Supplement of SOIL, 9, 277–300, 2023
https://doi.org/10.5194/soil-9-277-2023-supplement

Supplement of

Accuracy of regional-to-global soil maps for on-farm decision-making: are soil maps "good enough"?

Jonathan J. Maynard et al.

Correspondence to: Jonathan J. Maynard (jonathan.maynard@usda.gov)

Soil texture class transforms

This R notebook provides the code used to transform the soil textural class data presented in the following two studies from their national textural class systems (ASNIE and AU) to the USDA system.

Richer-de-Forges, A. C., Arrouays, D., Chen, S., Dobarco, M. R., Libohova, Z., Roudier, P., … & Bourennane, H. (2022). Hand-feel soil texture and particle-size distribution in central France. Relationships and implications. Catena, 213, 106155.

Minasny, B., McBratney, A. B., Field, D. J., Tranter, G., McKenzie, N. J., & Brough, D. M. (2007). Relationships between field texture and particle-size distribution in Australia and their implications. Soil Research, 45(6), 428-437.

Load packages

```
library(dplyr)

library(soiltexture)
```

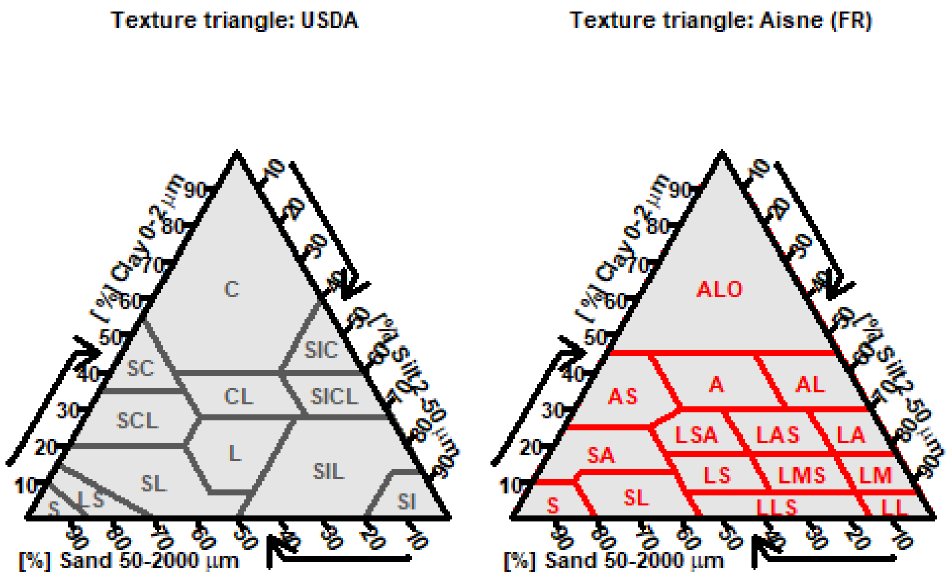
Create point grid within soil textural triangle. This creates a point for every possible particle size class combination (e.g., [sand=22,silt=10,clay=68])

```
sand <- seq(0, 100, 1)
clay <- seq(0, 100, 1)
txt <-  expand.grid(x = sand, y = clay)
txt <- txt %>% rowwise() %>% mutate(silt = (100 - (x + y))) %>% ungroup() %>%
purrr::set_names("SAND", "CLAY", "SILT") %>% filter(SILT >= 0)
```

Transform AISNE textural class data to USDA textural class data

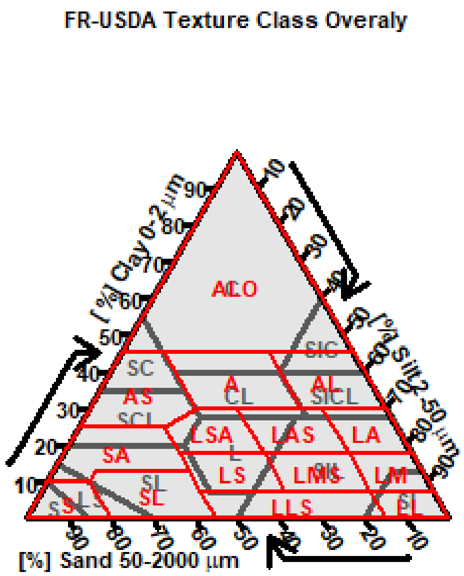
Plot Aisne and USDA texture triangles

```
old.par <- par(no.readonly=T)
par("mfcol" = c(1,2),"mfrow"=c(1,2))
TT.plot(class.sys = "USDA-NCSS.TT",main = "Texture triangle: USDA", cex=0.5,c
ex.axis = 0.7, cex.lab = 0.7, cex.main = 0.7, grid.show=F)
TT.plot(class.sys = "FR.AISNE.TT",main = "Texture triangle: Aisne (FR)", cex=
0.5,cex.axis = 0.7, cex.lab = 0.7, cex.main = 0.7, grid.show=F, class.line.co
l = "red", class.lab.col = "red")
```



Overlay AISNE and USDA

```
geo <- TT.plot(class.sys = "USDA-NCSS.TT",main = "FR-USDA Texture Class Overa
ly", cex=0.5,cex.axis = 0.7, cex.lab = 0.7,  cex.main = 0.7, grid.show=F)
TT.classes(
geo = geo,
class.sys = "FR.AISNE.TT",
# Additional "graphical" options
class.line.col = "red",
class.lab.col = "red",
lwd.axis = 2, cex.lab = 0.7
)
```



Assign each point to its corresponding AISNE class and USDA class

```
AISNE <-  txt %>% rowwise() %>% mutate(txt = TT.points.in.classes(
tri.data = data.frame(SAND, CLAY, SILT),
class.sys = "FR.AISNE.TT", PiC.type = "t", collapse=",")) %>% ungroup()
#extract only first texture label
AISNE <-  AISNE %>% rowwise() %>% mutate(AISNE = strsplit(txt, "[,]")[[1]][1]
) %>% ungroup() %>% as.data.frame()
AISNE <-  AISNE %>% rowwise() %>% mutate(txt = TT.points.in.classes(
tri.data = data.frame(SAND, CLAY, SILT),
class.sys = "USDA-NCSS.TT", PiC.type = "t", collapse=",")) %>% ungroup()
#extract only first texture label
AISNE <-  AISNE %>% rowwise() %>% mutate(USDA = strsplit(txt, "[,]")[[1]][1])
%>% ungroup() %>% as.data.frame()
AISNE <-  AISNE %>% dplyr::select(-c(txt))
```

Load in class accuracies (e.g., Producer's accuracy [PA]) for each AISNE class from Richer-de-Forges et al. (2022), join to our texture grid, and summarize by USDA texture class

```
AISNE_PA <- data.frame(c("A","AL","ALO","AS","LA","LAS","LL","LLS","LM","LMS"
,"LS","LSA","S","SA","SL"),
c(69, 64, 83, 74, 76, 66, 100, 30, 84, 72, 54, 64, 95, 73, 70)) %>% purrr::se
t_names("AISNE", "PA")
AISNE_PA_adj <- data.frame(c("A","AL","ALO","AS","LA","LAS","LL","LLS","LM","
LMS","LS","LSA","S","SA","SL"),
c(96, 97, 99, 98, 94, 94, 100, 100, 97, 95, 92, 91, 100, 98, 99)) %>% purrr::
set_names("AISNE", "PA_adj")
```

```r
AISNE <-  AISNE %>% left_join(AISNE_PA, by="AISNE") %>% left_join(AISNE_PA_adj, by="AISNE")

# This effectively gives us an area weighed average for each USDA texture class based on the relative area of each AISNE texture class that intersects USDA texture classes
USDA_AISNE_txt_class_PA <- AISNE %>% group_by(USDA) %>% summarise(PA = mean(PA) %>% round(digits = 0))
USDA_AISNE_txt_class_PA_adj <- AISNE %>% group_by(USDA) %>% summarise(PA_adj = mean(PA_adj) %>% round(digits = 0))
```

Aproximate the number of AISNE texture class samples that would fall into USDA texture classes

```r
# Calculates the area of each AISNE texture class that falls within a USDA texture class
USDA_AISNE_class_count <- AISNE %>% group_by(USDA) %>% count(AISNE)

# Calculate the gridded area of the textural triangle occupied by each AISNE texture class
AISNE_class_count <- AISNE %>% group_by(AISNE) %>% count(AISNE) %>% ungroup() %>% purrr::set_names("AISNE", "txt_tri_area")

# Add the number of samples analyzed for each AISNE texture class
AISNE_class_count$txt_samples <- c(1723, 2270, 2487, 870, 1723, 1331, 3, 10, 346, 532, 815, 1315, 1210, 1561, 1192)

# Join table of AISNE texture class area and corresponding sample numbers per class
USDA_AISNE_class_count <- USDA_AISNE_class_count %>% left_join(AISNE_class_count, by="AISNE")

# For each intersection area of an AISNE class that falls within a USDA texture class, calculate the relative proportion of samples from that AISNE class based on its relative area.

#For example, the USDA CL-Clay Loam texture class intersects three AISNE textural classes (A-clay, AL-silty clay, and ALO-heavy clay). ~32% of the AISNE clay class falls within the USDA clay loam class and there were 1723 clay samples in the Richer-de-Forges et al. 2022 study. Thus we attribute 556 of these samples (32%) to the USDA clay loam class. This procedure is applied to each intersecting AISNE class within each USDA texture class.

USDA_AISNE_class_count <- USDA_AISNE_class_count %>% rowwise() %>% mutate(prop_samp = (n/txt_tri_area)*txt_samples) %>% ungroup()

USDA_AISNE_samp_count <- USDA_AISNE_class_count  %>% group_by(USDA) %>% summarise(USDA_samp = sum(prop_samp) %>% round(digits = 0))  %>% ungroup()
```

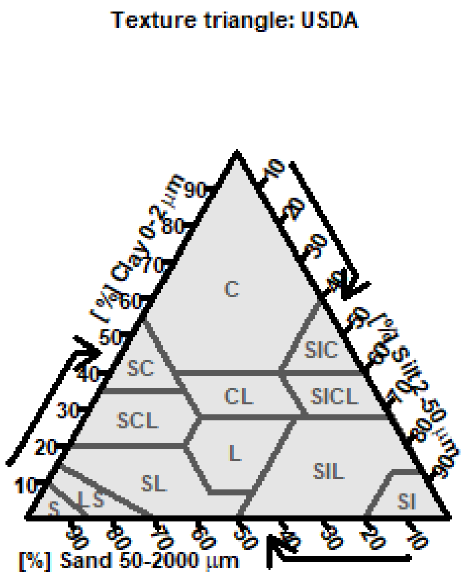
```r
#combine PA, PAadj, and USDA_samp

USDA_AISNE_summary <- USDA_AISNE_samp_count %>% left_join(USDA_AISNE_txt_clas
s_PA, by="USDA") %>% left_join(USDA_AISNE_txt_class_PA_adj, by="USDA") %>% pu
rrr::set_names("USDA", 'n', "PA (%)", "PA-adj (%)")

knitr::kable(USDA_AISNE_summary, align = "lccrr")
```

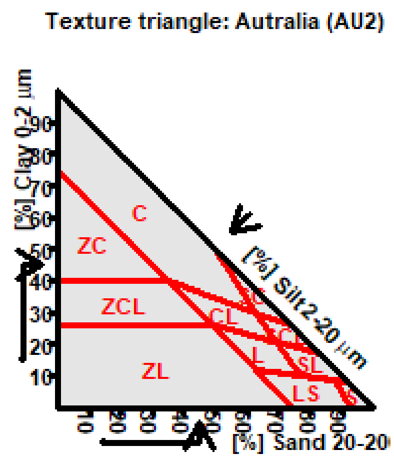
USDA	n	PA (%)	PA-adj (%)
C	2829	82	99
CL	1387	68	95
L	1985	61	92
LS	876	82	99
S	664	95	100
SC	493	76	98
SCL	1336	73	97
SI	111	92	99
SIC	868	75	98
SICL	1853	66	96
SIL	2946	61	96
SL	2039	66	98

Transform Australian (AU) textural class data to USDA textural class data

Plot Australian and USDA texture triangles

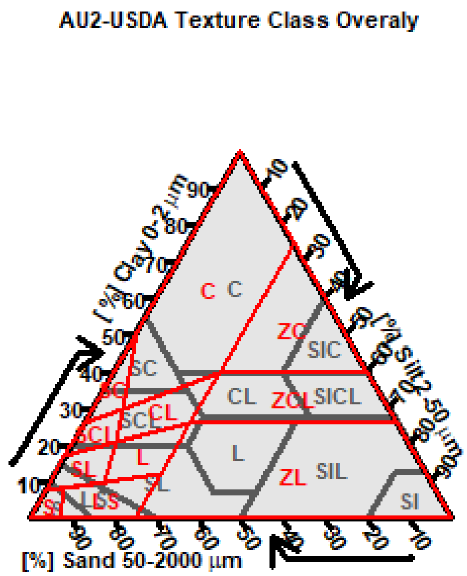
```r
old.par <- par(no.readonly=T)
par("mfcol" = c(1,2),"mfrow"=c(1,2))
TT.plot(class.sys = "USDA-NCSS.TT",main = "Texture triangle: USDA", cex=0.5,c
ex.axis = 0.7, cex.lab = 0.7,  cex.main = 0.7, grid.show=F)
TT.plot(class.sys = "AU2.TT",main = "Texture triangle: Autralia (AU2)", cex=0
.5,cex.axis = 0.7, cex.lab = 0.7,  cex.main = 0.7, grid.show=F, class.line.co
l = "red", class.lab.col = "red")
```





Overlay AU and USDA

```
geo <- TT.plot(class.sys = "USDA-NCSS.TT",main = "AU2-USDA Texture Class Overaly", cex=0.5,cex.axis = 0.7, cex.lab = 0.7, cex.main = 0.7,grid.show=F)
TT.classes(
geo = geo,
class.sys = "AU2.TT",
# Additional "graphical" options
class.line.col = "red",
class.lab.col = "red",
lwd.axis = 2, cex.lab = 0.7
)
```



Assign each point to its corresponding Australian class and USDA class

```
AU <-  txt %>% rowwise() %>% mutate(txt = TT.points.in.classes(
tri.data = data.frame(SAND, CLAY, SILT),
class.sys = "AU2.TT", PiC.type = "t", collapse=",")) %>% ungroup()
#extract only first texture label
AU <-  AU %>% rowwise() %>% mutate(AU = strsplit(txt, "[,]")[[1]][1]) %>% ung
roup() %>% as.data.frame()
AU <-  AU %>% rowwise() %>% mutate(txt = TT.points.in.classes(
tri.data = data.frame(SAND, CLAY, SILT),
class.sys = "USDA-NCSS.TT", PiC.type = "t", collapse=",")) %>% ungroup()
#extract only first texture label
AU <-  AU %>% rowwise() %>% mutate(USDA = strsplit(txt, "[,]")[[1]][1]) %>% u
ngroup() %>% as.data.frame()
AU <-  AU %>% dplyr::select(-c(txt))
```

Load in Producer accuracies for each Australian class from Minasny et al., 2007, join to our texture grid, and summarize by USDA texture class

```
AU_PA <- data.frame(c('S', 'LS', 'SL', 'L', 'ZL', 'SCL', 'CL', 'ZCL', 'SC', '
ZC', 'C'),
c(78, 32, 40, 35, 14, 37, 15, 6, 40, 6, 86)) %>% purrr::set_names("AU", "PA")
AU_PA_adj <- data.frame(c('S', 'LS', 'SL', 'L', 'ZL', 'SCL', 'CL', 'ZCL', 'SC
', 'ZC', 'C'),
c(94, 94, 92, 63, 55, 77, 66, 24, 86, 84, 96)) %>% purrr::set_names("AU", "PA
_adj")
AU <-  AU %>% left_join(AU_PA, by="AU") %>% left_join(AU_PA_adj, by="AU")
```

```
# This effectively gives us an area weighed average for each USDA texture cla
ss based on the relative area of each Australian texture class that intersect
s USDA texture classes
USDA_AU_txt_class_PA <- AU %>% group_by(USDA) %>% summarise(PA = mean(PA) %>%
round(digits = 0))
USDA_AU_txt_class_PA_adj <- AU %>% group_by(USDA) %>% summarise(PA = mean(PA_
adj) %>% round(digits = 0))
```

Aproximate the number of Australian texture class samples that would fall into USDA texture classes

```
# Calculates the area of each AU texture class that falls within a USDA textu
re class
USDA_AU_class_count <- AU %>% group_by(USDA) %>% count(AU)

# Calculate the gridded area of the textural triangle occupied by each AU tex
ture class
AU_class_count <- AU %>% group_by(AU) %>% count(AU) %>% ungroup() %>% purrr::
set_names("AU", "txt_tri_area")

# Add the number of samples analyzed for each AU texture class
AU_class_count$txt_samples <- c(1278, 1060, 1748, 1339, 557, 858, 1745, 635,
456, 704, 7599)


# Join table of AU texture class area and corresponding sample numbers per cl
ass
USDA_AU_class_count <- USDA_AU_class_count %>% left_join(AU_class_count, by="
AU")

# For each intersection area of an AU class that falls within a USDA texture
class, calculate the relative proportion of samples from that AU class based
on its relative area.

#For example, the USDA C-Clay Loam texture class intersects two AU textural c
lasses (C-clay and ZC-silty clay). ~66% of the AU silty clay class falls with
in the USDA clay class and there were 456 silty clay samples in the Minasney
et al. 2007 study. Thus we attribute 304 of these samples (66%) to the USDA c
lay class. This procedure is applied to each intersecting AU class within eac
h USDA texture class.

USDA_AU_class_count <- USDA_AU_class_count %>% rowwise() %>% mutate(prop_samp
= (n/txt_tri_area)*txt_samples) %>% ungroup()

USDA_AU_samp_count <- USDA_AU_class_count  %>% group_by(USDA) %>% summarise(U
SDA_samp = sum(prop_samp) %>% round(digits = 0))  %>% ungroup()

USDA_AU_summary <- USDA_AU_samp_count %>% left_join(USDA_AU_txt_class_PA, by=
```

```
"USDA") %>% left_join(USDA_AU_txt_class_PA_adj, by="USDA") %>% purrr::set_nam
es("USDA", 'n', "PA (%)", "PA-adj (%)")

knitr::kable(USDA_AU_summary, align = "lccrr")
```

USDA	n	PA (%)	PA-adj (%)
C	1388	64	93
CL	451	10	30
L	1600	13	51
LS	957	33	91
S	726	62	94
SC	662	73	93
SCL	3609	29	70
SI	838	14	55
SIC	152	6	84
SICL	277	6	24
SIL	3741	14	53
SL	3577	24	67

Calculate estimated field texture class accuracies (±SD) from global meta-analysis

```
library(Hmisc)

exact_match <- read.csv('C:/R_Drive/Data_Files/LPKS_Data/Projects/Ghana/Soil
Map Variability/Hand_Txt_Lit_Summary_Exact_Match.csv')
adj_match <- read.csv('C:/R_Drive/Data_Files/LPKS_Data/Projects/Ghana/Soil Ma
p Variability/Hand_Txt_Lit_Summary_Adj_Match.csv')

txt <- c('S', 'LS', 'SL', 'SC', 'SCL', 'L', 'SiL', 'CL', 'SiCL', 'SiC', 'C',
'Si')
txt_num <- c('S_n', 'LS_n', 'SL_n', 'SC_n', 'SCL_n', 'L_n', 'SiL_n','CL_n', '
SiCL_n', 'SiC_n', 'C_n', 'Si_n')


mean_exact <- list()
std_exact <- list()

for(i in 1:length(txt)){
  mean_exact[i] <- wtd.mean(exact_match[txt[i]], exact_match[txt_num[i]], na.
rm=T)
  std_exact[i] <- base::sqrt(wtd.var(exact_match[txt[i]], exact_match[txt_num
[i]], na.rm=T))
}

exact_match_summary <- data.frame(txt, unlist(mean_exact) %>% round(digits=0)
```

```
, unlist(std_exact) %>% round(digits=0)) %>% purrr::set_names("USDA", "PA" , "std")

mean_adj <- list()
std_adj <- list()

for(i in 1:length(txt)){
  mean_adj[i] <- wtd.mean(adj_match[txt[i]], adj_match[txt_num[i]], na.rm=T)
  std_adj[i] <- base::sqrt(wtd.var(adj_match[txt[i]], adj_match[txt_num[i]], na.rm=T))
}

adj_match_summary <- data.frame(txt, unlist(mean_adj) %>% round(digits=0), unlist(std_adj) %>% round(digits=0)) %>% purrr::set_names("USDA", "PA_adj" , "std_adj")

match_summary <- exact_match_summary %>% left_join(adj_match_summary, by="USDA")

knitr::kable(match_summary, align = "lccrr")
```

USDA	PA	std	PA_adj	std_adj
S	73	7	89	3
LS	45	12	94	3
SL	57	11	91	8
SC	28	22	71	21
SCL	41	19	81	18
L	54	10	91	8
SiL	79	8	98	6
CL	57	7	86	8
SiCL	69	4	95	3
SiC	59	7	93	2
C	74	4	93	2
Si	19	16	80	19

```
ghana_txt_freq <- data.frame(c("C", "CL","L","LS","S","SC","SCL","SL","SiL","SiC","SiCL"),c(107, 1283, 1274, 9213, 10605, 440, 2380, 6205, 1226, 524, 1157), c(0.003100461, 0.037176552, 0.036915766, 0.266958361, 0.307293327, 0.012749558, 0.068963519, 0.179797746, 0.035524905, 0.015183565, 0.033525543)) %>% purrr::set_names("USDA", "n", "prop")

#calculate weighted mean accuracy based on sample texture distribution
match_summary_ghana <- match_summary %>% left_join(ghana_txt_freq, by="USDA")
match_summary_ghana <- match_summary_ghana %>% mutate(ghana_PA = PA*prop) %>%
```

```
mutate(ghana_PA_adj = PA_adj*prop)  %>%
                        mutate(ghana_H_PA = (PA+std)*prop) %>% mutate(ghana_H
_PA_adj = (PA_adj+std_adj)*prop) %>%
                        mutate(ghana_L_PA = (PA-std)*prop) %>% mutate(ghana_L
_PA_adj = (PA_adj-std_adj)*prop)

sum(match_summary_ghana$ghana_PA, na.rm=T)

## [1] 58.23601

sum(match_summary_ghana$ghana_PA_adj, na.rm=T)

## [1] 90.21935

sum(match_summary_ghana$ghana_H_PA, na.rm=T)

## [1] 68.32552

sum(match_summary_ghana$ghana_H_PA_adj, na.rm=T)

## [1] 95.8326

sum(match_summary_ghana$ghana_L_PA, na.rm=T)

## [1] 48.1465

sum(match_summary_ghana$ghana_L_PA_adj, na.rm=T)

## [1] 84.6061

ghana_txt_freq_fg <- data.frame(c('C', 'CL', 'L', 'LS', 'S', 'SC', 'SCL', 'SL
', 'SiL', 'SiCL'), c(1, 254, 22, 11, 35, 1, 223, 102, 1, 22), c(0.001488095,
0.377976190, 0.032738095, 0.016369048, 0.052083333, 0.001488095, 0.331845238,
0.151785714, 0.001488095, 0.032738095)) %>% purrr::set_names("USDA", "n", "pr
op")

#calculate weighted mean accuracy based on sample texture distribution
match_summary_ghana_fg <- match_summary %>% left_join(ghana_txt_freq_fg, by="
USDA")
match_summary_ghana_fg <- match_summary_ghana_fg %>% mutate(ghana_PA = PA*pro
p) %>% mutate(ghana_PA_adj = PA_adj*prop)  %>%
                        mutate(ghana_H_PA = (PA+std)*prop) %>% mutate(ghana_H
_PA_adj = (PA_adj+std_adj)*prop) %>%
                        mutate(ghana_L_PA = (PA-std)*prop) %>% mutate(ghana_L
_PA_adj = (PA_adj-std_adj)*prop)

sum(match_summary_ghana_fg$ghana_PA, na.rm=T)

## [1] 52.6369

sum(match_summary_ghana_fg$ghana_PA_adj, na.rm=T)

## [1] 85.85119
```

```r
sum(match_summary_ghana_fg$ghana_H_PA, na.rm=T)

## [1] 64.32738

sum(match_summary_ghana_fg$ghana_H_PA_adj, na.rm=T)

## [1] 96.67113

sum(match_summary_ghana_fg$ghana_L_PA, na.rm=T)

## [1] 40.94643

sum(match_summary_ghana_fg$ghana_L_PA_adj, na.rm=T)

## [1] 75.03125
```