

AQP Introduction

D.E. Beaudette and P. Roudier

November 21, 2011

Abstract

Soils are routinely sampled and characterized according to genetic horizons, resulting in data that are associated with principle dimensions: location (x , y), depth (z), and property space (\mathbf{p}). The high dimensionality and grouped nature of this type of data can complicate standard analysis, summarization, and visualization. The “aqp” (algorithms for quantitative pedology) package was designed to support data-driven approaches to common soils-related tasks such as visualization, aggregation, and classification of soil profile collections. In addition, we sought to advance the study of numerical soil classification by building on previously published methods within an extensible and open source framework. Functions in the **aqp** package have been successfully applied to studies involving several thousand soil profiles. The stable version of the **aqp** package is hosted by CRAN (<http://cran.r-project.org/web/packages/aqp>), and the development version is hosted by R-Forge (<http://aqp.r-forge.r-project.org>).

1 Introduction

Soil profiles are usually described, sampled, and characterized by genetic horizons (defined by morphology and usually associated with an inferred process), extending from the surface to a lower boundary determined by bedrock contact or to a depth of 150-200 cm (?). Stratigraphy and morphology of horizons are usually the first data used to infer dominant pedogenic processes within the profile: i.e. degree of alteration relative to the parent material, expression of oxidized or reduced forms of iron, accumulation of organic matter, or evidence of cyclical deposition of new material. While the investigation of soil profile characteristics and horizon-level morphology are strongly based on visual and tactile cues, communication of these data is typically delivered via written narrative or tabular form— complicating integration or further meta-analysis. Soil profile sketches aligned with landscape, parent material, or vegetation gradients are commonly used as the foundation for the development of soil-landscape models, and ultimately take the form of block diagrams in soil survey documents (Figure 2). Despite artistic merits, the reliance on hand-drawn soil profile sketches highlights a missing component in the soil scientist’s digital toolkit. An automated approach to the creation, alignment (e.g. along environmental gradients), and styling of soil profile sketches from digital databases could contribute towards more powerful, data-driven exploration of soil-landscape relationships.

Depiction of basic soil morphology (e.g. horizonation and soil color) is commonly the first step towards understanding patterns in a collection of soil profiles. A simple diagram showing horizon depths, colors, and designations represents an ideal starting point for arranging soils along environmental gradients or for communicating soil variability to non-specialists.

```
> library("aqp")
> data(ca630)
> str(ca630)
```

List of 2

```
$ site:'data.frame':      137 obs. of  11 variables:
 ..$ user_site_id      : chr [1:137] "90CA009101" "90CA009102" "90CA009103" "90CA009104" ...
 ..$ mlra              : chr [1:137] "22" "22" "22" "22" ...
 ..$ county            : chr [1:137] "CA009" "CA009" "CA009" "CA009" ...
 ..$ ssa               : chr [1:137] "CA731" "CA731" "CA731" "CA731" ...
 ..$ lon               : num [1:137] -120 -120 -120 -120 -120 ...
 ..$ lat               : num [1:137] 38.4 38.4 38.4 38.4 38.3 ...
```

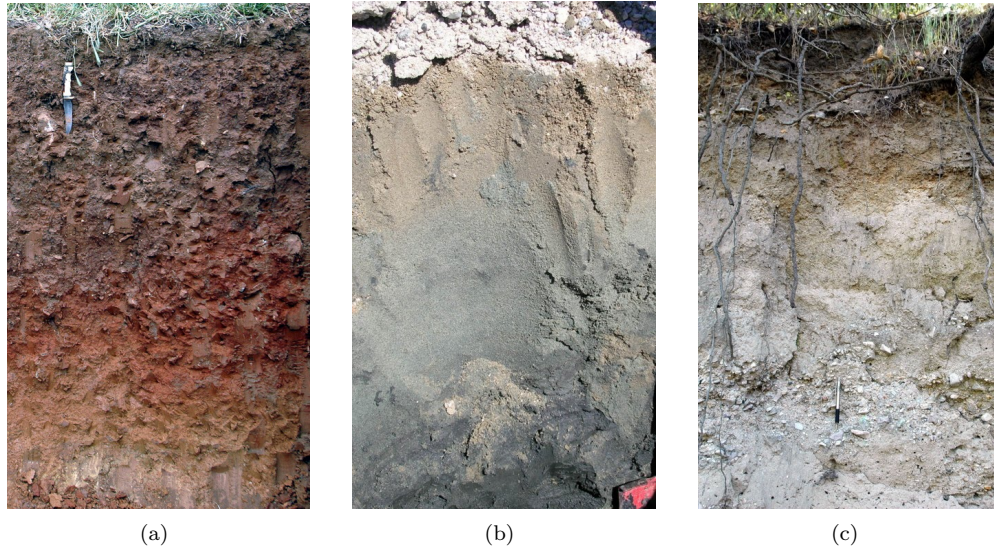


Figure 1: Examples of soil profiles illustrating how horizons change with depth. Color, texture, structure and root abundance are common visual indicators of near surface processes in soil.

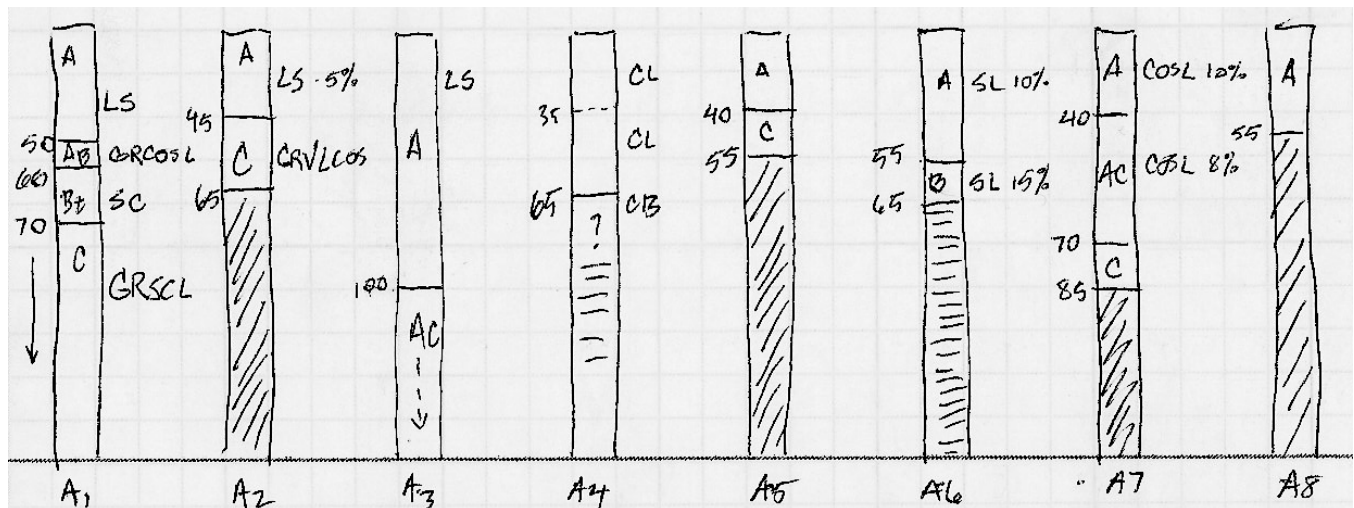


Figure 2: An example of profile sketches manually created from soil profile observations collected as part of the Pinnacles National Monument soil survey. Horizon designations, sequences, boundaries, and soil texture classes are usually sufficient for describing complex soil-landscape relationships.

```

..$ pedon_key      : chr [1:137] "91P0709" "91P0710" "91P0711" "91P0712" ...
..$ user_pedon_id  : chr [1:137] "90CA009101" "90CA009102" "90CA009103" "90CA009104" ...
..$ cntrl_depth_to_top: int [1:137] NA 25 NA 0 25 NA 25 25 0 28 ...
..$ cntrl_depth_to_bot: int [1:137] NA 100 NA 30 100 NA 97 100 30 78 ...
..$ sampled_taxon_name: chr [1:137] "Windy" "Tallac" "Gerle" "Nd" ...
$ lab : 'data.frame':      571 obs. of  13 variables:
..$ pedon_key      : chr [1:571] "91P0709" "91P0709" "91P0709" "91P0709" ...
..$ layer_key       : chr [1:571] "91P04048" "91P04049" "91P04050" "91P04051" ...
..$ layer_sequence  : int [1:571] 1 2 3 4 1 2 3 4 1 2 ...
..$ hzn_top         : int [1:571] 0 30 79 119 0 25 46 71 0 25 ...
..$ hzn_bot         : int [1:571] 30 79 119 218 25 46 71 107 25 58 ...

```

```

..$ hzn_desgn      : chr [1:571] "A" "Bw1" "Bw2" "C" ...
..$ texture_description: chr [1:571] "Sandy loam" "Sandy loam" "Fine sandy loam" "Fine sandy loam" ...
..$ nh4_sum_bases   : num [1:571] 9.9 1.3 0.9 0.8 10.1 1.4 0.8 0.2 1.7 0.5 ...
..$ ex_acid         : num [1:571] 21.9 18.8 13.7 12.2 11.3 18.6 18.6 12 11.6 9.3 ...
..$ CEC8.2          : num [1:571] 31.8 20.1 14.6 13 21.4 20 19.4 12.2 13.3 9.8 ...
..$ CEC7            : num [1:571] 25.3 13.2 10.8 9 16.7 14.1 11.6 6.5 14.9 9.7 ...
..$ bs_8.2          : int [1:571] 31 6 6 6 47 7 4 2 13 5 ...
..$ bs_7            : int [1:571] 39 10 8 9 60 10 7 3 11 5 ...

> # combine site+horizon data into single DF
> ca <- join(ca630$lab, ca630$site, type='inner')
> # promote to SoilProfileCollection
> depths(ca) <- pedon_key ~ hzn_top + hzn_bot
> # extract site data
> site(ca) <- ~ mlra + ssa + lon + lat + cntrl_depth_to_top + cntrl_depth_to_bot + sampled_taxon_name
> # extract spatial data as SpatialPoints
> coordinates(ca) <- ~ lon + lat
> # assign CRS data
> proj4string(ca) <- '+proj=latlong +datum=NAD83'
> # add new fake grouping column
> ca@site$group <- 1 # messy syntax ... will be fixed in AQP 1.0
> # aggregate CEC 7 for all profiles into 1 cm slices
> a <- soil.slot.multiple(ca, fm=group ~ bs_7)

|
|
|
|=====| 100%

> # 4. plot median & IQR by 1 cm slice
> p1 <- xyplot(
+ top ~ p.q50, data=a, lower=a$p.q25, upper=a$p.q75,
+ ylim=c(160,-5), alpha=0.5, scales=list(alternating=1, y=list(tick.num=7)),
+ panel=panel.depth_function, prepanel=prepanel.depth_function,
+ ylab='Depth (cm)', xlab='Base Saturation at pH 7', par.settings=list(superpose.line=list(col='black'))
+ )
> # 5. safely compute hz-thickness weighted mean CEC (pH 7)
> head(lab.agg.cec_7 <- ddply(ca630$lab, .(pedon_key),
+ .fun=summarise, CEC_7=wtd.mean(bs_7, weights=hzn_bot-hzn_top)))

  pedon_key    CEC_7
1  00P0780 13.888889
2  00P0781 11.191489
3  00P0783 19.254545
4  00P0784 29.825000
5  00P0785  9.917722
6  00P0786 11.514286

> # 6. extract a SPDF with horizon data along a slice at 25 cm
> s.25 <- slice(ca, fm=55 ~ bs_7 + CEC7 + ex_acid)

result is a SpatialPointsDataFrame object

> # 6.1 extract a data.frame with horizon data along a slices c(10,20,50)
> s.multiple <- slice(ca, fm=c(10,20,50) ~ bs_7 + CEC7 + ex_acid)

```

result is a SoilProfileCollection object

```
> # 7. Extract the 2nd horizon from all profiles as SPDF  
> ca.2 <- ca[, 2]
```

result is a SpatialPointsDataFrame object

```
> # 8. Extract the all horizons from profiles 1:10 as data.frame  
> ca.1.to.10 <- ca[1:10, ]
```

result is a data.frame object

...

plotting lines...

