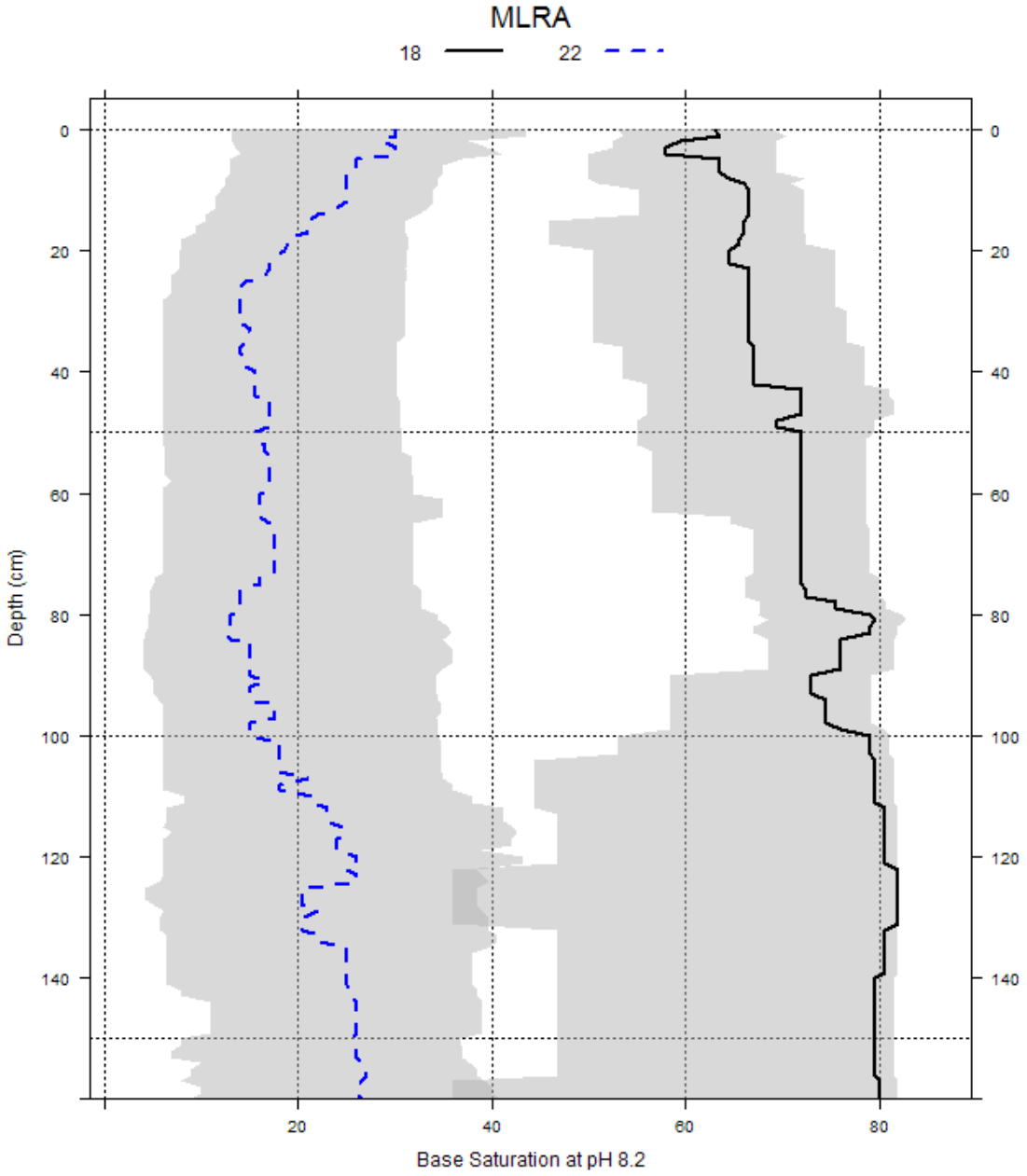
(Digital) Soil Profile Sketches & Landscape Gradients



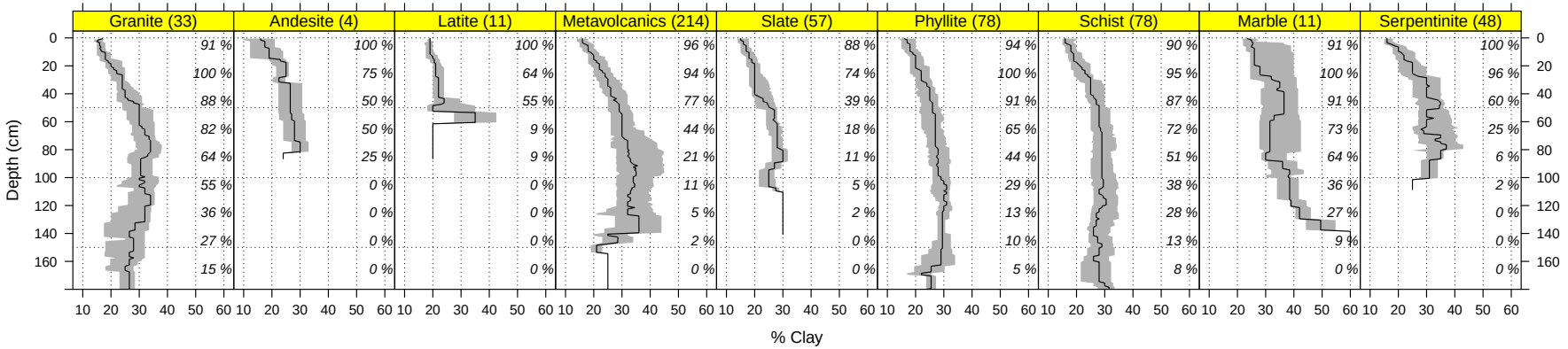
(Digital) Soil Profile Sketches & Climatic Gradients

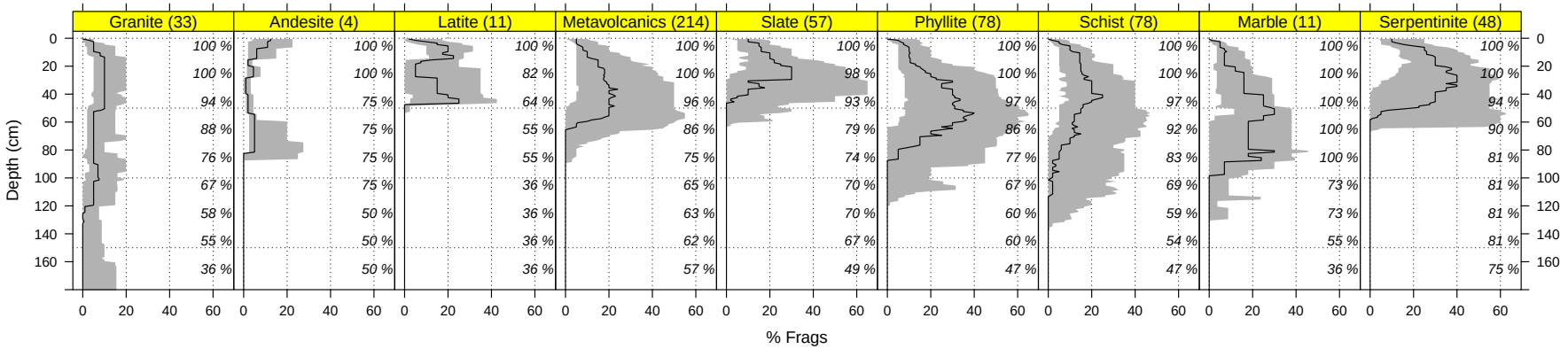


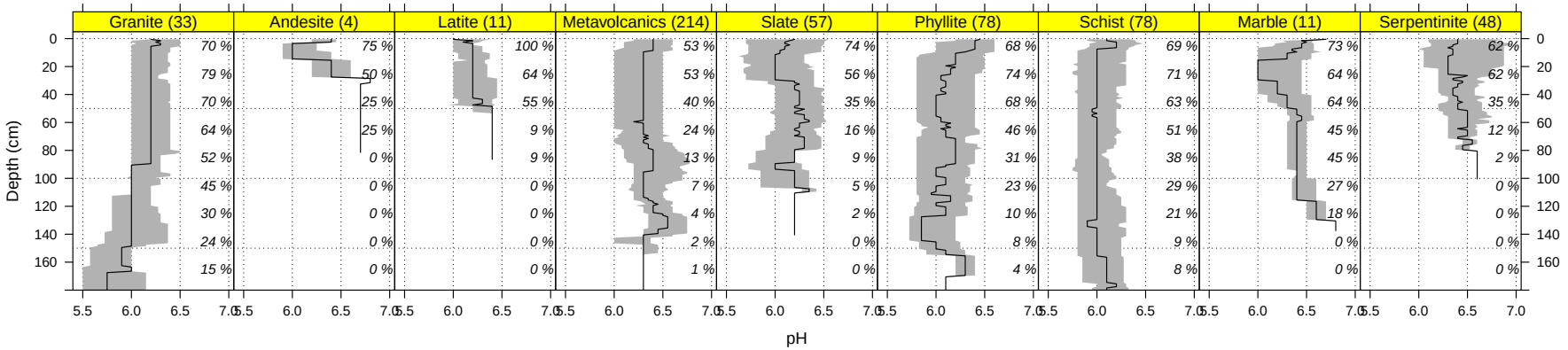
Aggregate Representation of Soil Profile Collections



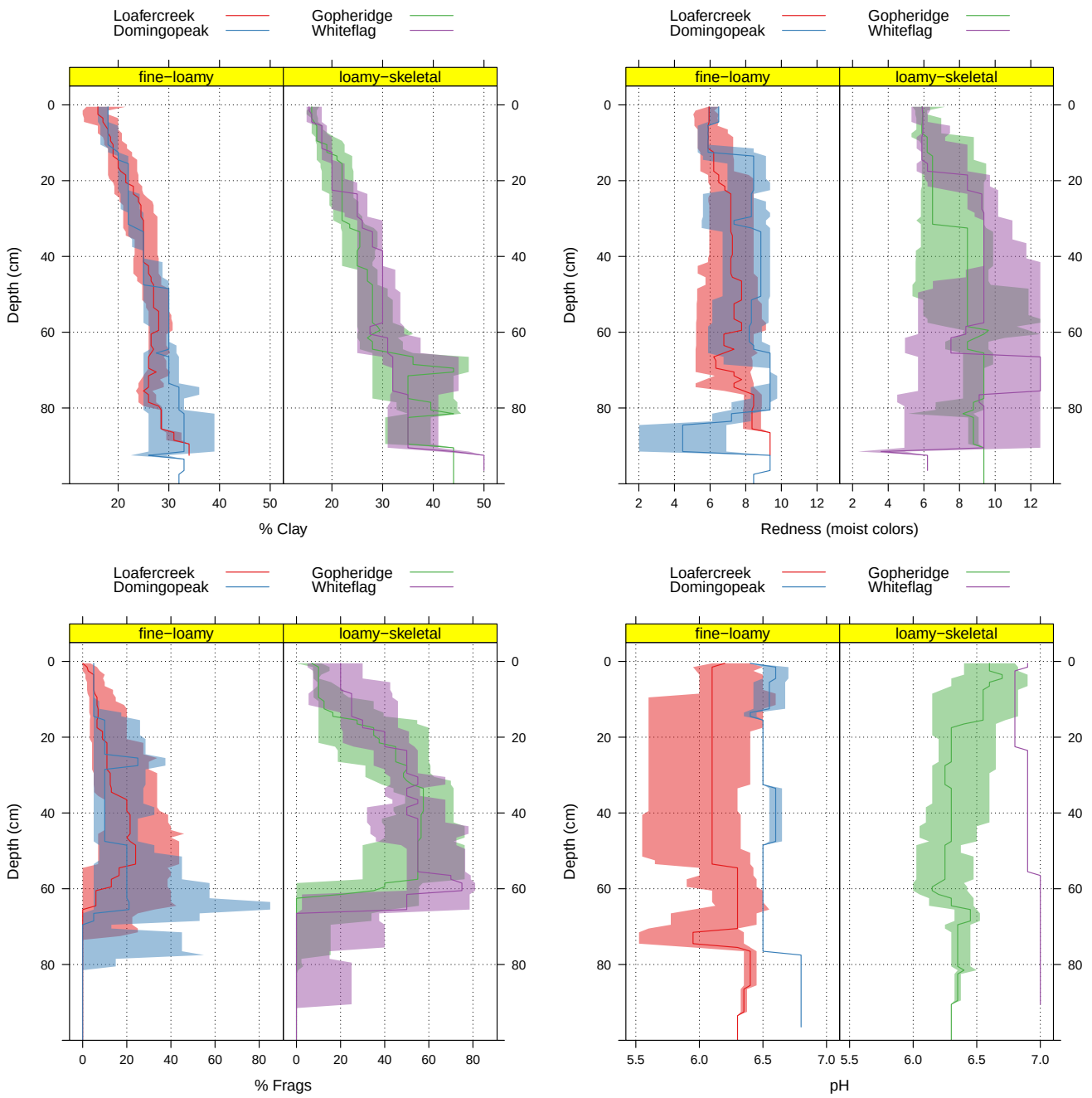
Aggregate Representation of Soil Profile Collections



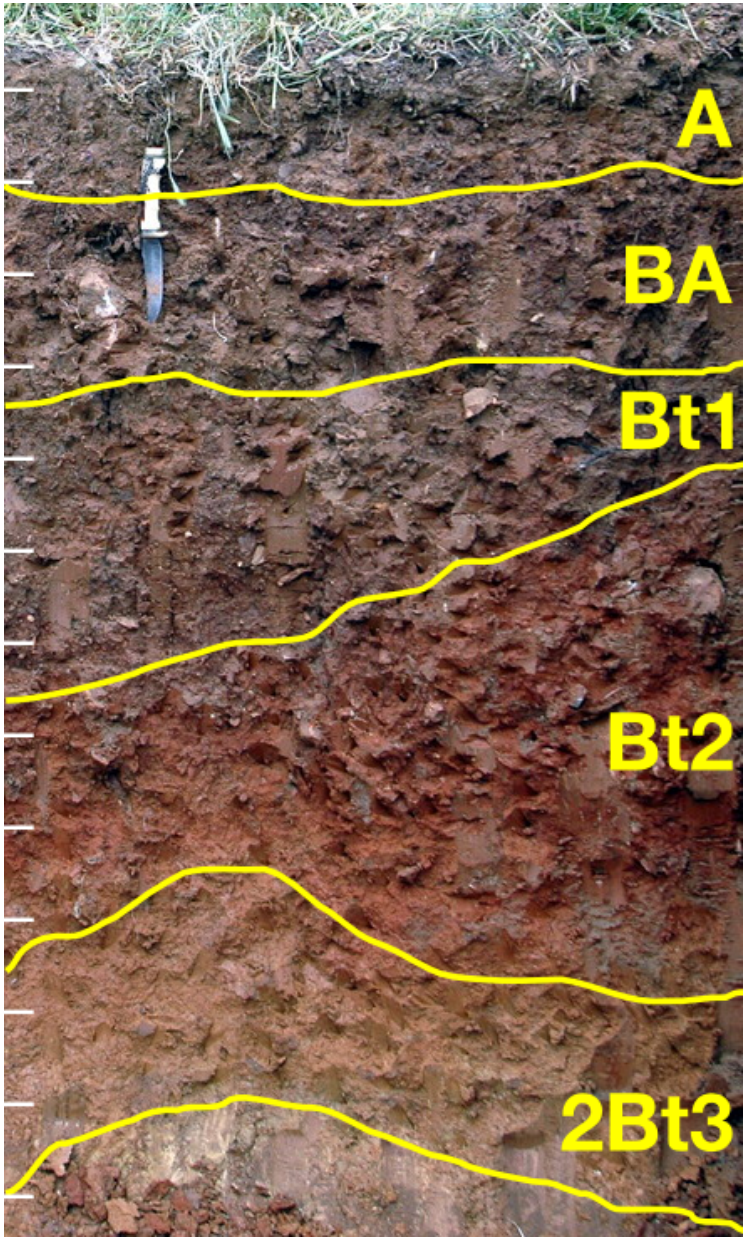


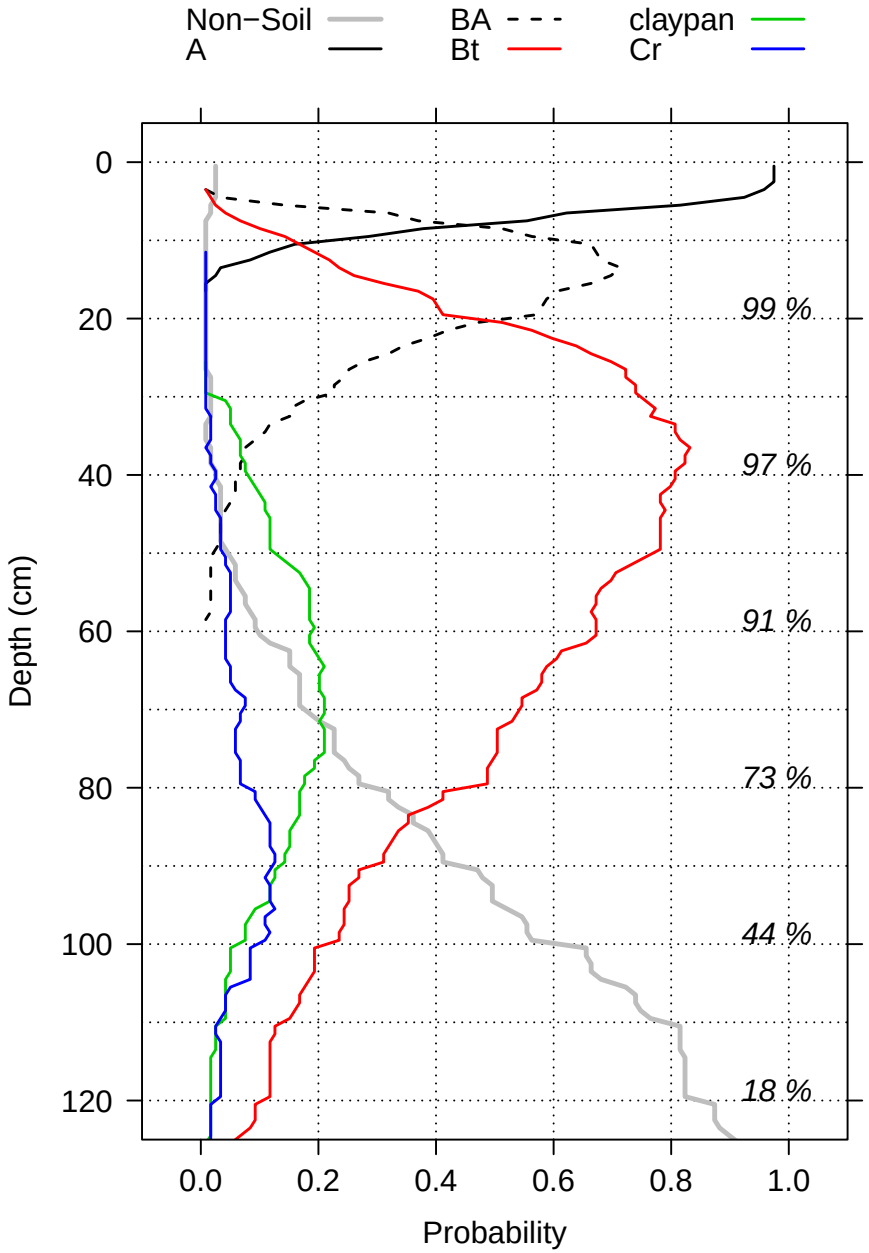


Aggregate Representation of Soil Profile Collections

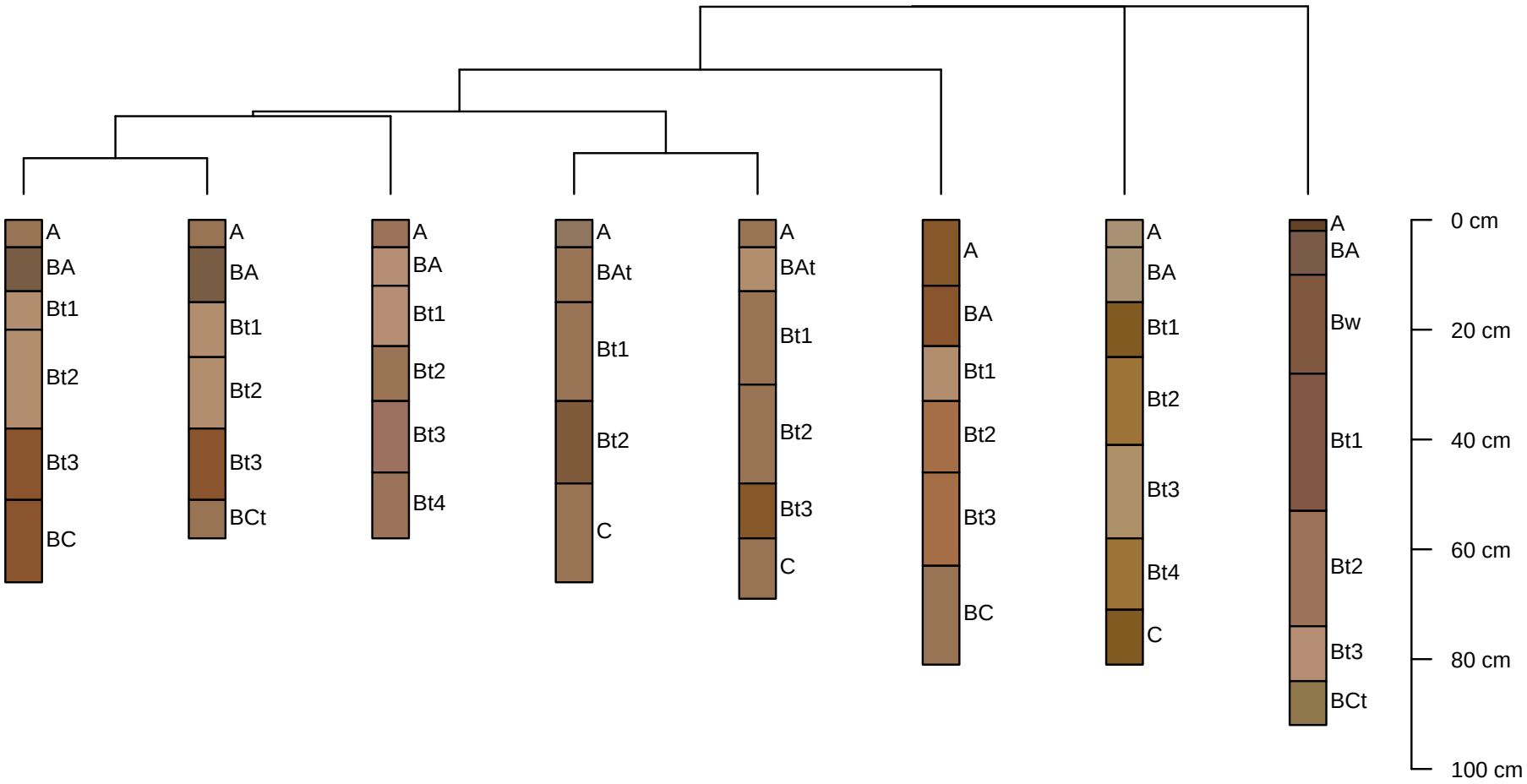


Aggregate Representation of Soil Profile Collections





Simmilar/Dissimilar Evaluation



Domingopeak

Quick Demos

Algorithms for Quantitative Pedology:
`http://aqp.r-forge.r-project.org`

Reports presented by Kevin Godsey

Reports are located in the NSSC Pangaea folder

NSSC Pangaea:!!NASIS QUICK TIPS !!

This report has tips on how to work in NASIS efficiently.

Just run the report. It does not need anything loaded.

NSSC Pangaea:!SDJR DMU Quality Check

This report uses sub reports to make over 250 checks and cross checks of the data for a final check before submitting to the SDQS.

It may fail to run if critical data is missing.

It may timeout if the local data set and the selected set is large.

Do not use to populate the database.

Target Table required: mapunit, component

Check only one DMU at a time, this report can become very large.

Remove any minor soils that are not completely populated.

The Derived values are from the National comparison reports.

Please ****do not copy**** just add to your favorites.

Please help to create a complete tool, notify me of any errors, omission, or just tweaks to make it better. Kevin Godsey---- kevin.godsey@mo.usda.gov

NSSC Pangaea:!NASIS DMU Check list

This is a check list of all the field that need to be populated and the order in which they need to be completed.

Save and edit this list in Word to preserve check marks

NSSC Pangaea:!NASIS SUMMARY OF QUERYING DATA

This is an explanation of querying data.

Just run the report. It does not need anything loaded.

NSSC Pangaea: SDQ - Evaluate correlated non-mlra data

This report makes 140 checks on the data.

It runs on one area symbol at a time.

It counts possible inconsistencies in the data.

Run it against the National database.

NSSC Pangaea: SDQ - Compare Mapunit name with data

Can be run on one area at a time.

Use: Check map unit name against Component data

This report will generate a map unit name from component data.

It then compares the populated map unit name with the generated name.

"a" will print all map units while "d" will only print the names that do not match.

Consociation has one major component.

A Complex has at least two major components.

Names in Map unit match the Component names in the Component table.

In a Consociation, the texture in the map unit name matches the Horizon Texture Group RV.

Component Slope Gradient, low and high, is within the range of slopes in the Map unit name.

NSSC Pangaea: PEDON - Quality check

This report checks pedon and site data tables for missing table or data and inconsistency among the tables. It flags these data and notifies the user of possible inconsistencies.

Load ONLY 1 pedon into the selected set!

NSSC Pangaea: PEDON - Create mini profiles for comparison

The report runs off the selected set.

Load all pedons to be evaluated.

The soil name in the parameter box is used for the OSD page. It does not affect the pedons selected.

The Choice list is r for redox features, f for fragments and t for texture.

The url link below will open the profile report.

You can change the name of the soil (cname) and the choice (t for texture and r for rocks).

https://nasis.sc.egov.usda.gov/NasisReportsWebSite/limsreport.aspx?report_name=WEB-Pedon_profiles&cname=firebaugh&choice=t