# NatureServe and NRCS regional collaboration: Data and technology transfer to enhance & Support Ecological Site Habitat and Wildlife Management

Agreement No. 69-3A75-12-181

Project Director: Shannon Menard

Period Covered by Report: 10/01/2012 - 09/30/2015

Report Delivered 12/28/2015

# Contents

Executive Summary ..... 1
Introduction ..... 2
Background ..... 6
Review of Methods and Quality Assurance ..... 8
Findings ..... 10
Conclusions and Recommendations ..... 13
Appendices ..... 15
Appendix A. References: ..... 15
Appendix B. Detailed Results of Comparison of Soil Variables to Presence or Absence of Ecological Systems in MLRA 105. ..... 16

## Executive Summary

This project was initiated with three objectives in mind. In partnership with NRCS staff, we wanted to

1. Use vegetation classification information developed and maintained by NatureServe to contribute to the development, and eventual use, of Ecological Site Descriptions (ESDs) in the upper Midwest,
2. Adapt and expand a database relating threatened and endangered wildlife to habitats using NatureServe's Ecological Systems classification, and
3. Provide this information to potential users so they had the opportunity to incorporate it in their decision-making process.

Toward these ends, we worked with NRCS staff who were developing ESDs in two pilot Major Land Resource Areas (MLRA) in the upper Midwest. These two MLRAs were MLRA 93A and MLRA 105. In the early stages of ESD development, we provided information on the US National Vegetation Classification (USNVC) and Ecological Systems classification units that were likely to be most similar to the Ecological Sites in question and we provided input on the ESDs, including State-and-Transition models, to strengthen the ESDs and keep the link between these classification units as clear as possible. A well-defined and unambiguous relationship between ESDs and the USNVC and Ecological Systems allows use of the extensive information already in the latter vegetation classifications as well as providing a bridge to other classifications that are linked to the USNVC.

Our second objective did not meet with as much success. We found the modifications to the existing database and the time required to determine which Ecological Systems would provide habitat for all of the wildlife species to be beyond the scope of this project. We adjusted the objectives to examine another possible route to aid in the development of ESDs and linking them to the USNVC and Ecological Systems. Using 776 previously acquired field vegetation plots, we examined the relationship between 46 soil characteristics from SSURGO data and Ecological Systems. Statistically significant relationships between these vegetation types and soils and other environmental variables indicate possible starting points in the consideration of new Ecological Sites.

The third objective was modified to take into account the results from the first two. Our potential user group had changed from both staff developing Ecological Sites and landowners wishing to know potential wildlife on their property to just staff developing Ecological Sites. We presented results of our work to this group at 2015 National Cooperative Soil Survey National Conference in Duluth, MN, June 2015.

The project was initially scoped to run from October 2012 to March 2014. Due to slower-than-initially-expected receipt of ESDs from NRCS to review, to difficulties attempting to develop the Wildlife-Habitat database, and changes in NRCS priorities, we received an extension to September 2015. Funding was spent according to this expanded schedule and changes in priorities per approval by the NRCS CIG technical contact, Curtis Talbot.

## Introduction

The goals of this project were to increase ecological and wildlife knowledge and data accessibility, on a local and regional basis, to support, enhance, and accelerate the development of Ecological Site Descriptions (ESDs), and to enhance Wildlife Interpretation sections of ESDs by linking classification information to wildlife species of concern. To achieve this, we selected two Major Land Resource Areas (MLRA) where ESD development was just beginning to serve as pilots. These were MLRA 93A and MLRA 105. These MLRAs exhibit distinct landscapes and land management practices within the region thus providing data necessary to demonstrate the applicability of these data and technologies across the entire Midwest region.

MLRA 93A is in northeast Minnesota and encompasses approximately 22,205 square kilometers (Figure 1). It is relatively unaltered with most of the area in the Superior National Forest and the Boundary Waters Canoe Area Wilderness. It is considered part of the true forested region of Minnesota. Prior to European settlement, this area was almost entirely forested. It also contains many lakes, ponds, rivers, marshes, and bogs. Ecological communities in the region such as pine and hardwood forests, kettle lakes, and bogs support a high percentage of species of concern in Minnesota including the bald eagle, Canada lynx, and the eastern timber wolf. Game species such as white-tailed deer, ruffed grouse, walleye, northern pike, and smallmouth bass are also common in the area. Timber harvesting is the primary land resource management in the region. It impacts erosion and water quality along with wildlife habitat quality. Conservation practices on these timber lands include forest stand improvement, management of wildlife habitat, and management of riparian areas to protect water quality, improve wildlife habitat, and protect streams and rivers (http://www.mo10.nrcs.usda.gov/mlras/93A/description.pdf).

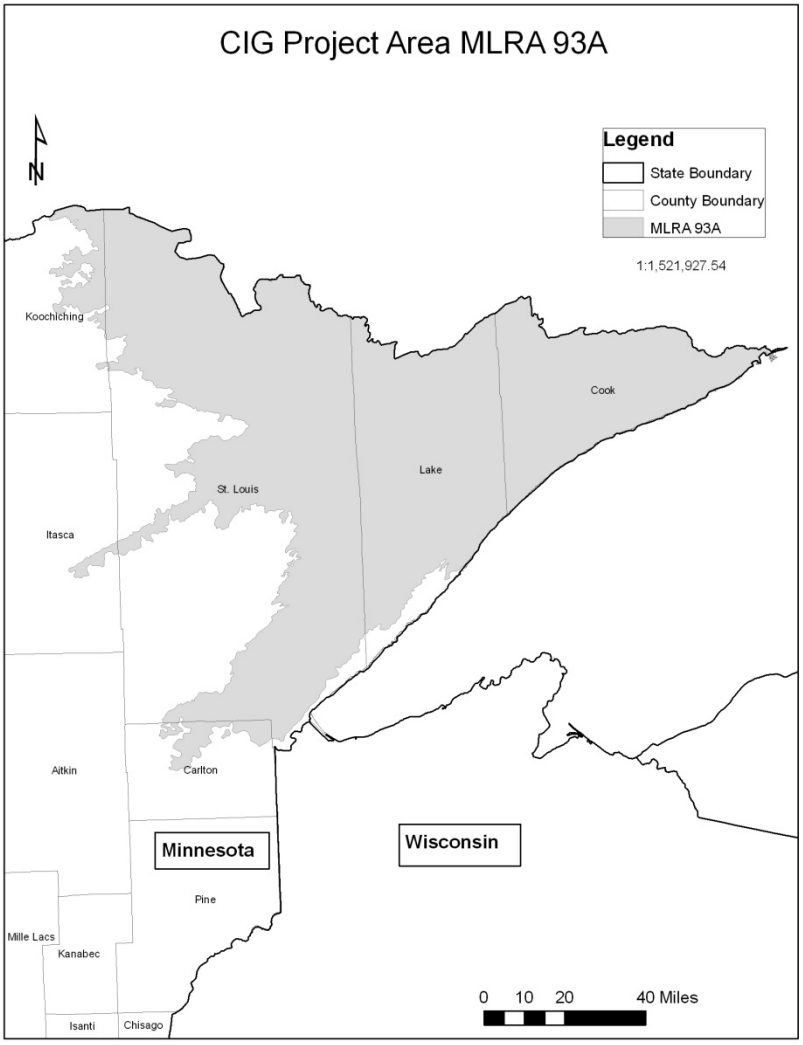



Figure 1: Boundary of MLRA 93A (Clarke, 2012, pers. communication).

MLRA 105 is found in the Wisconsin Driftless section of MN, WI, IA, and IL (Figure 2) and covers approximately 46,515 square kilometers. This area is unique within the upper Midwest as it has been only slightly impacted by glacial ice. As a result, the landscape includes features such as deep valleys, high bluffs, caves, and sinkholes. The area also includes the upper reaches and tributaries of the Mississippi River. Upland hardwood and mixed hardwood-conifer forests, lowland forests, savannas, and tallgrass prairie habitat all occur within this MLRA. Riverine, lake, and wet meadow habitats provide an abundance of aquatic habitat and resources. This unique landscape supports numerous wildlife and plant species such as white-tailed deer, gray fox, red fox, beaver, fisher, otter, Sandhill crane, bald eagle, peregrine falcon, and the great horned owl to name a few. Numerous waterfowl following the Mississippi Flyway also occur in this MLRA on a seasonal basis. However, unlike MLRA 93A, nearly all of MLRA 105 is heavily impacted by agriculture, in particular row crops, or residential and business development. One-half of the area is cropland and 15% in permanent pasture. Farm woodlots are often used for commercial timber production or farm products (http://www.mo10.nrcs.usda.gov/mlras/105/description.pdf).

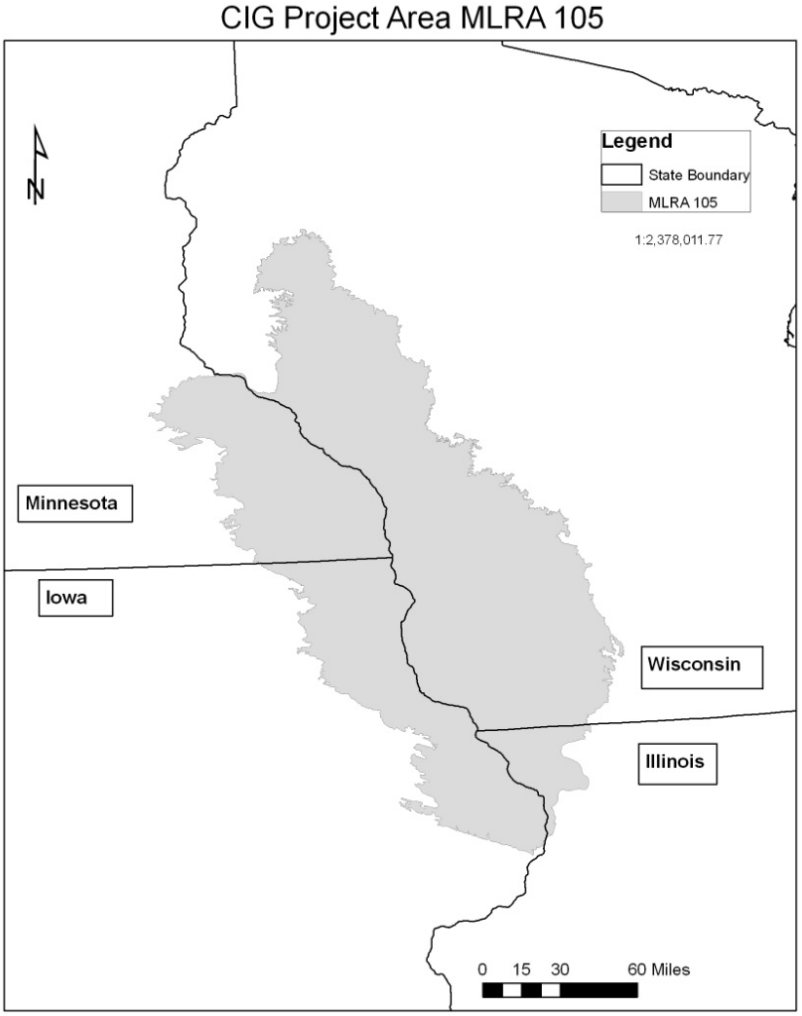



Figure 2: Boundary of MLRA 105 (Clarke, 2012, pers. communication).

At the outset of the project, there were three primary objectives.

1. Contribute to the development of ESDs with NatureServe's expertise in developing classifications and relevant information from NatureServe classifications. This project began by linking NatureServe data and expertise with ESD development in MLRAs 93A and 105. Connecting NatureServe data to ESDs would enhance the regional significance and context appropriateness of ESDs. Linking the development of ESDs to the USNVC/Ecological Systems, that is, using this classification information during the development of ESDs rather than determining the relationship post hoc, would allow the resources and information already present in the USNVC/Ecological Systems to be used in the application of ESDs. The classification units are defined in a national context and, to the extent that ESDs use USNVC units in their development, this link will allow a straightforward comparison of different ESDs. That is, since individual USNVC and Ecological System units are consistent across political and administrative boundaries, it will be easier to compare ESDs within or across MLRA boundaries based on their links to USNVC/Ecological Systems units. The lower level USNVC units, which would be most appropriate to link to ESDs, also have information on floristic composition, community

dynamics, and environmental characteristics that could help inform the Plant Community section of the ESDs.

2. Expand the NatureServe Wildlife Habitat Characterization database to all of the project area. After the ESDs and USNVC are linked, NatureServe would develop an innovative enhancement to the Wildlife Interpretation section of ESDs. NatureServe has developed a "Habitat Characterization" database to monitor the relationship of species to habitats in the Lake Superior region of MN, WI, and MI (Comer et al. 2010). This project would expand the Habitat Characterization database to include parts of the Upper Mississippi River basin (MLRA 105) and to add state listed wildlife species of concern in the Arrowhead region of MN (MLRA 93A), and include a generalized process for applying a habitat-based approach to addressing at-risk biodiversity. If these habitat associations could be reliably discerned in the field or from existing maps and other information sources (e.g., remote sensing, forest inventory systems), it would enhance the ability to apply standards efficiently for conserving at-risk biodiversity.
3. Assist NRCS in presenting wildlife-habitat data to producers. Through this project, NatureServe will assist NRCS in presenting the data linking wildlife species of concern to ecological communities in the region to a select group of EQIP eligible producers in that region. NRCS staff will identify these producers. These data can help producers both identify possible habitat and species of concern on their lands and identify possible ways that conservation management (e.g., through the Conservation Reserve Program or WHIP program) would enhance this habitat, thereby enriching their property. In particular, a survey approach can be utilized to determine possible use of NRCS data by producers. Some possible questions could include:
    a. How many producers have land not in agricultural production;
    b. How many producers would be interested in the information contained within an ESD, in particular the habitat information;
    c. How many would use this information to review landuse practices; and
    d. How many would take part in the NRCS conservation practice or program as a result of habitat information in the ESD. Ideally this would lead to producers allowing NRCS the means to collect further ecological and habitat data on lands not in agricultural production.

As the project developed, the objectives were adjusted to account for the number of ESDs available for consideration and for our growing understanding of the difficulties in achieving Objective 2 and the subsequent impact on Objective 3.

Primary NatureServe staff for this project were

Dr. Shannon Menard, Senior Vegetation Ecologist, Midwest Region of NatureServe. She served as lead scientist from NatureServe and was responsible for overseeing and managing NatureServe's work along with developing many of the techniques to link the classification information.
She manages the strategic direction and management of the Midwest regional ecology program and directs and participates on several national ecology initiatives for NatureServe. Her responsibilities include the quantitative analysis of ecological data, development and application of standard methods of ecological sampling and inventory, and the mapping and classification of vegetation communities. She also works with other NatureServe ecologists to maintain and interpret the vegetation classification systems developed by NatureServe (US National Vegetation Classification System and Classification of Ecological Systems) along with other community data for the research and conservation of ecological communities and ecosystems, especially in the Midwest and Great Plains. She has worked directly with developing EO Ranking Criteria for upland

and wetland ecological systems throughout the Great Plains and Midwest. She helps with project development and management as the operations manager for NatureServe ecology. Dr. Menard holds a B.A. in Biology from Gustavus Adolphus College, a Master of Forest Biology from Purdue University, and a Ph.D. in Forest Science (emphasis: Ecology) from Michigan Technological University.

Jim Drake, Regional Vegetation Ecologist, Midwest Region of NatureServe. He assisted in developing the classification information and worked on integrating that information into the Wildlife Habitat Characterization database. He also reviewed draft ESDs provided by NRCS. His work focuses on development and application of the US National Vegetation Classification (USNVC) and Classification of Ecological Systems in the Midwest Region. This includes quantitative analysis of ecological data, developing new and revising existing vegetation units, training users in field data collection and field application of USNVC concepts, and mapping vegetation communities. He has worked extensively with classifying and mapping vegetation in National Park Service lands in the Midwest and western US as well as other projects mapping vegetation or landcover (Landfire, USGS GAP) and applying vegetation classification information to natural resource management and conservation decision making. Mr. Drake has a B.S. in Biology from Lewis and Clark College and a M.S. in Conservation Biology from the University of Minnesota.

Regan Smyth, Spatial Ecology Project Manager, NatureServe. She performed all the GIS and statistical analyses determining the relationship of soil variables to Ecological Systems in MLRA 105. In her role at NatureServe, Ms. Smyth is intricately involved in the modeling and mapping of ecological systems, as well as in the development of GIS and statistical methods to assess the relationship between environmental variables (e.g. soils, disturbance, etc.) and ecosystem occurrence and condition. Ms. Smyth has a B.S. in Environmental Science and a M.E.M. in Ecosystem Science and Management from Duke University.

This project was funded by NRCS CIG with a match from a NatureServe project funded by NCEAS (National Center for Ecological Analysis and Synthesis) and NatureServe internal funds.

## Background

Ecological Sites are used by the (NRCS) and others to classify and map the landscape. Ecological Sites are defined as "a distinctive kind of land with specific physical characteristics that differs from other kinds of land in its ability to produce a distinctive kind and amount of vegetation and in its response to disturbance." (NRCS 2003). The goal of NRCS's Ecological Site Inventory program is to identify and describe Ecological Sites in all lands, seamlessly, across the United States, and to make these descriptions available to internal users, to partners, and to the public for conservation planning. Hundreds of approved Ecological Site Descriptions (ESD) exist across 35 states (USDA 2015). Ecological Site Descriptions provide a consistent framework for stratifying and describing soil, vegetation, and abiotic features and delineating units that share similar capabilities to respond to management activities or disturbance processes. They also identify restoration pathways and conservation practices which are most appropriate for that unique Ecological Site to either restore the site to a historical reference condition state, or to improve and maximize the performance of the site for a sustainable balance between ecosystem function, wildlife habitat use, and economic productivity.

NatureServe has worked with many agencies to develop the US National Vegetation Classification (USNVC) and the classification of ecological systems. The USNVC is a hierarchical classification of existing vegetation, which allows users to work at several scales (Fig. 1). It is the Federal Geographic Data Committee (FGDC) reporting standard for vegetation classification (FGDC 2008). NatureServe also has developed an Ecological Systems classification, which uses the base community (association) level of the USNVC to describe communities that co-occur on the landscape, tied together by underlying ecological processes and patterns (Comer et al. 2003). Both the USNVC and Ecological Systems classifications have been used by other mapping and classification efforts across the US, e.g. the NPS Vegetation Inventory Program (Lea 2011), Gap Analysis Program (GAP 2012), and LandFire (Vogelmann et al. 2011).

Figure 1. US National Vegetation Classification Hierarchical Structure with example association from MLRA 93A.

**US National Vegetation Classification Hierarchical Structure**
**Class**: Forest & Woodland)
**Subclass**: Temperate & Boreal Forest & Woodland
**Formation**: Cool Temperate Forest & Woodland
**Division:** Eastern North American & Great Plains Cool Temperate Forest & Woodland
**Macrogroup:** Laurentian & Acadian Northern Hardwood - Conifer Mesic Forest
**Group:** Laurentian & Acadian Hardwood Forest
**Alliance:** (broad floristic) Laurentian-Acadian Sugar Maple Rich Mesic Forest
**Association:** (base unit) Maple - Yellow Birch - Basswood Northern Forest

ESDs are developed within individual MLRAs and are not explicitly compared to potentially similar ESDs in nearby MLRAs. This affects the ability to roll-up ESDs for analyses at broader scales as well as the ability to easily develop and apply attributes to all similar ESDs across multiple MLRAs. That is, if it is determined that an ESD provides suitable habitat for certain wildlife species there is no easy way to determine what other ESDs might also provide habitat for that same species. These issues are, to our knowledge, not currently addressed in any systematic way. This results in duplication of effort in writing ESDs and also in inefficient use of information in ESDs since individual ESDs have to be examined and compared for similarity for larger scale analyses or uses.

Establishing a clear link between Ecological Sites and the USNVC and Ecological Systems would allow the information in those classifications to be related to Ecological Sites. NatureServe and its Natural Heritage Network members have datasets and analytical methods that would assist with the identification, classification, and description of Ecological Sites. NatureServe is adept at taking these data from state, local, and federal government agencies and standardizing them across regions to make one seamless dataset available to a broader range of users. NatureServe has completed projects linking different classification and mapping systems throughout the development of the USNVC and Ecological Systems. Linking the USNVC and Ecological Systems to ESDs during the development of ESDs would also allow the use of this suite of mapped and classified vegetation data to be available for defining provisional ESDs.

Establishing a clear and defined link between Ecological Sites and the USNVC and Ecological Systems will also allow the enhancement of the Wildlife Interpretation section of the ESD.

Standards for Wildlife Interpretation section of ESDs are currently being revised (S. Clark, pers comm). Information about wildlife species potentially occurring in Ecological Sites is currently not a standard part of the Site Interpretation section but that information could be valuable to land managers, private land owners, or others making natural resource management decisions. NatureServe has experience establishing habitat relationships for at-risk species using Ecological Systems (Comer et al. 2010) and would apply these proven methods to enhancing the Wildlife Interpretation Section of the selected ESDs. This project would facilitate the transfer of conservation data, expertise, and technologies between NatureServe and NRCS and help with the development of ESDs, including ecological data and wildlife habitat information. Specifically, this project addresses "**Priority Need #4: Wildlife**: Demonstrate new techniques and/or technologies for monitoring and evaluating wildlife habitat both on site and via remote sensing".

# Review of Methods and Quality Assurance

## Contribution to the Development of Ecological Site Descriptions

In the process of defining ESDs in MLRA 93A and MLRA 105, NRCS staff looked to incorporate and make use of existing data, where possible. NatureServe provided links between the USNVC hierarchy and Ecological Systems and the draft ESDs as well as the relationship between the USNVC and Ecological Systems and the Minnesota Native Plant Community classification. This allowed use of the extensive information in these classification systems to be used in the development of ESDs. NatureServe staff also reviewed drafts of ESDs, including State-and-Transition models, to validate the relationship of the ESDs and NatureServe classification units as well as providing general review of the ecological information in the ESDs. The only change to the project deliverables was the addition of some ESDs from MLRA 103. NRCS had moved their efforts from MLRA 93A to MLRA 103 based on changes in their priorities. We reviewed those as well and added data to the descriptions based on our experience in that area. This was an enhancement to our stated deliverables. These steps went smoothly and the final ESDs have a firm relationship to the USNVC and Ecological Systems, which should allow continued easy use and transfer of information as both classifications are further developed.

## Connecting the Wildlife-Habitat Database to Ecological Site Descriptions

During the execution of the project, it became clear that producing a wildlife-habitat database was beyond what could be done within the limits of this project. The steps required for our original plan were to

1. Modify an existing database created to link wildlife and Ecological Systems for the Great Lakes area to help meet the needs for the Wildlife Interpretation Section of the ESDs.
2. Create a list of rare wildlife species in the pilot MLRAs and describe the link between their habitat requirements and Ecological Systems. That is, which Systems would they be most likely to inhabit.
3. Gather occurrence data for tracked species from the individual states and overlay that with the US national map of Ecological Systems maintained by NatureServe.

As we began the modifications to the existing database and determine who would describe the link between wildlife habitats and Ecological Systems, the information that NRCS needed for this step in ESD development evolved. The priorities in NRCS ESD development decreased the need for this deliverable as originally scoped. An existing database had been developed with the NCEAS funded match project, which initially gave the information that NRCS needed with just some small additions. However, as the wildlife habitat data needed by NRCS changed, it became evident that this effort was well beyond the scope of this project. The number of species and the labor required to determine which Systems they could be linked to was not achievable given the current funding level and the need to complete other steps of the project. With approval from our NRCS partner and CIG technical contact, Curtis Talbot, we adjusted the project goal to determine if we could establish links between soil properties and Ecological Systems (as currently mapped by NatureServe; data available from http://www.natureserve.org/conservation-tools/terrestrial-ecological-systems-united-states).

## Comparison of Soil Variables to Ecological System Occurrence

In order to better understand the relationship between the occurrence of specific Ecological Systems sand underlying soil characteristics, we undertook a series of exploratory statistical analyses in MLRA105 (Driftless Area of MN, WI, IA, and IL). We used two sources for our comparisons. We compiled vegetation point data from the Minnesota Department of Natural Resources (https://gisdata.mn.gov/) and Effigy Mounds National Monument (http://science.nature.nps.gov/im/inventory/veg/project.cfm?ReferenceCode=1047704) to generate a list of 776 georeferenced sites where the vegetation was classified to the System. We then compared this to a seamless national SSURGO raster representing 46 soil variables compiled by the US Forest Service, as well as 30-meter resolution data on slope, elevation, and stream distance, to generate a series of boxplots to visualize relationships between environmental (e.g. soils, elevation) variables and ecological systems. We tested these relationships using t-tests. We also attempted Classification and Regression Tree (CART) models to ascertain if soil, elevation, and slope could successfully predict the occurrence of the Systems or shed light on the primary factors that might be driving vegetation patterns. For simplicity of analysis, we converted the multiple percent soil type (e.g. percent alfisols, percent entisols, etc.) and percent drainage class (e.g. percent very poorly drained, percent poorly drained, etc.) variables into a single categorical soil type variable and single categorical drainage class variable by assigning each pixel to the type or class of the highest percentage.

For purposes of comparison, we generated a series of random points from across the project area and used these as “absence” data for the t-tests, boxplots, and CART models. Any random points falling within an area of NatureServe’s National Map of Ecological Systems mapped as the same type of the ecological system being analyzed was excluded from the absence data. This gave us a set of points representing each mapped Ecological System and a set for all other map units. Using these methods, we were able to identify key soil characteristics related to Ecological Systems in MLRA105.

Results were compiled for the nine Ecological Systems with at least five occurrence points (field observation data) in the MLRA and are summarized below. Details of the statistical results can be

found in Appendix B. This portion of the project was demonstrated at the 2015 National Cooperative Soil Survey National Conference in Duluth, MN, June 2015.

# Findings

## Contribution to the Development of Ecological Site Descriptions

Although Ecological Systems and the USNVC are based on existing vegetation, and existing vegetation can be greatly affected by landuse, weather, and other short-term processes, the NVC was developed from examples of "high quality" (i.e., relatively undisturbed and subject to natural disturbance regimes), usually late-seral vegetation. Because vegetation is so dependent on its environmental setting, it is a highly reliable indicator of soils, hydrologic regime, nutrient availability, slope position, and other factors from which the Ecological Site is derived. The USNVC and Ecological System vegetation classification units are described, geographically bounded, and standardized. In addition, the classification units were used by LandFire in the development of State and Transition models that reflect the successional pathways between types influenced by various disturbance regimes. All of these factors contributed to the utility of using Ecological Systems and the USNVC in the development of ESDs. NatureServe was provided portions of three draft ESDs to comment on. A brief summary of these activities follows.

A93Y001 – Till Upland Mesic Hardwood Forests – NatureServe assisted in establishing the link between the various Community Phases in this ESD and Ecological Systems and associations and between NatureServe vegetation associations and Natural Communities as defined by the Minnesota Department of Natural Resources. A draft of the ESD was provided to NatureServe in the fall of 2013 for review, particularly of the Ecological Concept, the State and Transition model, and the State and Community Phases.

103XY001 – Loamy Wet Prairie – NatureServe was asked to provide crosswalks between NatureServe vegetation associations and Natural Communities associated with this ESD, as defined by the Minnesota Department of Natural Resources. NatureServe was provided with the State-and-Transition model in the fall of 2015 and asked to provide text describing the States, Community Phases, and Pathways.

103XY002 – Pothole Marsh – NatureServe was asked to provide crosswalks between NatureServe vegetation associations and Natural Communities associated with this ESD, as defined by the Minnesota Department of Natural Resources. NatureServe was provided with the State-and-Transition model in the fall of 2015 and asked to provide text describing the States, Community Phases, and Pathways.

## Comparison of Soil Variables to Ecological System Occurrence

Our results showed that soil and other environmental characteristics can have meaningful relationships to Ecological Systems. These relationships could be used to help define Ecological Sites as they are being developed and could make using other existing data sources easier, given

the established relationship between the USNVC and Ecological Systems and other existing data (e.g., local and regional maps and vegetation classifications and descriptions). Not all Ecological Systems were found to have significant relationships to soil and environmental variables but many of these had few samples so determining significance was not possible. Below are summaries of the statistically significant relationships between our field data and soil and environmental variables.

*Central Tallgrass Prairie (CES205.683)*

Our results, based on 20 occurrence points within the analysis area, indicate that, as expected, Central Tallgrass Prairie ecosystems are most likely to be located on alfisols and mollisols. Due to the relatively small sample size (n=20), t-tests show no significant differences (p-value < 0.05) between presence and absence points for most other variables, with the exception of elevation and water volume. Mean water volume (1/10 bar) was 9.5 percent at occurrence locations, as opposed to 13.9 percent (standard deviation = 4.5) elsewhere. Mean elevation was 360 meters (standard deviation = 49.99) at occurrence locations, as opposed to 332 meters (standard deviation = 60.29) elsewhere. Occurrence points were also more likely to be located at locations with southerly aspect (mean Beer's transformed aspect of 0.765), though this relationship was not statistically significant with 95% confidence (p-value = 0.09)

*North Central Interior Sand and Gravel Prairie (CES202.695)*

A relatively high percent sand (mean 40.4%, standard deviation 26.8%) was a defining and statistically significant (p-value=0.013) characteristic of soils where NCI Sand and Gravel Prairies are located. Percent sand at absence locations had a mean of 2.8% and standard deviation of 18.3%. Other soil measures characterizing grain size, such as the soil fraction passing through a #200 sieve also showed statistically significant differences between presence and absence points for this system. Available Water Capacity (AWC) was also lower for this system (mean = 46.4, standard deviation = 18.2) than for the absence locations (mean = 59.2, standard deviation = 14.1). Elevations where this system occurs were also found to be lower (mean = 247 meter, standard deviation = 45) than for the area as a whole. The boxplots indicated other differences in soil characteristics, particularly those related to texture and volume content, but these were not statistically significant at the 95% level of confidence, likely due to the relatively small number of sample points.

*North-Central Interior Dry Oak Forest and Woodland (CES202.047)*

Despite a small sample size (n=5) statistically significant differences between presence and absence points were uncovered for several soil variables as well as for percent slope. North Central Interior Dry Oak Forest and Woodland was observed to occur on soils with a relatively high percent sand (mean 39.6, standard deviation 36.05) and relatively low pH (mean 5.78, standard deviation 4.09) as compared to the study area as a whole. As might be expected, water volume (1/10 bar) was lower than for the region as a whole (mean = 12%). While other metrics relating to soil measure where not statistically significant at the 95% level of confidence, the boxplots indicate that WHC and AWC are also relatively lower for this system than for the area as a whole. Percent slope (mean = 14.6%, standard deviation = 9%) was greater for this system than for the absence points.

*North-Central Interior Dry-Mesic Oak Forest and Woodland (CES202.048)*

NCI Dry-Mesic Oak Forests and Woodland occurrence points (n = 150) were overwhelmingly located on well-drained, finer-textured soils with comparatively higher water holding capacity. T-tests between presence and absence samples for percent sand, percent coarse, soil fraction passing a no. 4 sieve, AWC, and bulk density all indicated a statistically significant difference between the groups at the 95% level of confidence. Percent slopes at occurrence locations (mean = 14.61, standard deviation = 8.68) were relatively high for the region, and elevations (mean = 287.7 meters, standard deviation = 46.46) were relatively low.

*North-Central Interior Maple-Basswood Forest (CES202.696)*

Well-drained and coarser-textured soils characterized NCI Maple-Basswood Forest occurrence points, with statistically significant lower mean values for the soil fraction passing through no. 4, no. 10, and no. 40 sieves and bulk density and statistically significant higher mean values for percent coarse for occurrence points as opposed to absence points. The percent by weight of the horizon occupied by rock fragments 3 to 10 inches in size was also higher for the Maple Basswood sample than for the regional sample. The mean percent coarse fragments for this type was 14.8, with a standard deviation of 15.31.

Water content was also generally lower for this type than for other vegetation types within the region, with mean AWC, kSat, and water volume (1/3 and 15 bar) all statistically lower for the sample of NCI Maple-Basswood Forest sample points than for the absence points. Statistically significant differences were also observed for the percent slope, stream distance, and elevation variables, with NCI Maple-Basswood forest characterized by steeper slopes (mean = 12.93%), closer proximity to streams (mean = 340.67 meters), and lower elevations (mean = 278.13).

*North-Central Interior and Appalachian Rich Swamp (CES202.605)*

Relatively few statistically significant differences were found for NCI and Appalachian Rich Swamps, likely because of the small sample size for this system (n=14). The data indicate finer soil textures for the presence sample (statistically significant lower percent coarse and percent sand). Statistically significant higher mean slope (mean = 12.7%), smaller stream distance (mean = 181 meters), and lower elevations (mean = 286 meters) were also observed.

*North-Central Interior Floodplain (CES202.694)*

North Central Interior Floodplains (n=199) had statistically significant higher proportions of clay and sand than the absence sample, with mean percent clay of 20.1% and mean percent sand of 17.9% (standard deviation = 18.0%); these differences in soil texture were also reflected in statistically higher means for the fraction of the soil passing through various size sieves (no. 10, no. 40, and no. 200). As might be expected, the mean AWC (mean=65.0, standard deviation = 16.1) and water volume (1/10 bar) (mean=12.5%, standard deviation =2.3%) were higher for the Floodplain sample than for the region as a whole.

*North-Central Interior Wet Meadow-Shrub Swamp (CES202.701)*

Occurrence points for NCI Wet Meadow-Shrub Swamp (n=23) were disproportionally located on poorly drained soils. Mean water volume (1/10 bar) was higher than for the absence sample, with

a mean value of 16.5% (standard deviation 0.58) compared to 14.0% (standard deviation 3.6) for the region as a whole (p-value = 0.00007). Mean saturated hydrologic conductivity (kSAT) was statistically lower for this type (mean = 6.65, standard deviation =6.48) than for the absence sample (mean = 10.8, standard deviation = 14.5). Soil textures were generally finer for this type, with a mean percent silt of 52.7% (standard deviation 19.9) for the presence sample as opposed to 46.4% (standard deviation 19.3) for the absence sample. Other soil measures characterizing grain size, such as the soil fraction passing through both a #4 and #10 sieve also showed statistically significant differences between presence and absence points for this system. NCI Wet Meadows-Shrub Swamp were also characterized by statistically significant lower percent slope (mean = 1.6, standard deviation = 2.4) and lower elevations (mean = 296.1, standard deviation = 75.1).

*North-Central Oak Barrens (CES202.727)*

North-Central Oak Barrens sites were associated with drier, coarser soils. Percent sand was 40% vs. 23% in presence vs. absence sites and bulk density was 80% vs. 74%, respectively. Variables that were statistically lower for this System were water holding capacity (98 vs. 145), percent that passed through a #40 sieve (68% vs. 86%), Plasticity Index (4 vs. 37), and available water capacity (39 vs. 66).

## Conclusions and Recommendations

The USNVC and Ecological Systems and Ecological Sites are widely used national classifications; the USNVC is the current FGDC standard for mapping vegetation across the US. Establishing a strong link between them, while retaining their own standards and purposes, will allow more efficient sharing of information from one classification to the other and increase the applicability of ESDs to other agencies already using the USNVC such as National Park Service, US Forest Service Forest, and Bureau of Land Management (BLM). Few ESDs are completed in the upper Midwest and other parts of the eastern United States so there is substantial potential for using USNVC and Ecological System information in developing ESDs. Integrating the USNVC and Ecological Systems into the drafting of ESDs, rather than a post hoc crosswalk, ensures the best possible linkage between the classifications, which will allow the most efficient use of information in these vegetation classifications. The USNVC and Ecological Systems have a well-developed set of vegetation units at different scales, descriptive material for these units, ancillary products in some areas (e.g., local and regional maps, field keys, rarity metrics, ecological integrity metrics), and established relationships with other classification systems developed for individual states, National Forests, academic studies, etc.

Our project demonstrated that using information and vegetation units in the USNVC and Ecological Systems classification to develop ESDs is possible, practical, and, we believe, useful. We also believe that having these established links between ESDs and the USNVC and Ecological Systems will make continued use and development of ESDs easier in the future, as developments in the vegetation classifications can be easily transferred to appropriate sections of ESDs. The process of creating the link between ESDs and the vegetation classifications requires personnel familiar with the USNVC and Ecological Systems and access to classification information. As a

result, complementary collaborative projects are occurring across the US. NatureServe is working with NRCS personnel in the southeast, northeast and western US to develop ESDs using expertise from both NatureServe and NRCS. Likewise, given the success of this project and others, NatureServe is working with NRCS staff to develop a larger proposal to join these efforts and develop standard methodology in ESD development and expand current NRCS-NatureServe partnerships, which will help further ESD development in the US.

Essential descriptive information on the vegetation classifications can be downloaded from the NatureServe Explorer website (http://explorer.natureserve.org/servlet/NatureServe?init=Ecol). Other, more detailed information, including other relevant projects and products in the area of interest or providing classification data in other formats that may be more useful to a project, will need to be gathered from NatureServe, state and federal agencies, or other groups that use the USNVC and Ecological Systems classifications extensively.

# Appendices

## Appendix A. References:

Comer, P., D. Faber-Langendoen, R. Evans, S. Gawler, C. Josse, G. Kittel, S. Menard, M. Pyne, M. Reid, K. Schulz, K.Snow, and J. Teague. 2003. *Ecological Systems of the United States: A Working Classification of U.S. Terrestrial Systems.* NatureServe, Arlington, Virginia.

Comer, P., M. Pyne, K. Snow, J. Cordiero, R. Crawford, J. Drake, J. Fleckenstein, J. Gamon, S. Hall, C. Nordman, M. Reid, J. Rocchio, R. Lyons Smyth, and B. Sorrie. 2010. Habitat Relationships Database Pilot: Supporting a Habitat-Based Approach to Conservation of At-risk Biodiversity on SFI Certified Lands. Final Report submitted to National Council for Air and Stream Improvement.

Faber-Langendoen, D., D. Tart, and R. Crawford. 2009. Contours of the Revised U.S. National Vegetation Classification Standard. Bulletin of the Ecological Society of America. Pp 87-93.

Federal Geographic Data Committee (FGDC). 2008. National Vegetation Standard, Version 2. Federal Geographic Data Committee, Vegetation Subcommittee FGDC-STD-005-2008 (Version 2).

GAP 2012. USGS GAP Analysis Program. GAP Land Cover website. http://gapanalysis.usgs.gov/gaplandcover/

Grigal, D. and L. Ohmann. 1975. Classification, Description, and Dynamics of Upland Plant Communities within a Minnesota Wilderness Area. Ecological Monographs 45:389-405.

Lea, C. 2011. Vegetation classification guidelines: National Park Service Vegetation Inventory, version 2.0. Natural Resource Report NPS/NRPC/NRR—2011/374. National Park Service, Fort Collins, Colorado.

USDA. 2003. National Range and Pasture Handbook. Natural Resources Conservation Service, USDA. 493 pp.

USDA. 2015. Ecological Site Description website, accessed 9/2015. https://esis.sc.egov.usda.gov/Welcome/pgReportLocation.aspx?type=ESD

Vogelmann, J., J. Kost, B. Tolk, S. Howard, K. Short, X. Chen, C. Huang, K. Pabst, and M. Rollins. 2011. Monitoring Landscape Change for LANDFIRE Using Multi-Temporal Satellite Imagery and Ancillary Data. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, Vol. 4, No. 2.

## Appendix B. Detailed Results of Comparison of Soil Variables to Presence or Absence of Ecological Systems in MLRA 105.

Statistically significant results were found for the following Ecological Systems

Central Tallgrass Prairie (CES205.683)

North Central Interior Sand and Gravel Prairie (CES202.695)

North-Central Interior Dry Oak Forest and Woodland (CES202.047)

North-Central Interior Dry-Mesic Oak Forest and Woodland (CES202.048)

North-Central Interior Maple-Basswood Forest (CES202.696)

North-Central Interior and Appalachian Rich Swamp (CES202.605)

North-Central Interior Floodplain (CES202.694)

North-Central Interior Wet Meadow-Shrub Swamp (CES202.701)

North-Central Oak Barrens (CES202.727)

# Notes

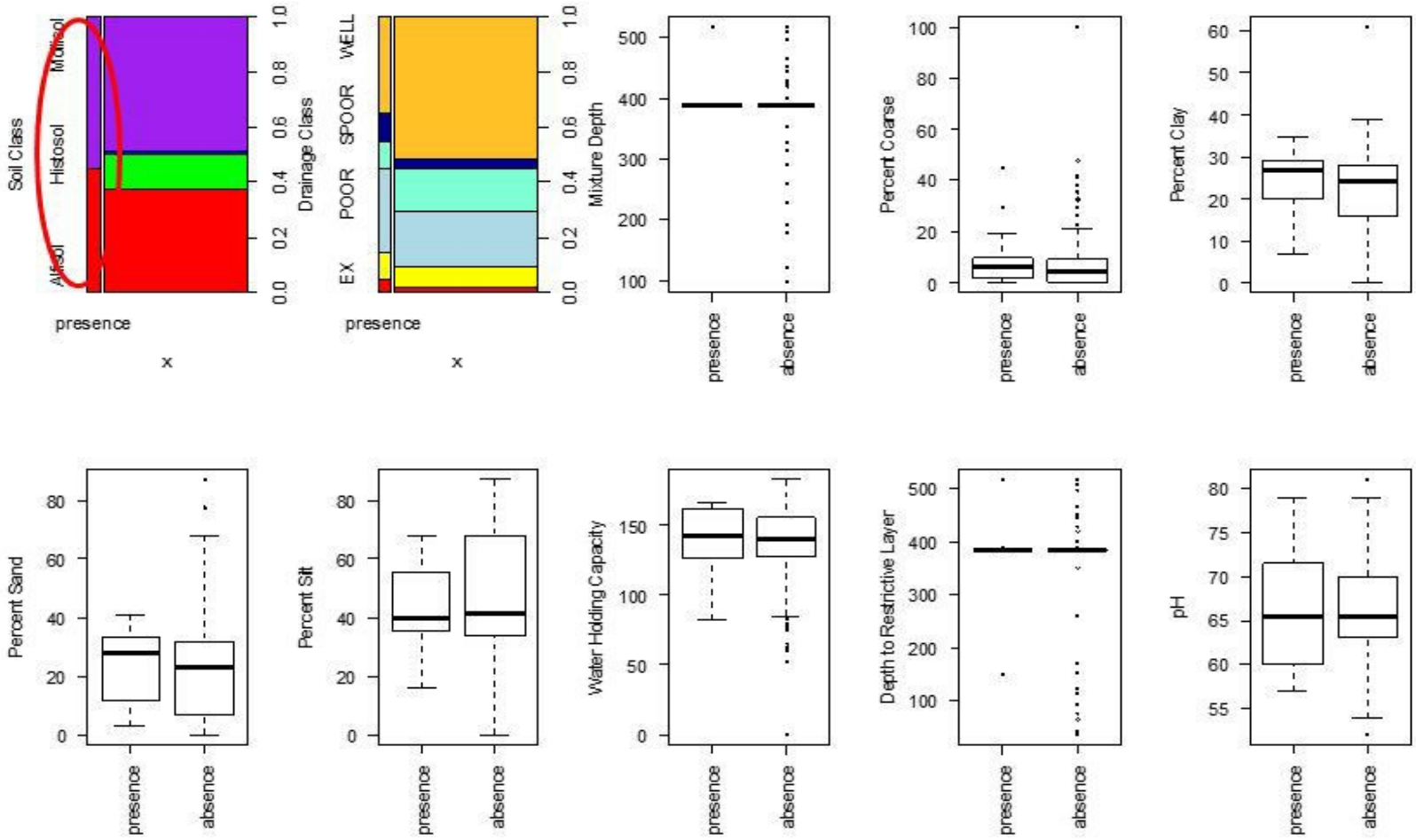
Soil types:
- Red = Alfisols
- Green = Entisols
- Blue = Histosols
- Yellow = Inceptisols
- Puple = Mollisols

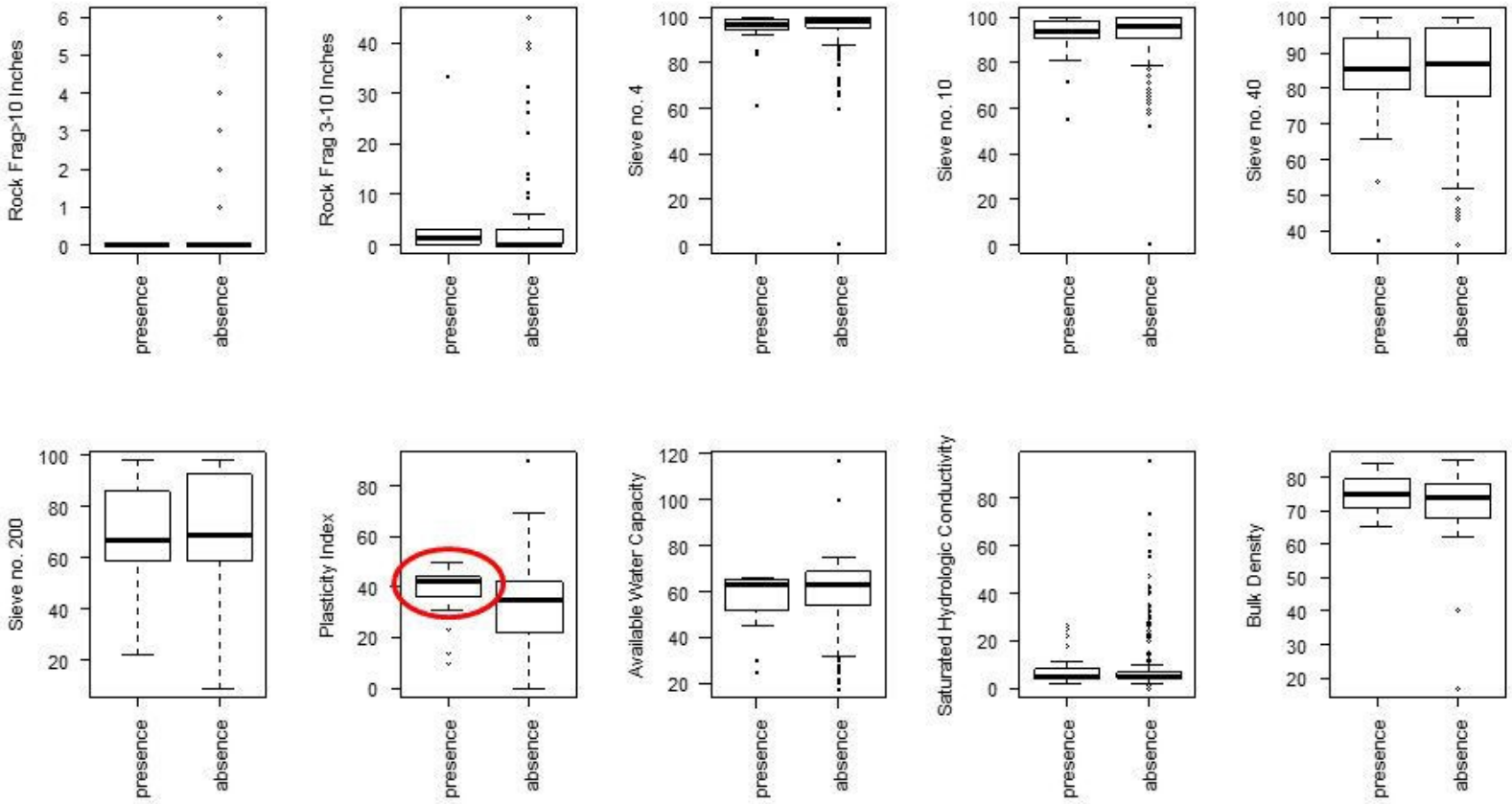
Sample size for presence listed for each system
Sample size for absence ~240 (differs slightly by system due to removal of random points falling within mapped area of system

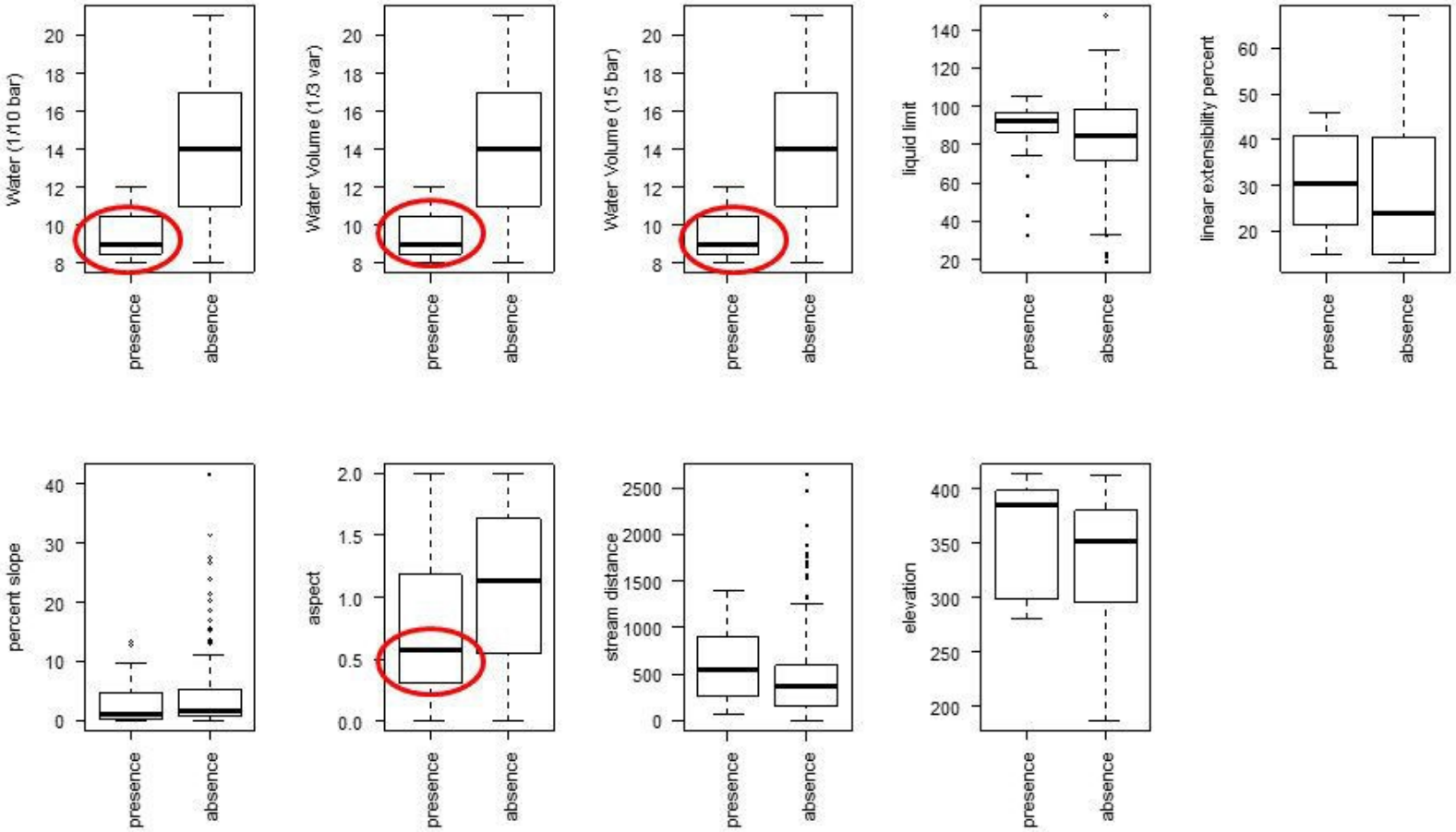
Drainage Class:
- Red = Excessively Well drained
- Dark Orange = Somewhat Excessively Drained
- Goldenrod = Well Drained
- Yellow = Moderately Well Drained
- Aquamarine = Somewhat Poorly Drained
- Light Blue = Poorly Drained
- Dark Plue = Poorly Drained

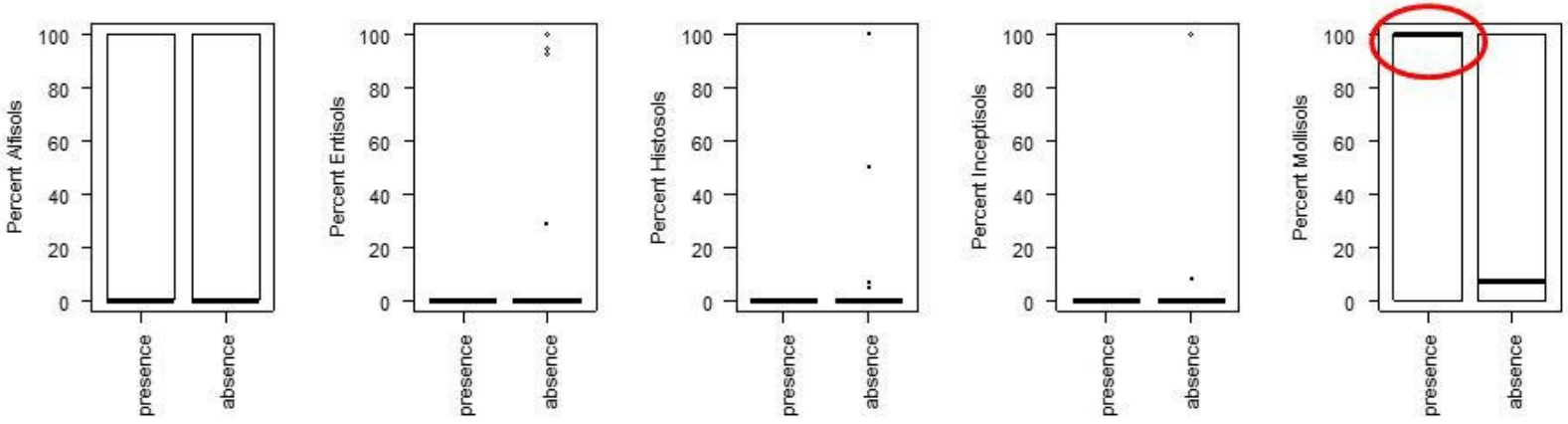
N = 20

# Central Tallgrass Prairie




Rock Frag>10 Inches
Rock Frag 3-10 Inches
Sieve no. 4
Sieve no. 10
Sieve no. 40
Sieve no. 200
Plasticity Index
Available Water Capacity
Saturated Hydrologic Conductivity
Bulk Density
presence
absence


Water (1/10 bar)
Water Volume (1/3 var)
Water Volume (15 bar)
liquid limit
linear extensibility percent
percent slope
aspect
stream distance
elevation
presence
absence


Percent Alfisols
Percent Entisols
Percent Histosols
Percent Inceptisols
Percent Mollisols
0
20
40
60
80
100
presence
absence

## Plasticity Index – not significant

```
Welch Two Sample t-test

data:  data$PI_R by data$group
t = 1.5641, df = 25.752, p-value = 0.13
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -1.272636  9.357155
sample estimates:
mean in group presence  mean in group absence
          37.90000            33.85774
```

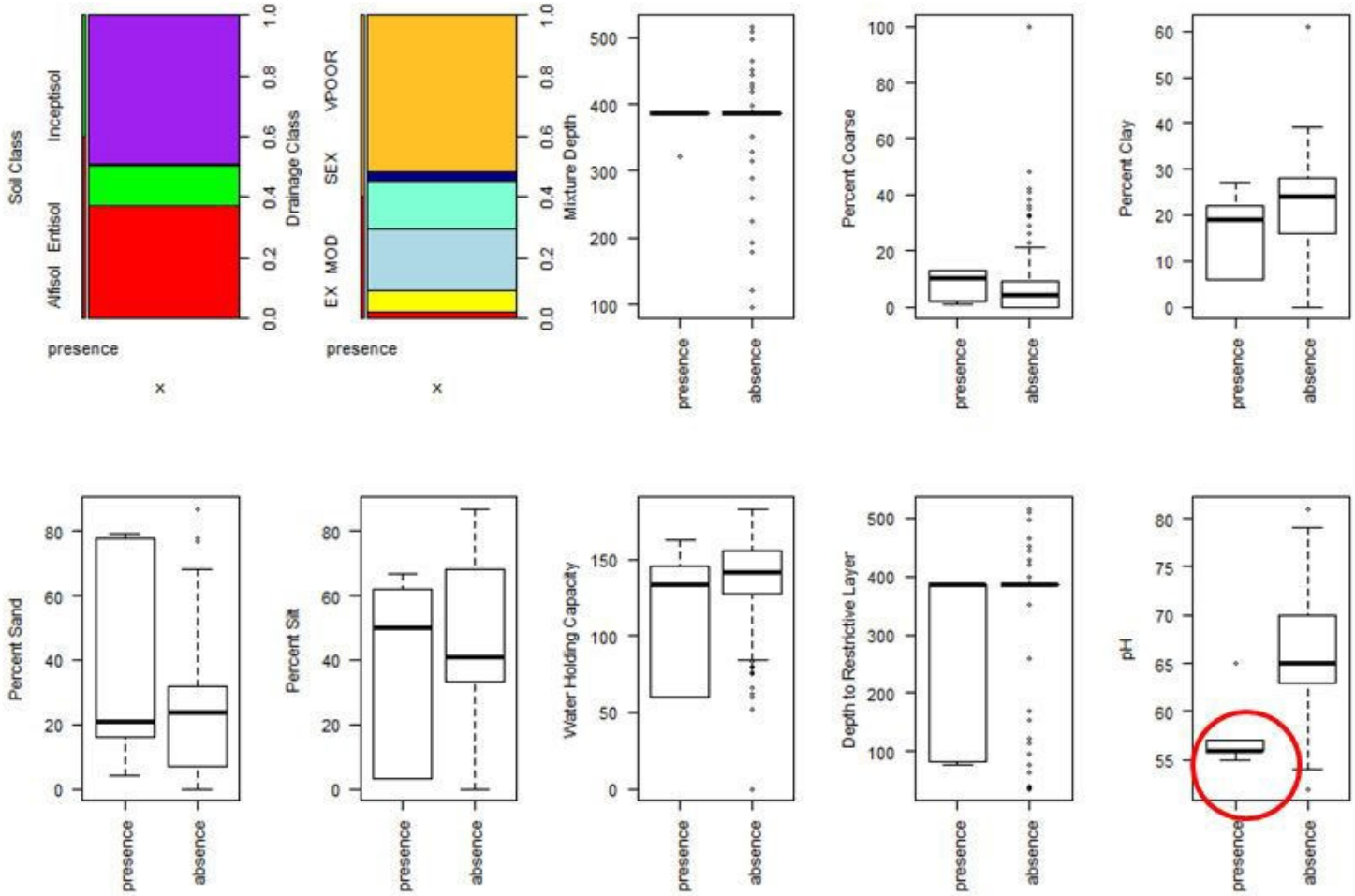
```
Welch Two Sample t-test

data:  data$slope by data$group
t = -0.7432, df = 25.697, p-value =
0.4641
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -2.915396  1.367763
sample estimates:
mean in group presence  mean in group absence
          3.160521            3.934338
```

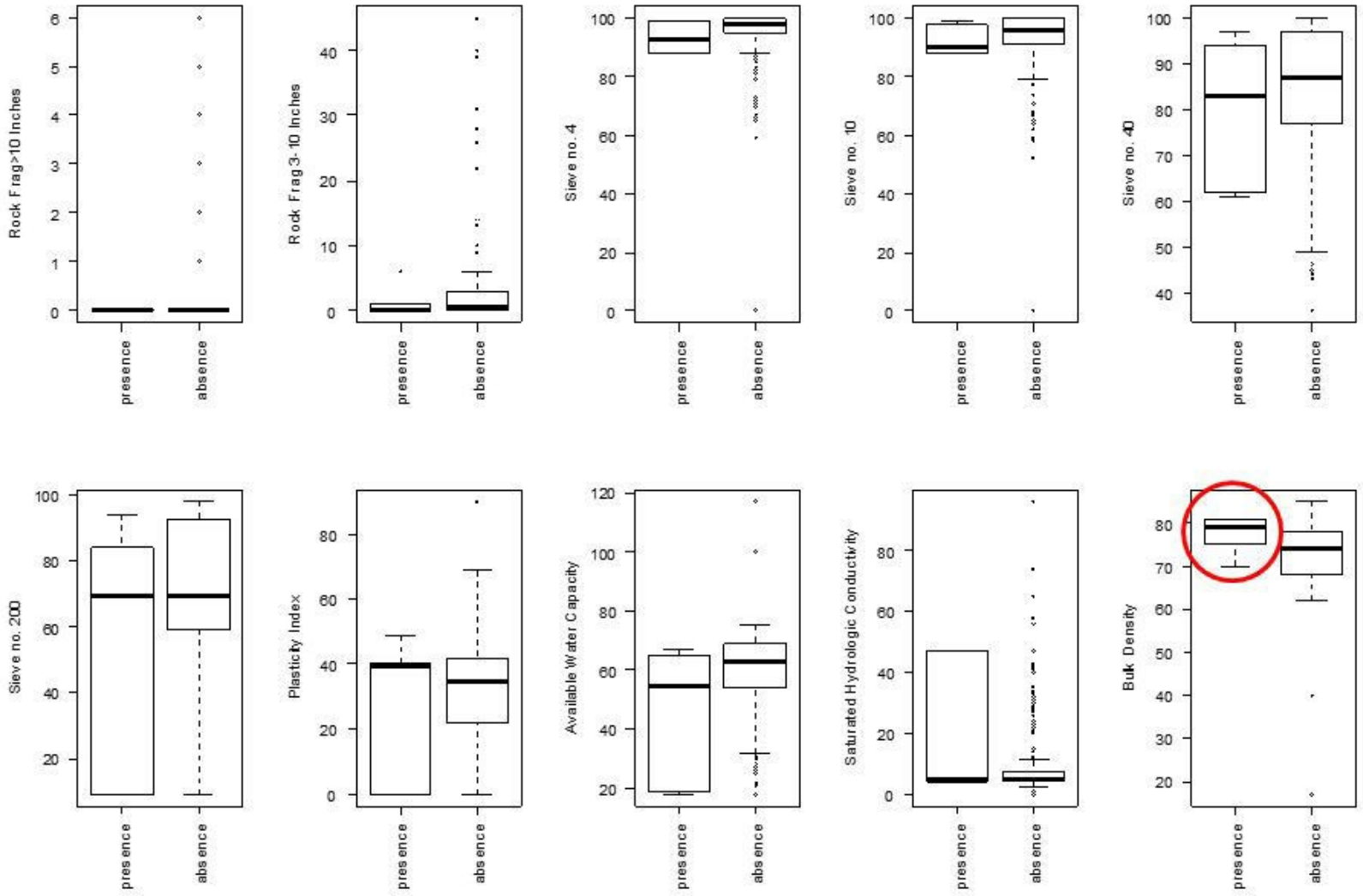
## Water Volume (1/10 bar) - significant

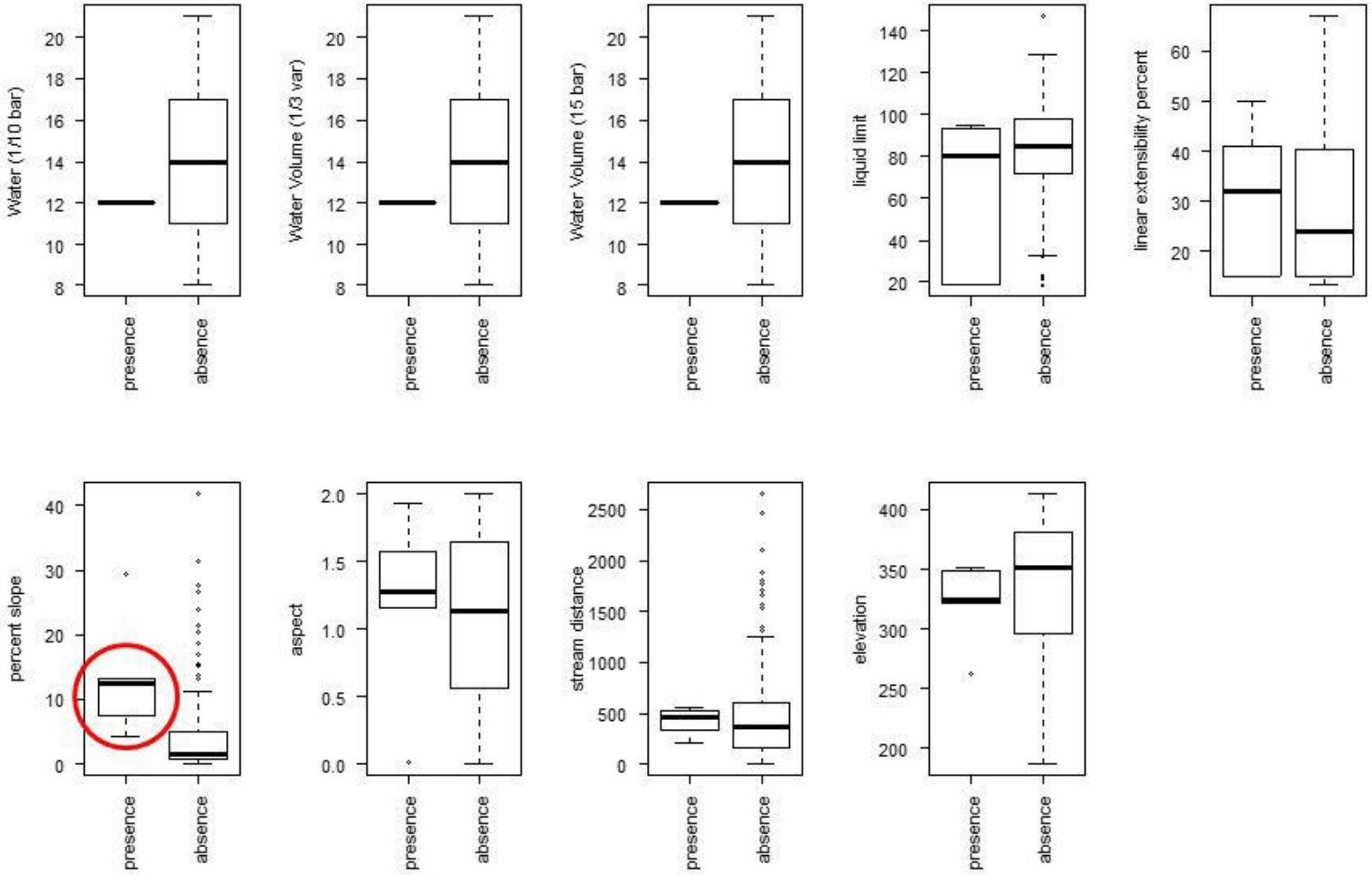
```
Welch Two Sample t-test

data:  data$WTENTHBAR_R by data$group
t = -4.5293, df = 5.013, p-value =
0.00619
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -6.996975 -1.932849
sample estimates:
mean in group presence  mean in group absence
          9.50000            13.96491
```

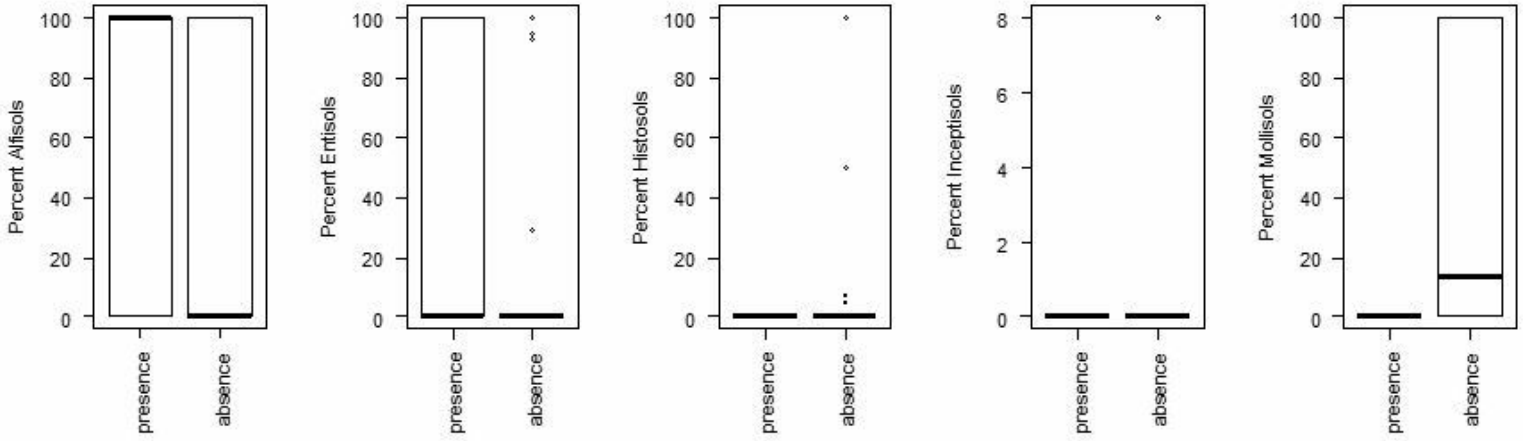
# NCI Dry Oak Forest

N = 5
NOTE SMALL SAMPLE SIZE




Rock Frag>10 Inches
Rock Frag 3-10 Inches
Sieve no. 4
Sieve no. 10
Sieve no. 40
Sieve no. 200
Plasticity Index
Available Water Capacity
Saturated Hydrologic Conductivity
Bulk Density
presence
absence


Water (1/10 bar)
Water Volume (1/3 var)
Water Volume (15 bar)
liquid limit
linear extensibility percent
percent slope
aspect
stream distance
elevation
presence
absence


Percent Alfisols
Percent Entisols
Percent Histosols
Percent Inceptisols
Percent Mollisols
presence
absence

## pH - significant

```
Welch Two Sample t-test

data:  data$PH by data$group
t = -4.4235, df = 4.304, p-value =
0.009723
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -13.260739  -3.206488
sample estimates:
mean in group presence  mean in group absence
          57.80000            66.03361
```

## Slope – not quite significant

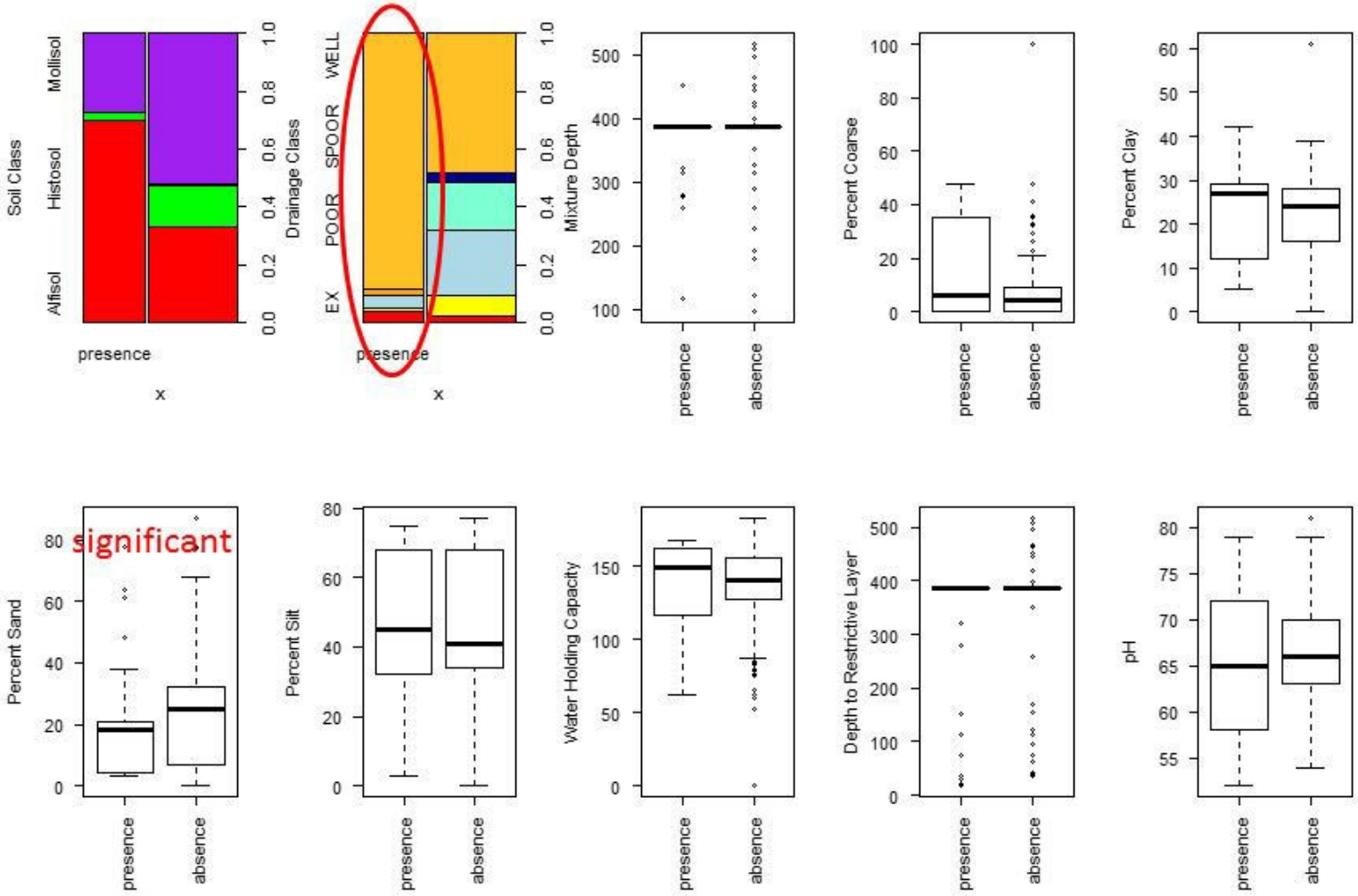
```
Welch Two Sample t-test

data:  data$slope by data$group
t = 2.1737, df = 4.065, p-value = 0.09432
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -2.551652 21.504467
sample estimates:
mean in group presence  mean in group absence
          13.398002            3.921594
```

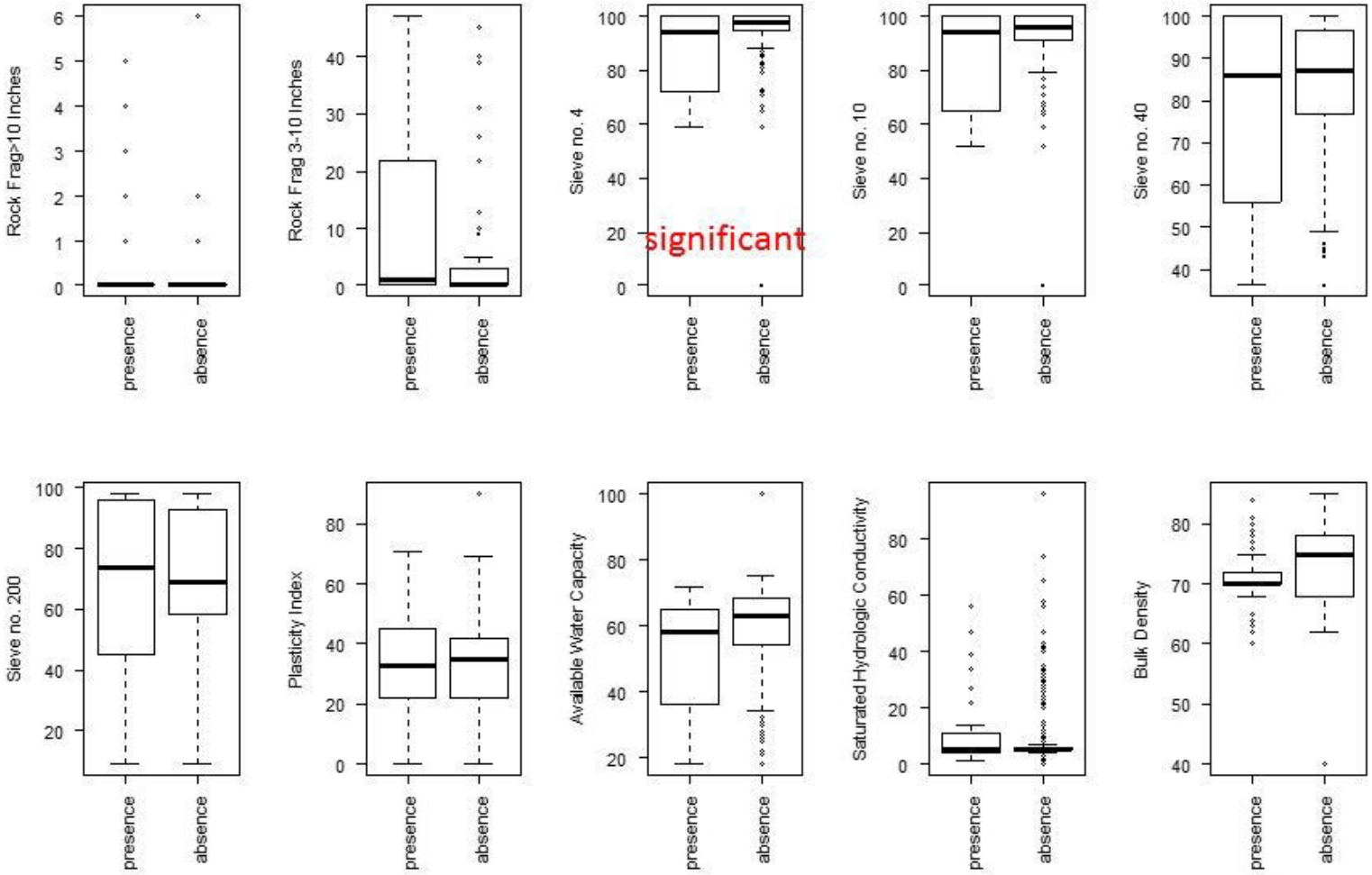
## Bulk density – not significant

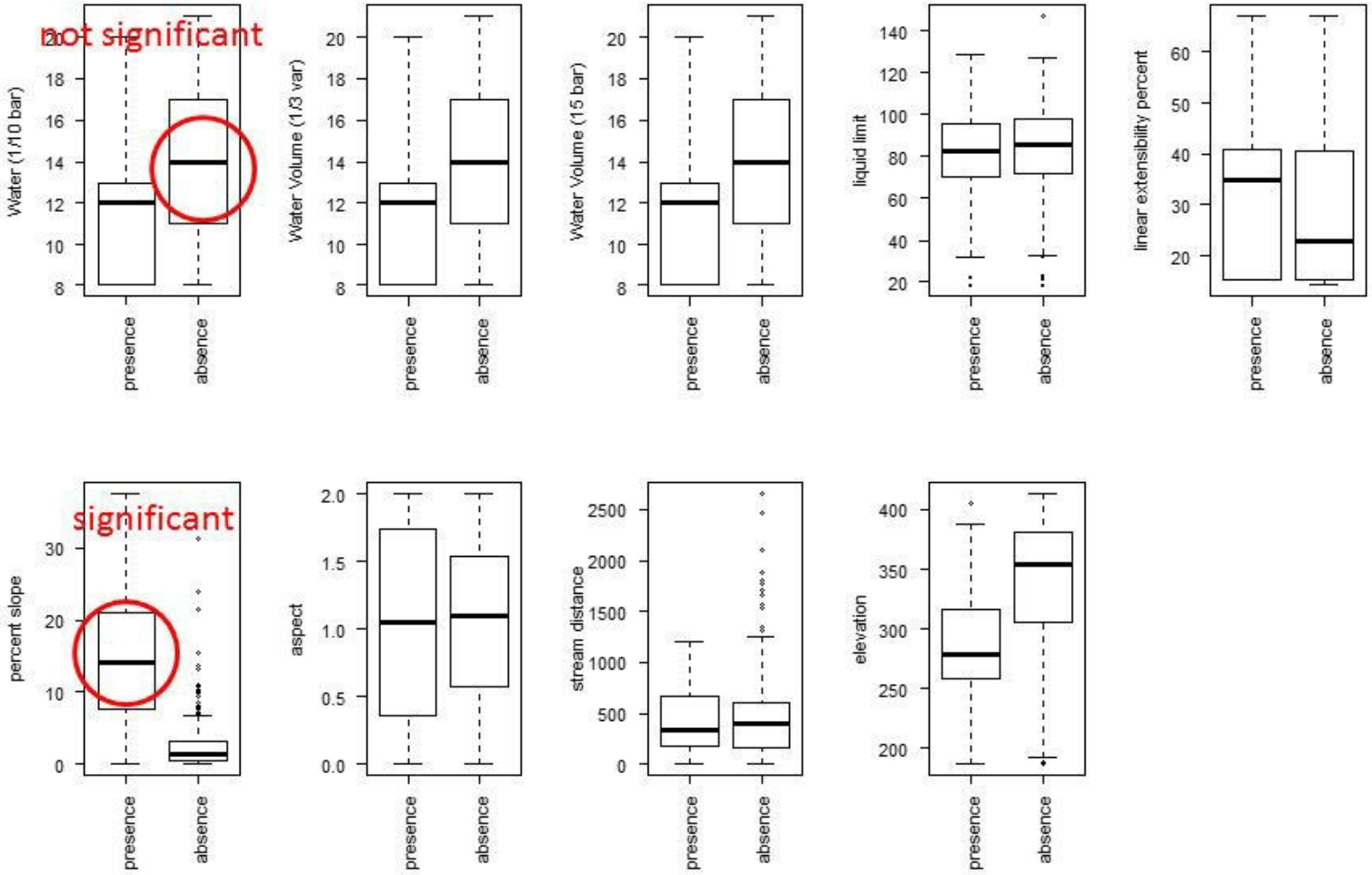
```
Welch Two Sample t-test

data:  data$DBTHIRDBAR_R by data$group
t = 1.8824, df = 4.417, p-value = 0.1261
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -1.714349  9.846566
sample estimates:
mean in group presence  mean in group absence
          77.20000            73.13389
```

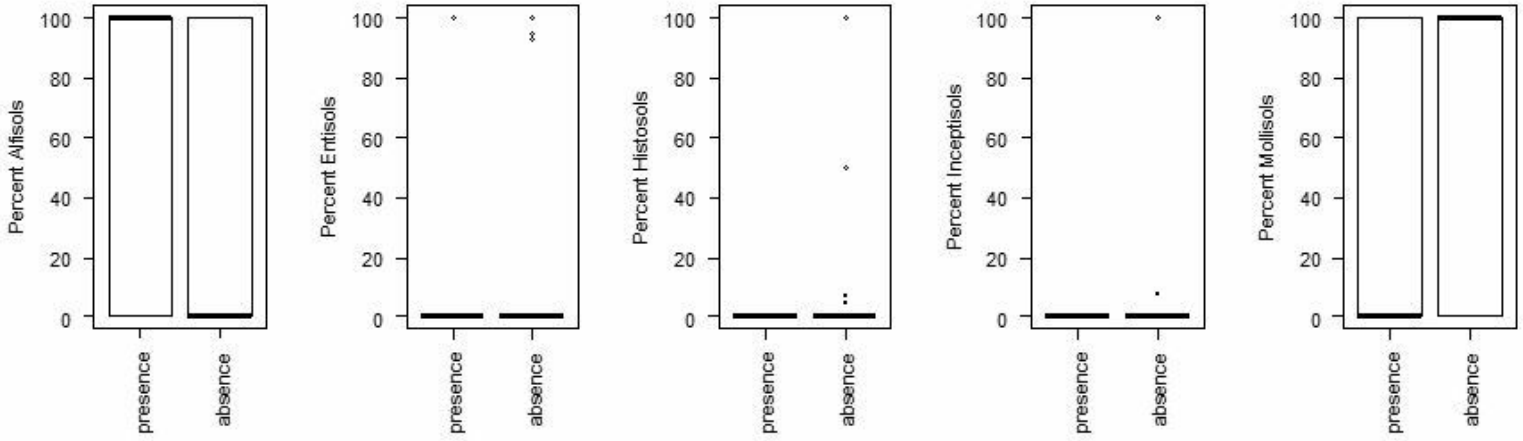
# NCI Dry-Mesic Forest

N presence = 150
N absence = 217




Rock Frag>10 Inches
Rock Frag 3-10 Inches
Sieve no. 4
significant
Sieve no. 10
Sieve no. 40
Sieve no. 200
Plasticity Index
Available Water Capacity
Saturated Hydrologic Conductivity
Bulk Density
presence
absence


not significant
significant
Water (1/10 bar)
Water Volume (1/3 var)
Water Volume (15 bar)
liquid limit
linear extensibility percent
percent slope
aspect
stream distance
elevation
presence
absence


Percent Alfisols
Percent Entisols
Percent Histosols
Percent Inceptisols
Percent Mollisols
0
20
40
60
80
100
presence
absence

```
Welch Two Sample t-test

data:  data$SIEVENO4_R by data$group
t = -5.6971, df = 244.409, p-value =
3.491e-08
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -9.605326 -4.669820
sample estimates:
mean in group presence  mean in group absence
          88.10667            95.24424
```

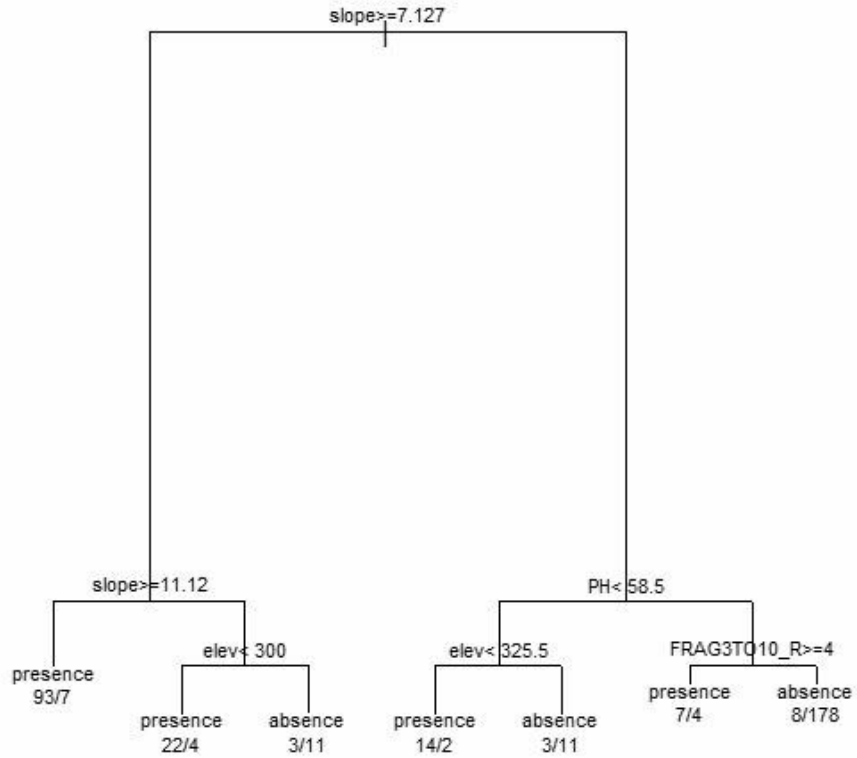
```
Welch Two Sample t-test

data:  data$WTENTHBAR_R by data$group
t = -1.804, df = 12.398, p-value =
0.09556
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -4.886951  0.451237
sample estimates:
mean in group presence  mean in group absence
          11.80000            14.01786
```

```
Welch Two Sample t-test

data:  data$SAND by data$group
t = -4.8112, df = 363.768, p-value =
2.202e-06
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -10.968627  -4.603662
sample estimates:
mean in group presence  mean in group absence
          16.88667            24.67281
```

```
Welch Two Sample t-test

data:  data$slope by data$group
t = 15.2056, df = 204.104, p-value <
2.2e-16
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 10.20408 13.24458
sample estimates:
mean in group presence  mean in group absence
          14.610881            2.886552
```

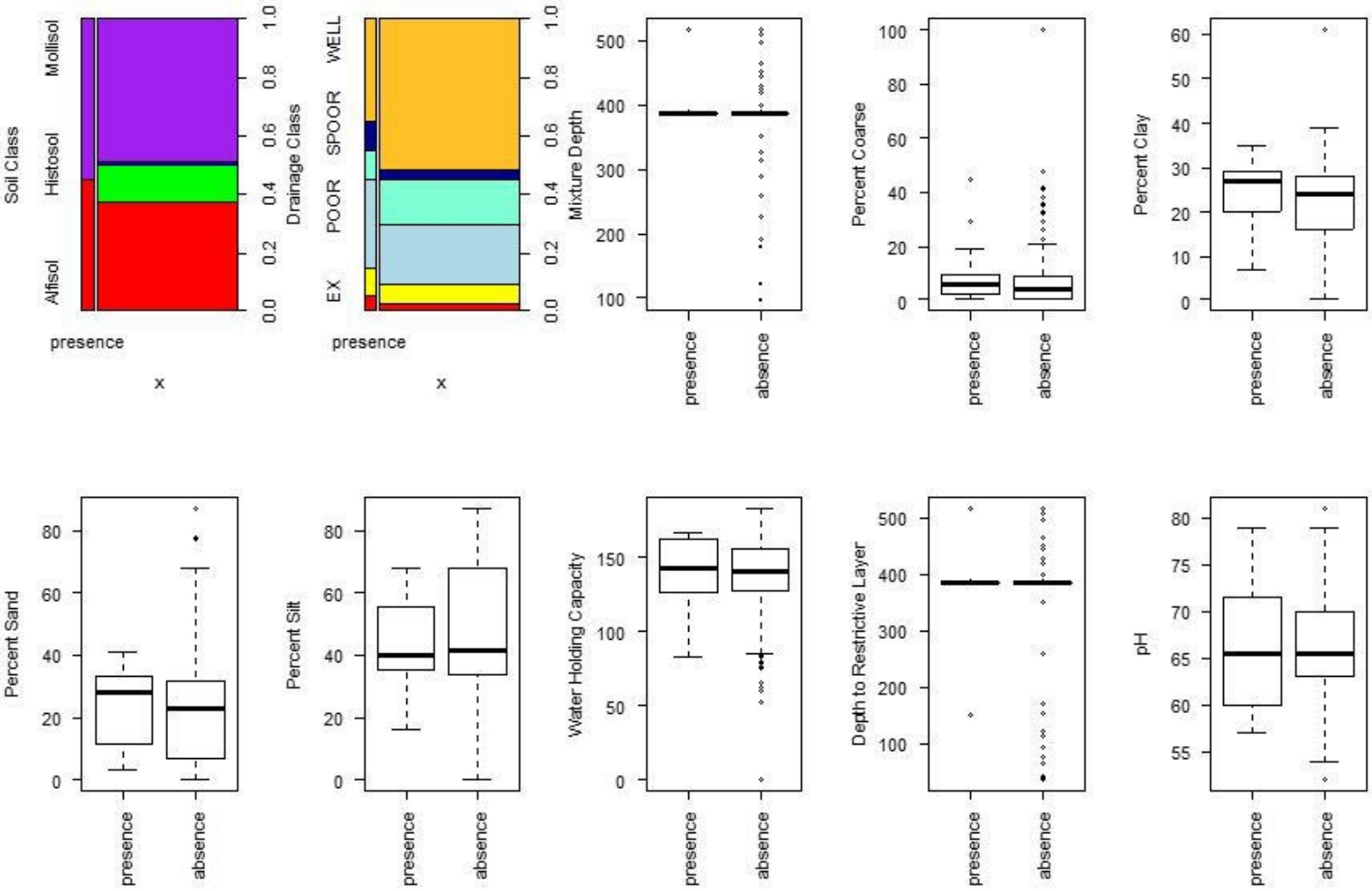
```
Call:
rpart(formula = as.factor(dataSgroup) ~ ., data = data)
 n= 367

       CP nsplit rel error   xerror     xstd
1 0.64000000      0 1.00000000 1.00000000 0.06278432
2 0.04000000      1 0.36000000 0.40000000 0.04723028
3 0.02666667      3 0.28000000 0.31333333 0.04267756
4 0.02000000      5 0.22666667 0.30666667 0.04228703
5 0.01000000      6 0.20666667 0.24666667 0.03845329

Variable importance
     slope FRAG3TO10_R      elev    COARSE SIEVENO4_R
        29        12        12        11        11
SIEVENO10_R        PH    AWC_R      SAND       WHC
         9         6         2         1         1
DRAIN_CLASS      CLAY    DC_SEX SIEVENO40_R P_ALFISOLS
         1         1         1         1         1
 SOIL_CLASS
         1

Node number 1: 367 observations,    complexity param=0.64
  predicted class=absence  expected loss=0.4087193  P(node) =1
    class counts:  150  217
   probabilities: 0.409 0.591
  left son=2 (140 obs) right son=3 (227 obs)
  Primary splits:
     slope      < 7.126729 to the right, improve=85.32051, (0 missing)
     elev       < 325.5   to the left,  improve=39.51652, (0 missing)
     DRAIN_CLASS  splits as  LRRLRRL,    improve=38.52899, (0 missing)
     FRAG3TO10_R  < 4.5     to the right, improve=37.26840, (0 missing)
     DBTHIRDBAR_R < 73.5    to the left,  improve=32.46365, (1 missing)
  Surrogate splits:
     FRAG3TO10_R < 4.5     to the right, agree=0.757, adj=0.364, (0 split)
     COARSE      < 30      to the right, agree=0.733, adj=0.300, (0 split)
     SIEVENO4_R  < 77      to the left,  agree=0.733, adj=0.300, (0 split)
     SIEVENO10_R < 70      to the left,  agree=0.733, adj=0.300, (0 split)
     elev        < 319.5   to the left,  agree=0.725, adj=0.279, (0 split)

Node number 2: 140 observations,    complexity param=0.02666667
  predicted class=presence  expected loss=0.1571429  P(node) =0.3814714
    class counts:  118   22
   probabilities: 0.843 0.157
  left son=4 (100 obs) right son=5 (40 obs)
  Primary splits:
     slope      < 11.12036 to the right, improve=5.315714, (0 missing)
     AWC_R      < 67      to the left,  improve=4.247005, (0 missing)
     CLAY       < 29.5    to the left,  improve=4.002381, (0 missing)
     SAND       < 16.5    to the right, improve=3.085714, (0 missing)
     SILTTOTAL_R < 52.5   to the left,  improve=3.049321, (0 missing)
  Surrogate splits:
     DRAIN_CLASS splits as  LRRLR-L,    agree=0.771, adj=0.200, (0 split)
     CLAY        < 35.5    to the left,  agree=0.750, adj=0.125, (0 split)
     DC_WELL     < 39      to the right, agree=0.743, adj=0.100, (0 split)
```

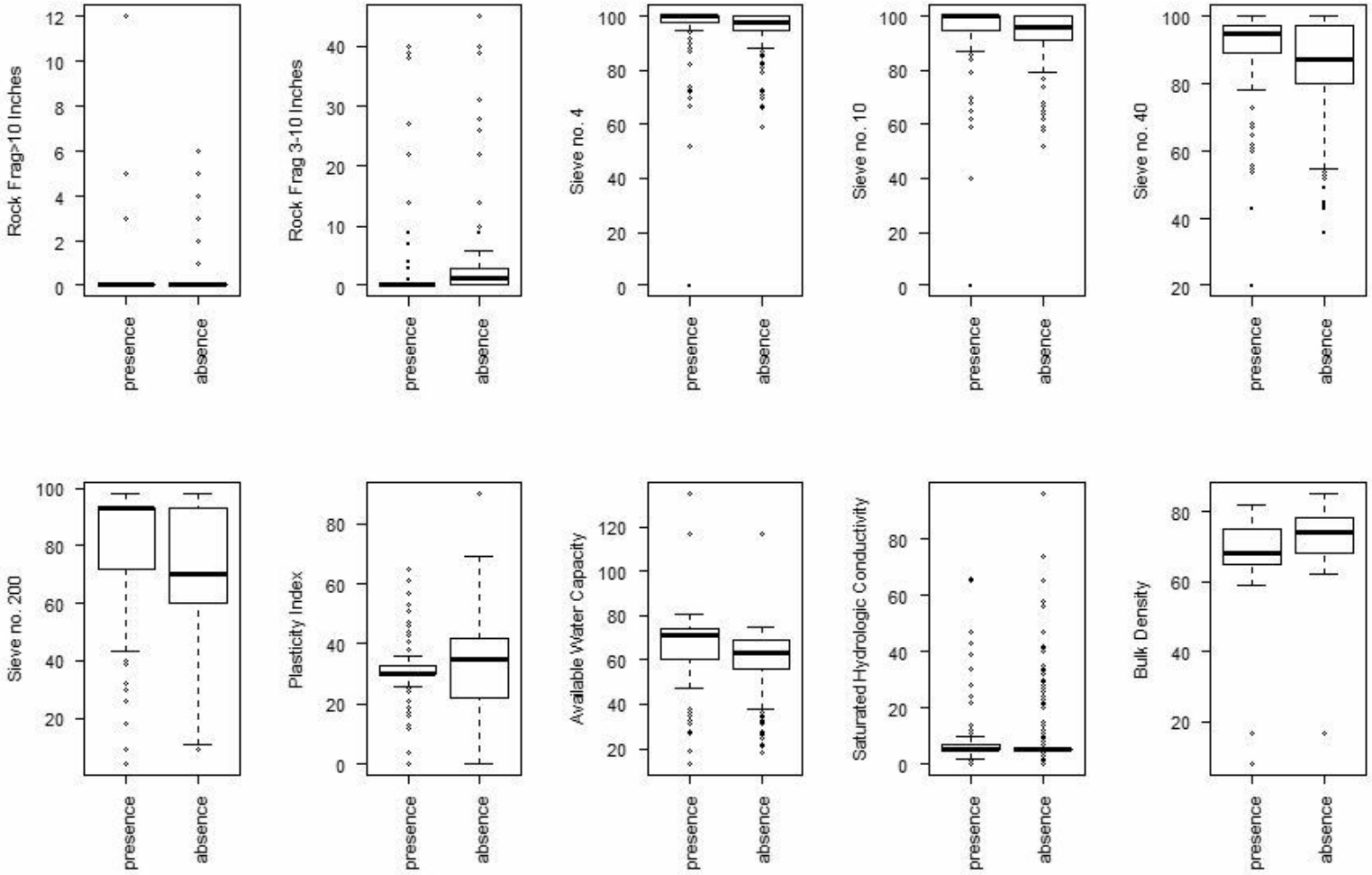
## NCIDryMesic

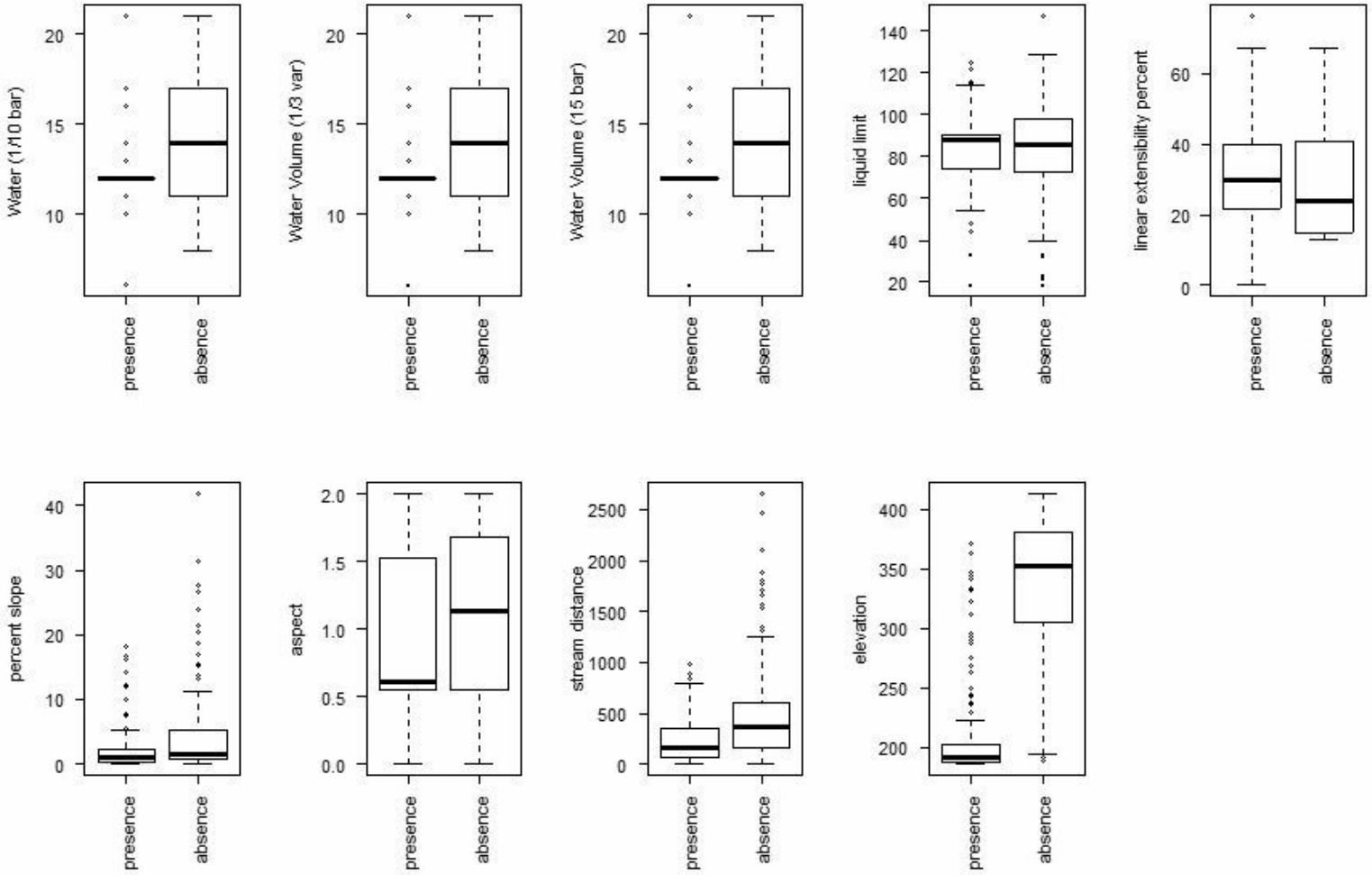


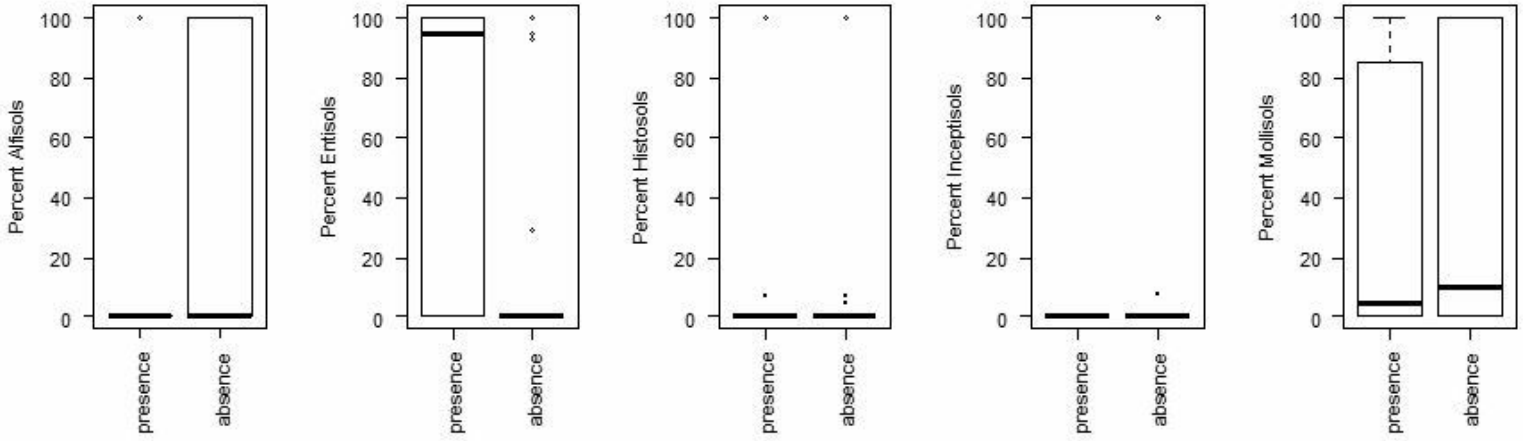
# NCI Floodplain

N presence = 199
N absence = 231

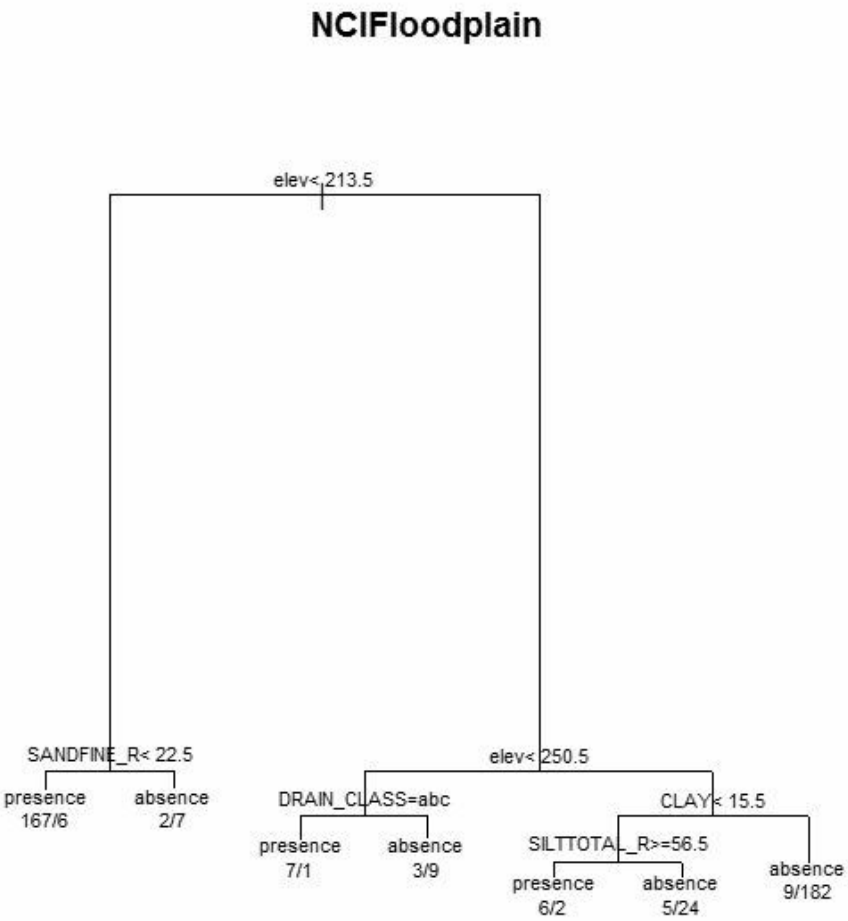



Rock Frag>10 Inches
Rock Frag 3-10 Inches
Sieve no. 4
Sieve no. 10
Sieve no. 40
Sieve no. 200
Plasticity Index
Available Water Capacity
Saturated Hydrologic Conductivity
Bulk Density
presence
absence


Water (1/10 bar)
Water Volume (1/3 var)
Water Volume (15 bar)
liquid limit
linear extensibility percent
percent slope
aspect
stream distance
elevation
presence
absence

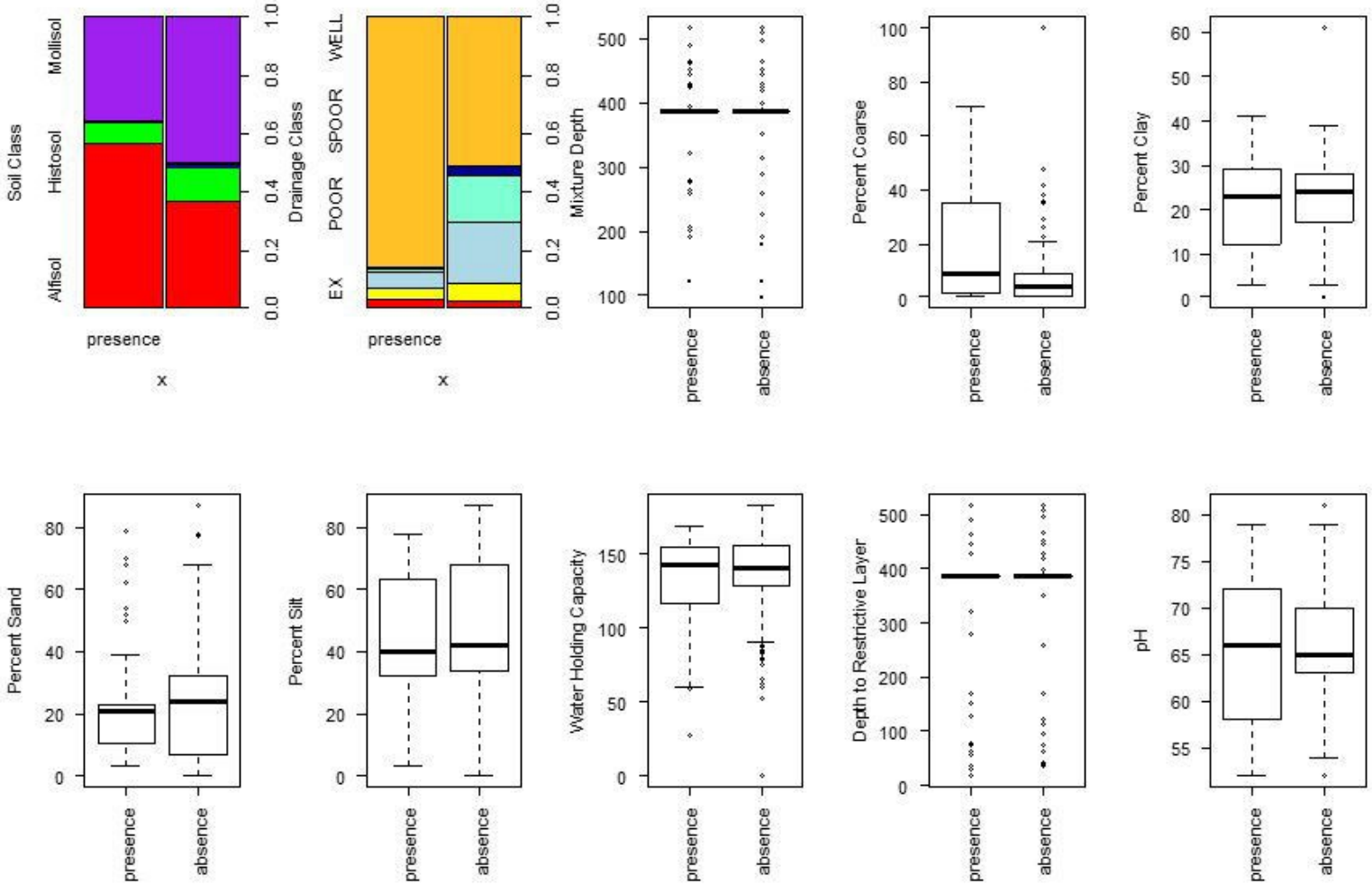

Percent Alfisols
Percent Entisols
Percent Histosols
Percent Inceptisols
Percent Mollisols
0
20
40
60
80
100
presence
absence

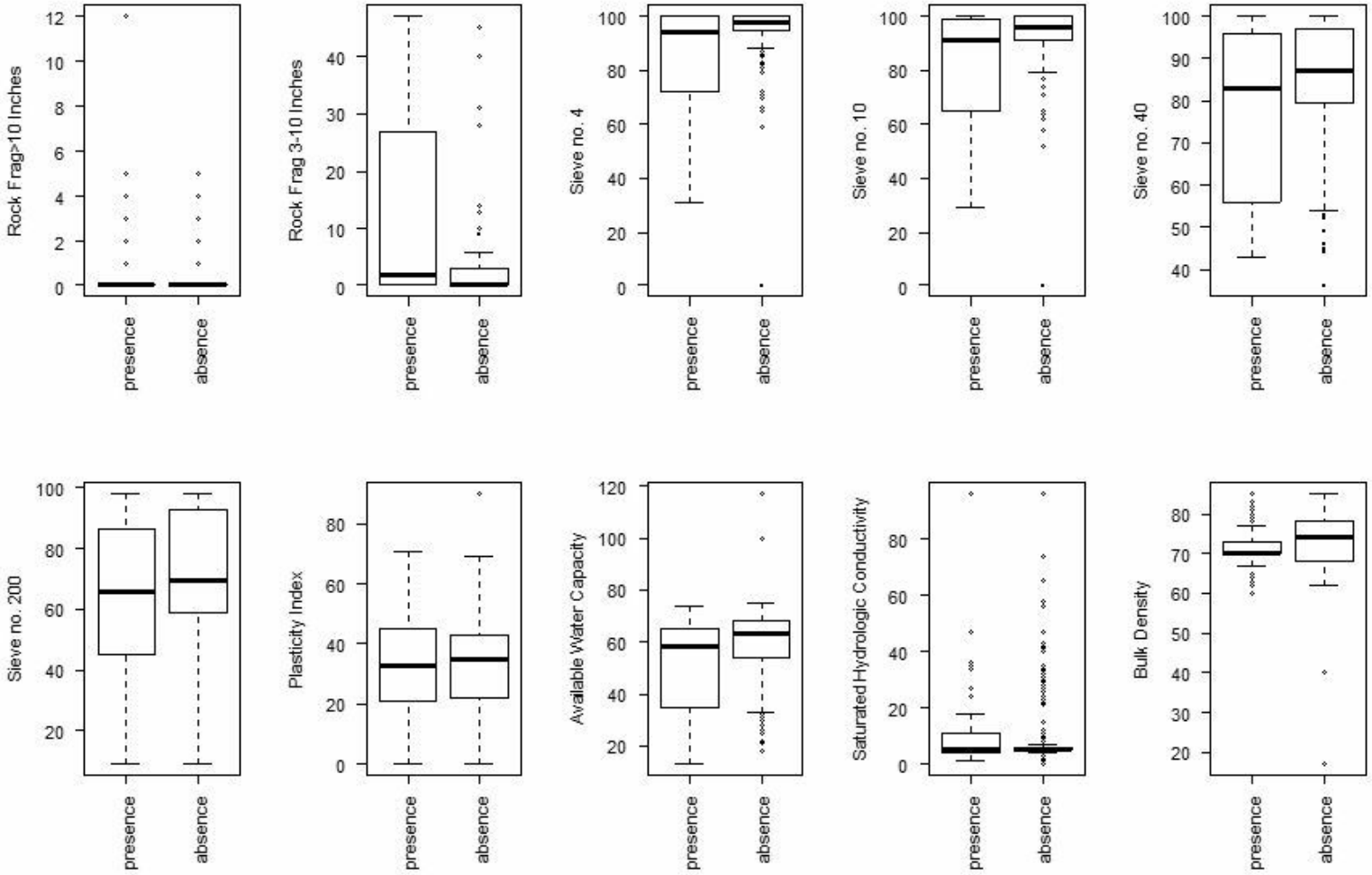
```
Call:
rpart(formula = as.factor(data$group) ~ ., data = data)
 n= 430

       CP nsplit rel error   xerror
1 0.78391960     0 1.0000000 1.0000000
2 0.02512563     1 0.2160804 0.2613065
3 0.01507538     2 0.1909548 0.2763819
4 0.01005025     4 0.1608040 0.2814070
5 0.01000000     6 0.1407035 0.3065327
      xstd
1 0.05195714
2 0.03397506
3 0.03480244
4 0.03507054
5 0.03635723

Variable importance
     elev  P_ENTISOLS    DC_POOR
       25        14         14
 SOIL_CLASS DRAIN_CLASS        PH
       12         11         10
 SANDFINE_R SANDTOTAL_R SILTTOTAL_R
        2          2          1
  SANDVF_R      SAND SIEVENO40_R
        1          1          1
     SILT  SANDMED_R       CLAY
        1          1          1
   DC_WELL   SANDCO_R
        1          1

Node number 1: 430 observations,    complexity param=0.7839196
  predicted class=absence  expected loss=0.4627907  P(node) =1
    class counts:  199  231
   probabilities: 0.463 0.537
  left son=2 (182 obs) right son=3 (248 obs)
  Primary splits:
     elev      < 213.5 to the left,  improve=136.92450, (0 missing)
     P_ENTISOLS < 2.5   to the right, improve= 83.87339, (0 missing)
     SOIL_CLASS  splits as  RLLRR,    improve= 66.69948, (0 missing)
     DC_POOR    < 4     to the right, improve= 54.97808, (0 missing)
     DRAIN_CLASS splits as  RLL-RLR,  improve= 50.69502, (0 missing)
  Surrogate splits:
     P_ENTISOLS < 2.5   to the right, agree=0.830, adj=0.599, (0 split)
     DC_POOR    < 4     to the right, agree=0.821, adj=0.577, (0 split)
     SOIL_CLASS  splits as  RLLRR,    agree=0.800, adj=0.527, (0 split)
     DRAIN_CLASS splits as  RLL-RRR,  agree=0.758, adj=0.429, (0 split)
     PH         < 72.5  to the right, agree=0.753, adj=0.418, (0 split)

Node number 2: 182 observations,    complexity param=0.02512563
  predicted class=presence  expected loss=0.07142857  P(node) =0.4232558
    class counts:  169   13
   probabilities: 0.929 0.071
```

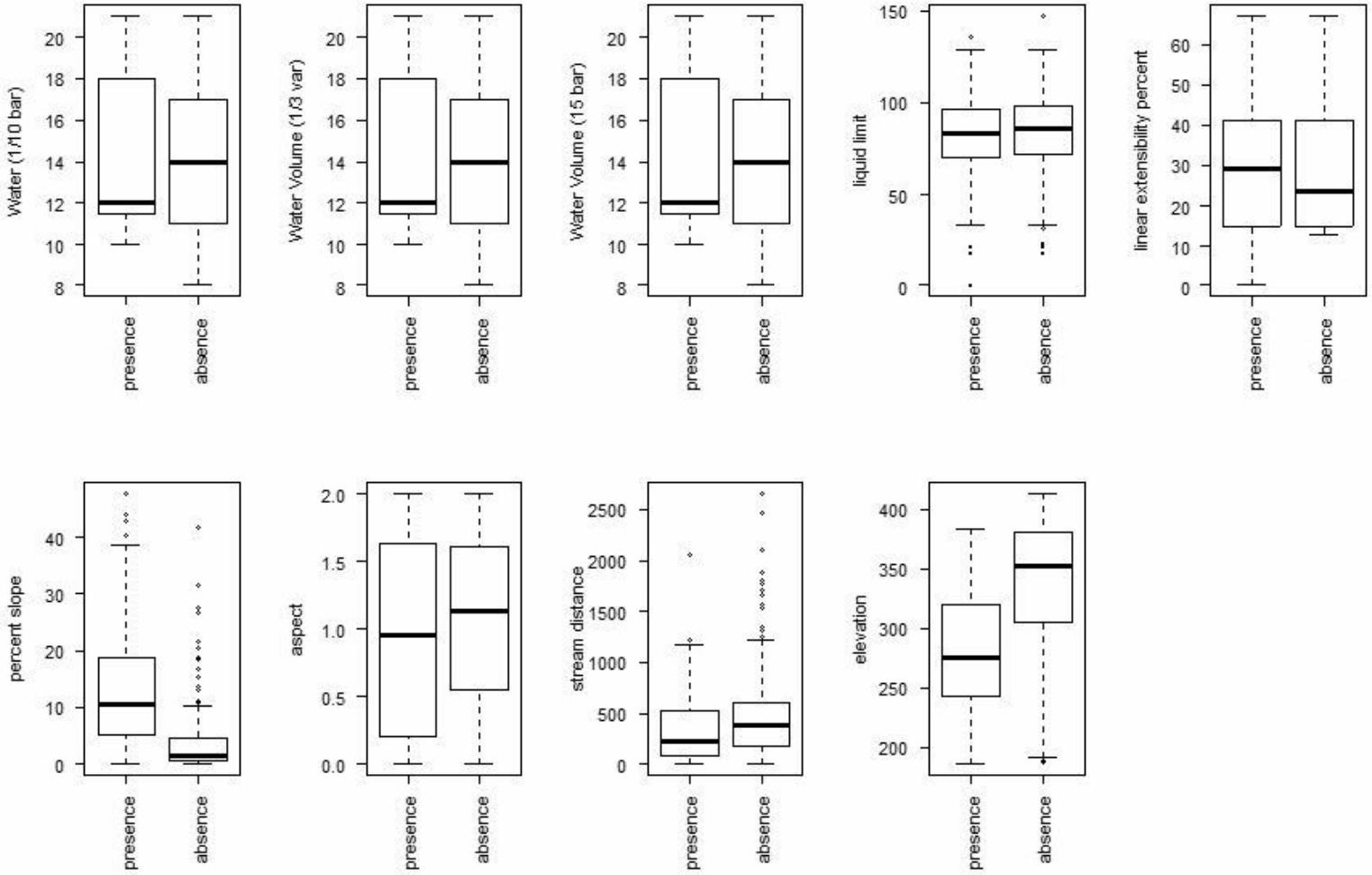


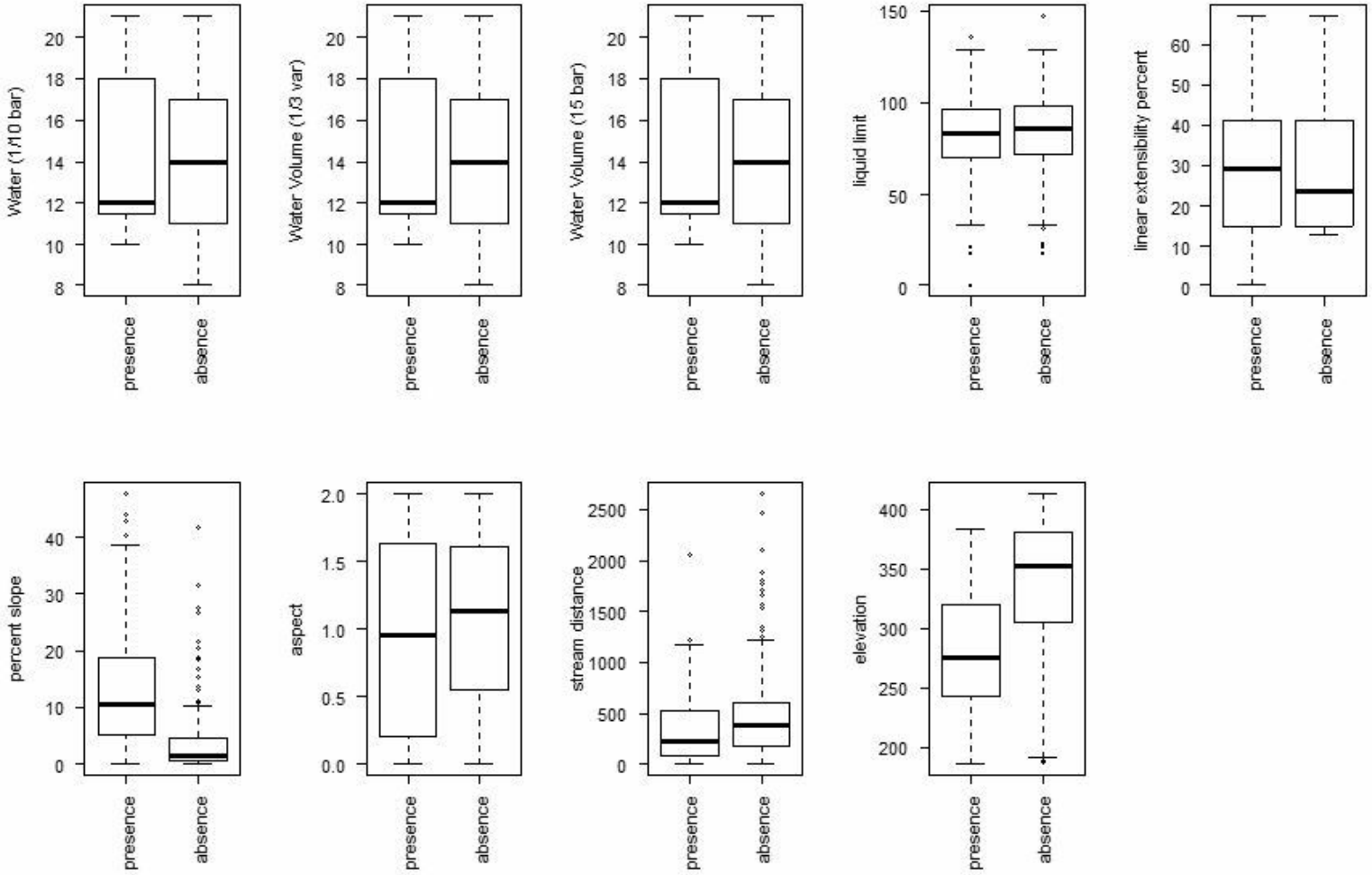
# NCI Maple-Basswood Forest

N presence = 243
N absence = 233



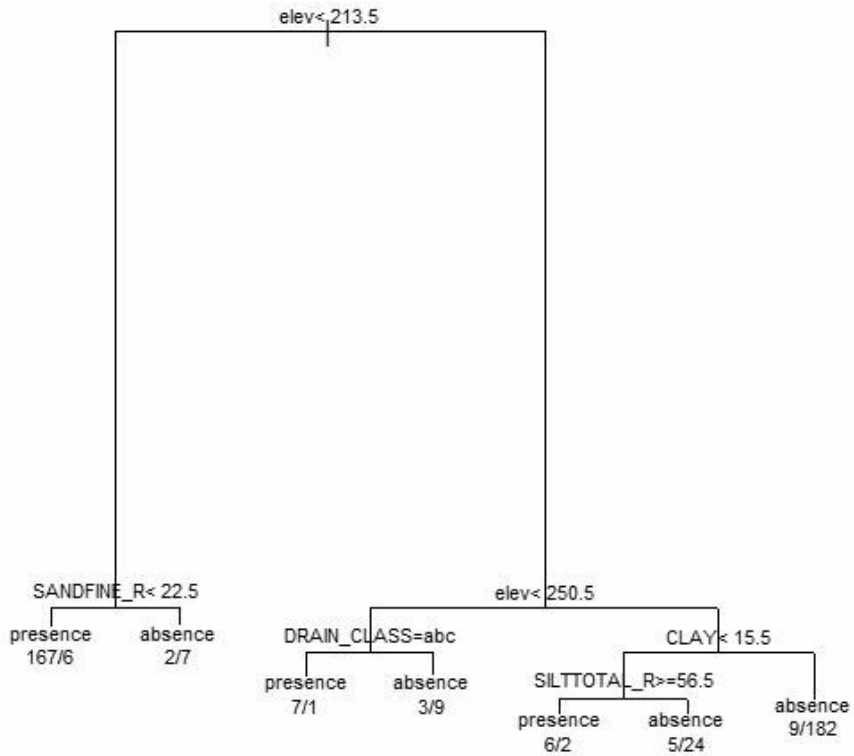

Rock Frag>10 Inches
Rock Frag 3-10 Inches
Sieve no. 4
Sieve no. 10
Sieve no. 40
Sieve no. 200
Plasticity Index
Available Water Capacity
Saturated Hydrologic Conductivity
Bulk Density
presence
absence


Water (1/10 bar)
Water Volume (1/3 var)
Water Volume (15 bar)
liquid limit
linear extensibility percent
percent slope
aspect
stream distance
elevation
presence
absence


Water (1/10 bar)
Water Volume (1/3 var)
Water Volume (15 bar)
liquid limit
linear extensibility percent
percent slope
aspect
stream distance
elevation
presence
absence

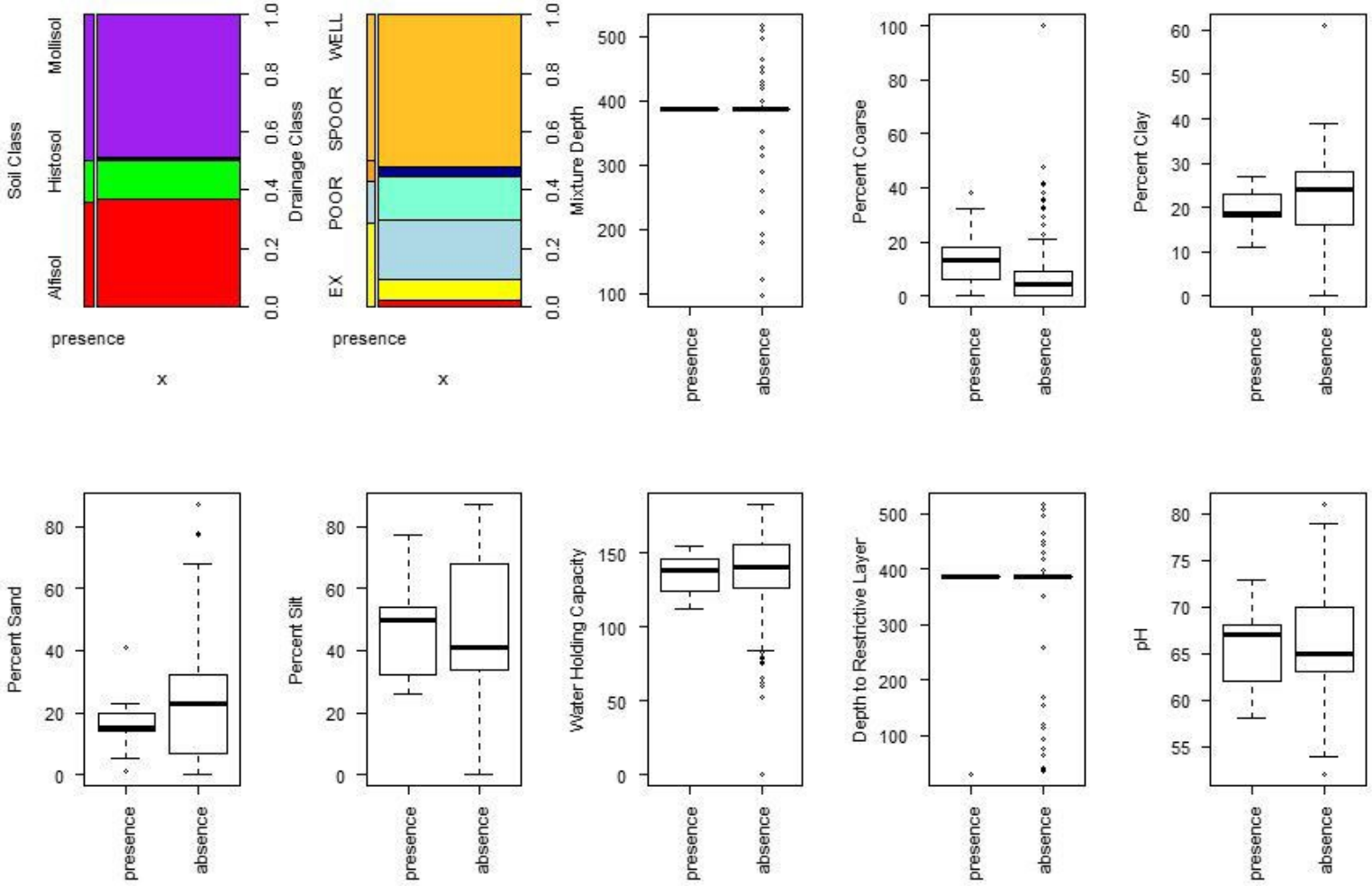
```
Call:
rpart(formula = as.factor(data$group) ~ ., data = data)
 n= 476

       CP nsplit rel error   xerror
1 0.54077253     0 1.0000000 1.0515021
2 0.03648069     1 0.4592275 0.4849785
3 0.03433476     3 0.3862661 0.5193133
4 0.01931330     5 0.3175966 0.4420601
5 0.01000000     7 0.2789700 0.4248927
      xstd
1 0.04679825
2 0.03984127
3 0.04077063
4 0.03855792
5 0.03800396

Variable importance
     slope       elev  DRAIN_CLASS
        23         16            9
      SAND  DBTHIRDBAR_R     DC_WELL
         9          8            8
   strmdist      AWC_R  FRAG3TO10_R
         7          3            2
 SIEVENO4_R       SILT     COARSE
         2          2            2
SIEVENO10_R   asp_beers       CLAY
         2          1            1
   RESDEPTH    DC_POOR    SANDCO_R
         1          1            1
   SANDVC_R        WHC    MXDEPTH
         1          1            1
```

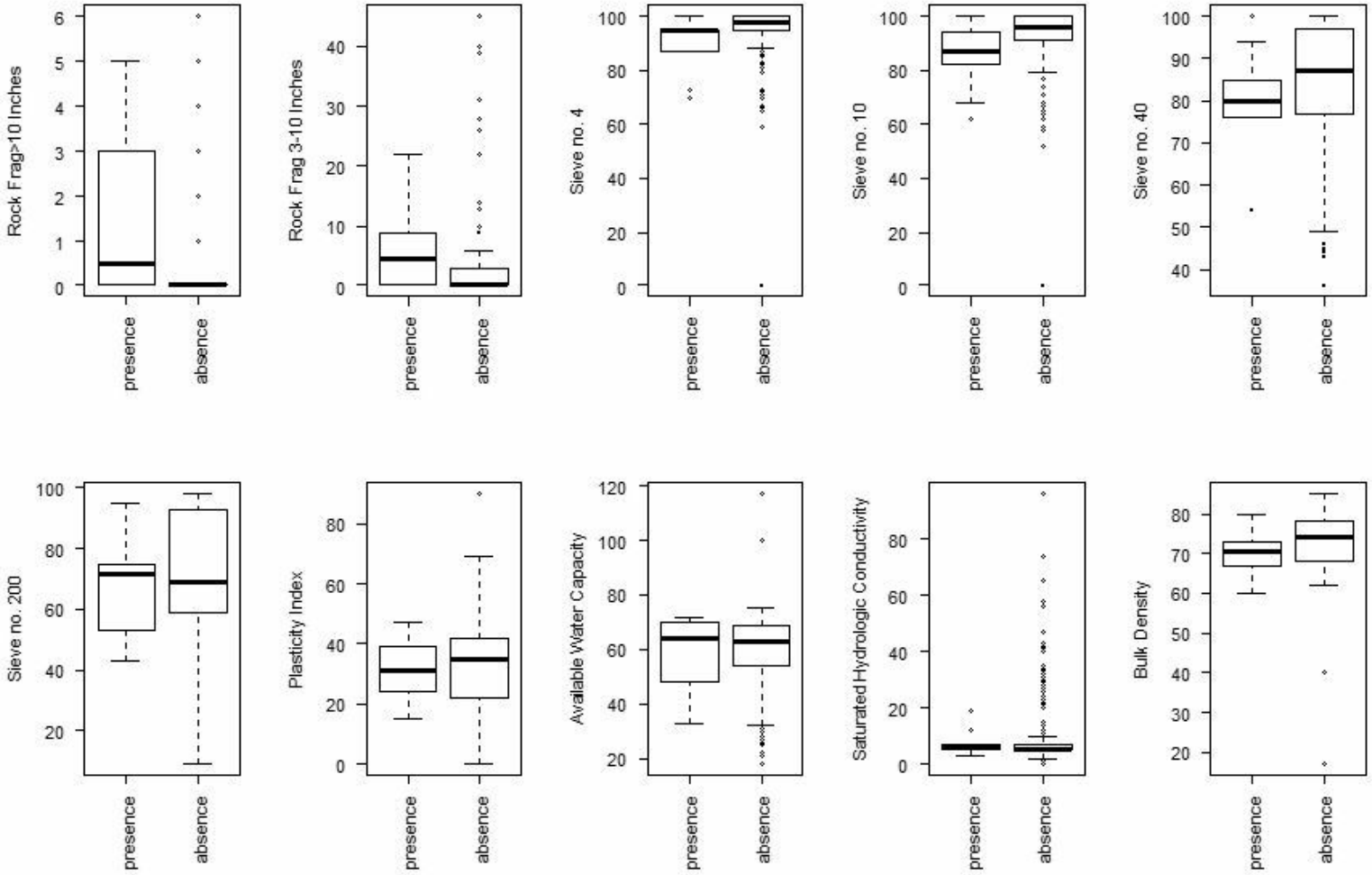
## NCIFloodplain

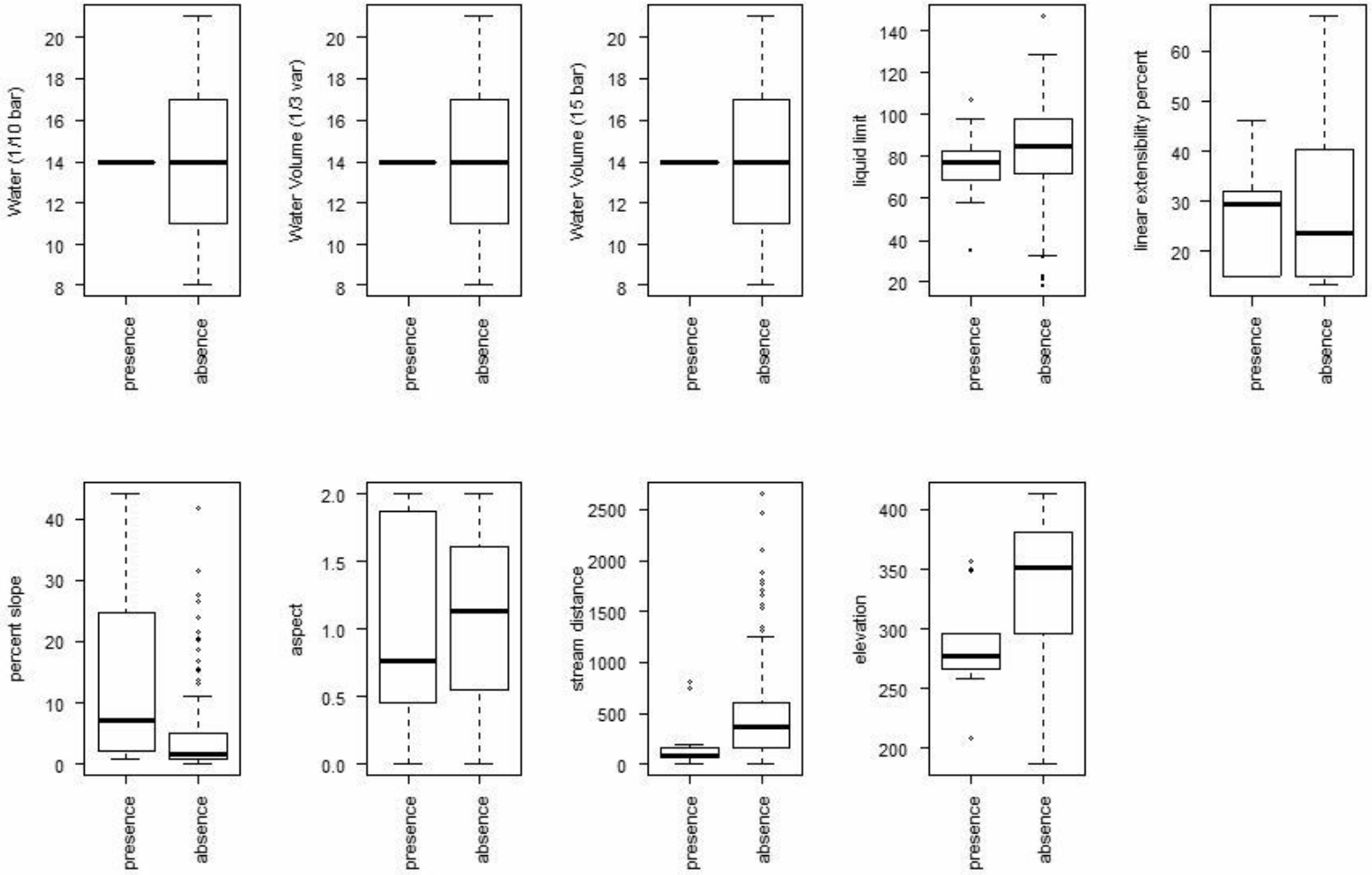


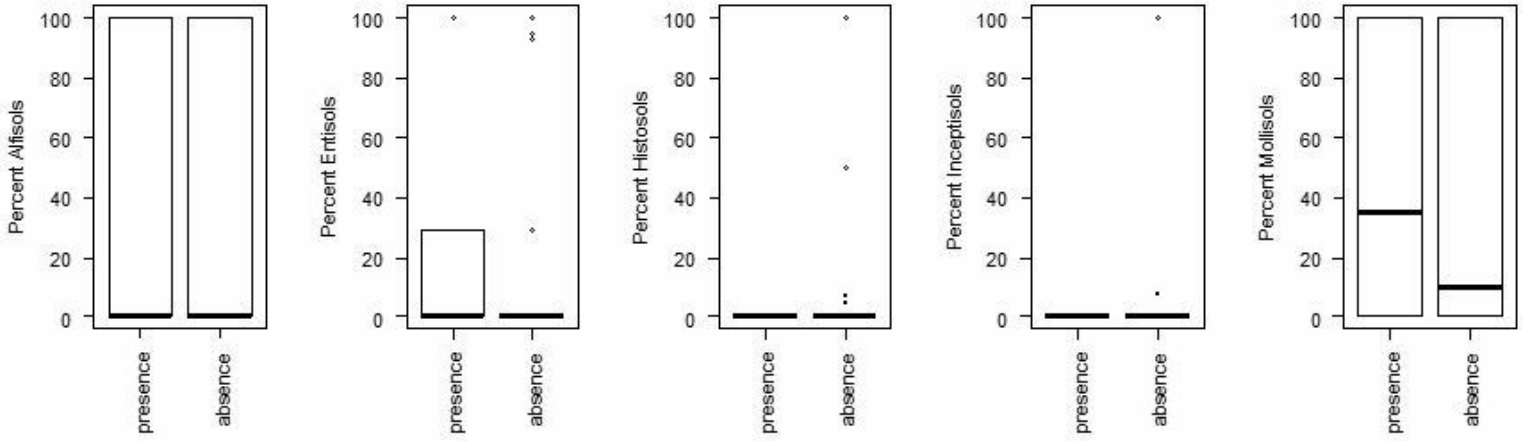
# NCI Rich Swamp

N presence = 14
N absence = 241



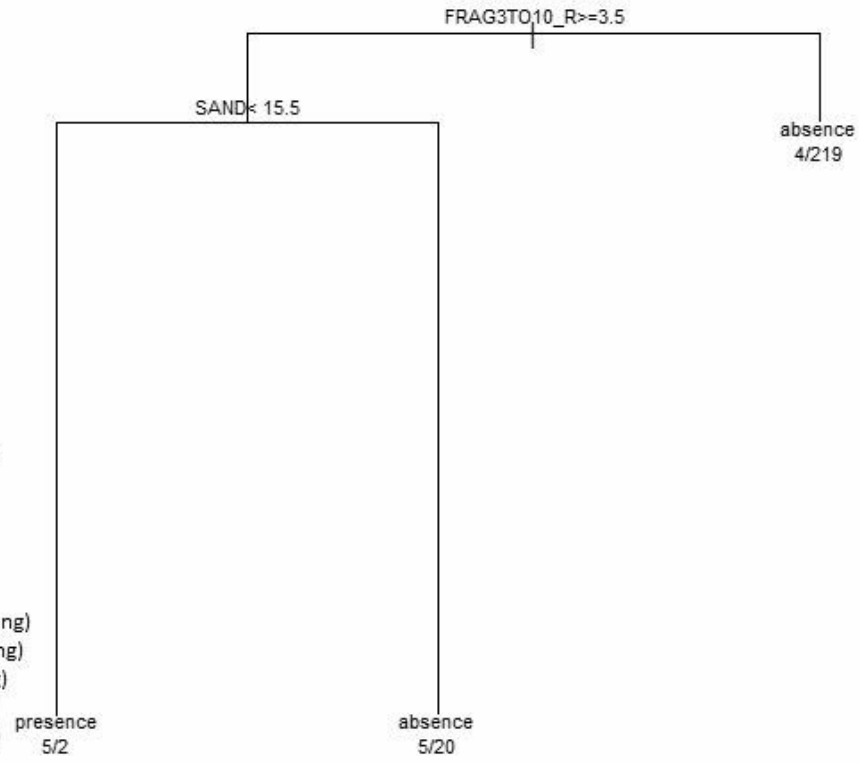

Rock Frag>10 Inches
Rock Frag 3-10 Inches
Sieve no. 4
Sieve no. 10
Sieve no. 40
Sieve no. 200
Plasticity Index
Available Water Capacity
Saturated Hydrologic Conductivity
Bulk Density
presence
absence


Water (1/10 bar)
Water Volume (1/3 var)
Water Volume (15 bar)
liquid limit
linear extensibility percent
percent slope
aspect
stream distance
elevation
presence
absence


Percent Alfisols
Percent Entisols
Percent Histosols
Percent Inceptisols
Percent Mollisols
0
20
40
60
80
100
presence
absence

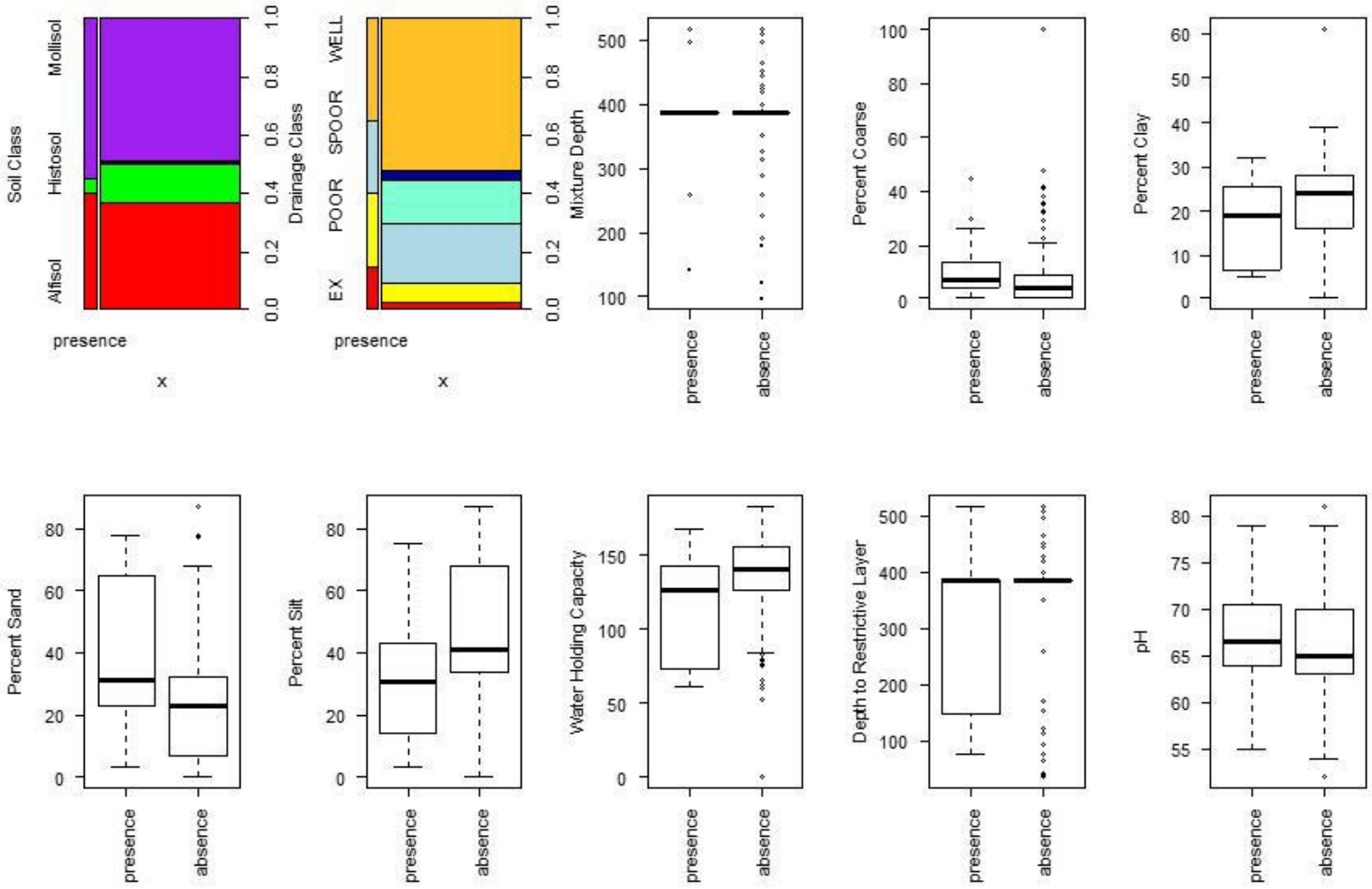
```
Call:
rpart(formula = as.factor(data$group) ~ ., data = data)
 n= 255

      CP nsplit rel error   xerror
1 0.1071429     0 1.0000000 1.000000
2 0.0100000     2 0.7857143 1.357143
     xstd
1 0.2598211
2 0.2995261

Variable importance
 FRAG3TO10_R   SIEVENO4_R        SAND
          19          15          11
      COARSE  SIEVENO10_R        SILT
           8           8           8
       slope      DC_MOD  DRAIN_CLASS
           6           6           6
SIEVENO200_R   FRAGGT10_R
           6           6

Node number 1: 255 observations,   complexity param=0.1071429
  predicted class=absence  expected loss=0.05490196 P(node) =1
    class counts:   14  241
   probabilities: 0.055 0.945
  left son=2 (32 obs) right son=3 (223 obs)
  Primary splits:
     FRAG3TO10_R < 3.5     to the right, improve=4.856243, (0 missing)
     FRAGGT10_R  < 2.5     to the right, improve=4.676837, (0 missing)
     slope       < 24.31818 to the right, improve=2.831309, (0 missing)
     strmdist    < 101.5   to the left,  improve=2.618843, (0 missing)
     COARSE      < 12.5    to the right, improve=2.079291, (0 missing)
  Surrogate splits:
     COARSE      < 30.5    to the right, agree=0.929, adj=0.438, (0 split)
     SIEVENO4_R  < 76      to the left,  agree=0.929, adj=0.438, (0 split)
     SIEVENO10_R < 69.5    to the left,  agree=0.929, adj=0.438, (0 split)
     slope       < 10.21329 to the right, agree=0.918, adj=0.344, (0 split)
     FRAGGT10_R  < 1.5     to the right, agree=0.914, adj=0.312, (0 split)
```

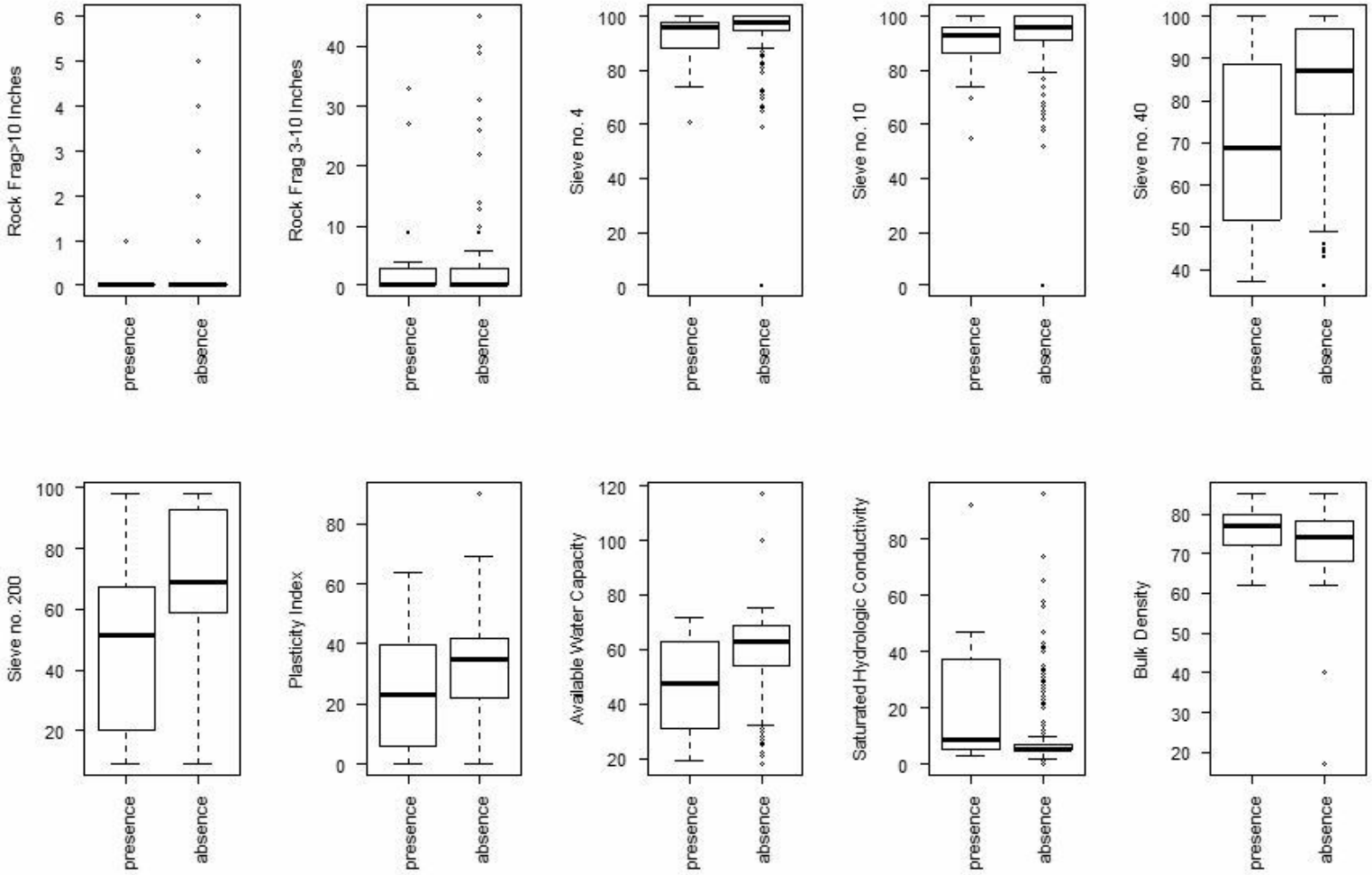
**NCIRichSwamp**

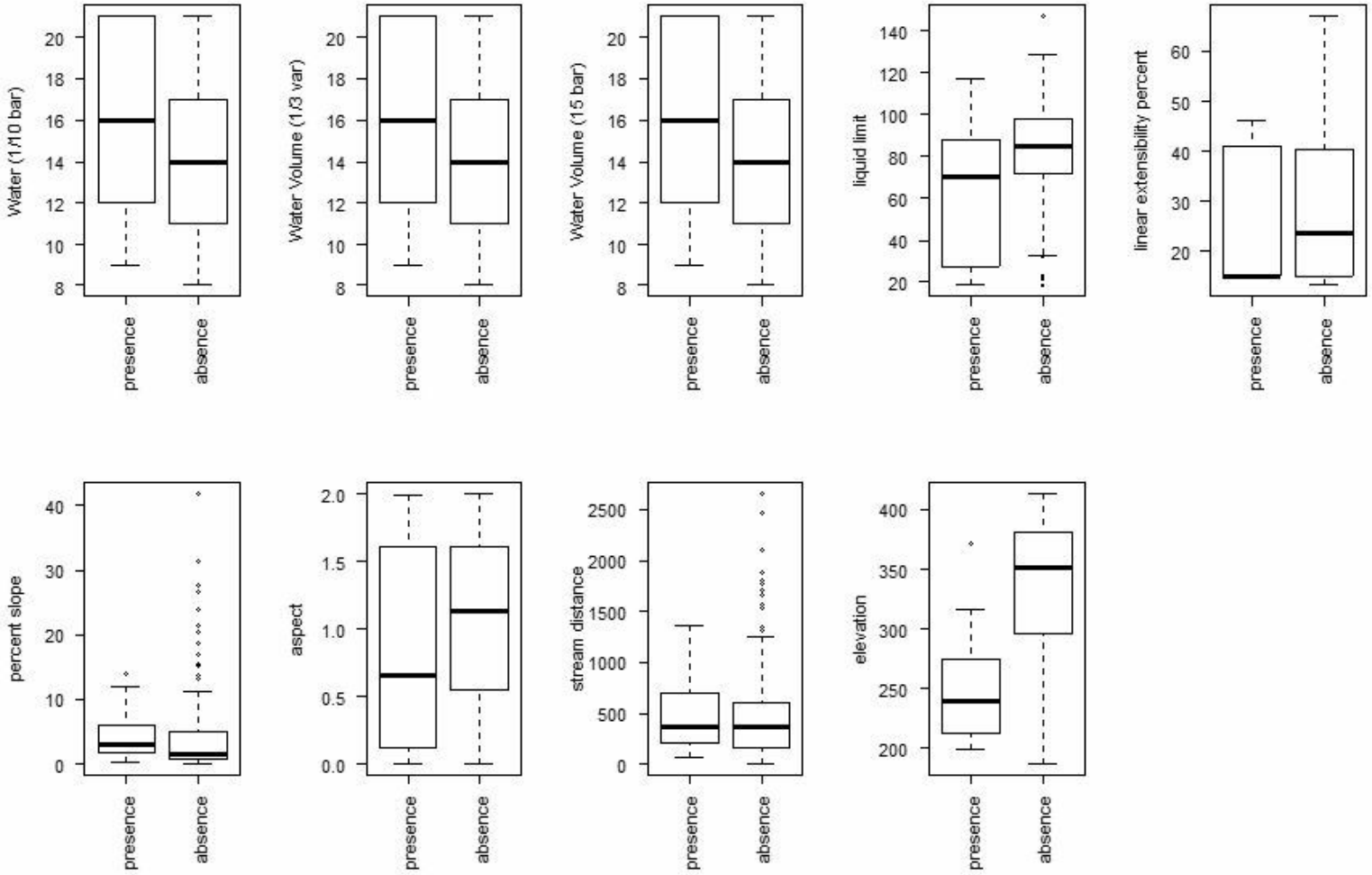


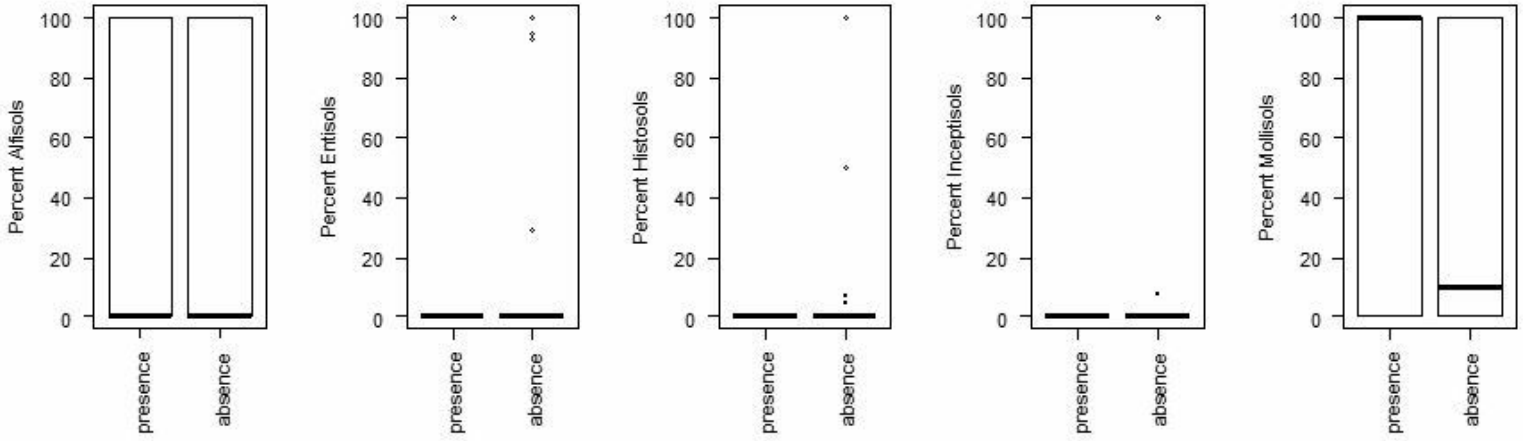
# NCI Sand & Gravel Prairie

N presence = 20
N absence = 241

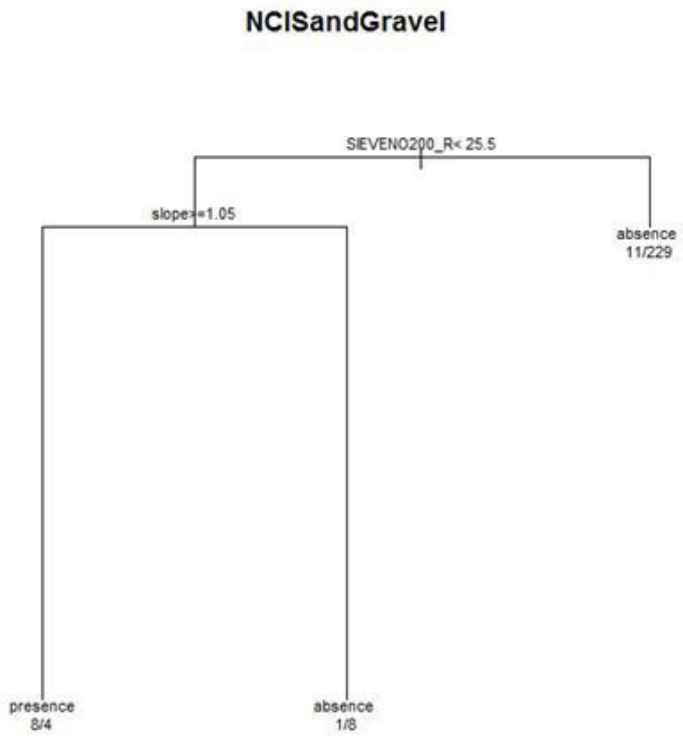



Rock Frag>10 Inches
Rock Frag 3-10 Inches
Sieve no. 4
Sieve no. 10
Sieve no. 40
Sieve no. 200
Plasticity Index
Available Water Capacity
Saturated Hydrologic Conductivity
Bulk Density
presence
absence


Water (1/10 bar)
Water Volume (1/3 var)
Water Volume (15 bar)
liquid limit
linear extensibility percent
percent slope
aspect
stream distance
elevation
presence
absence

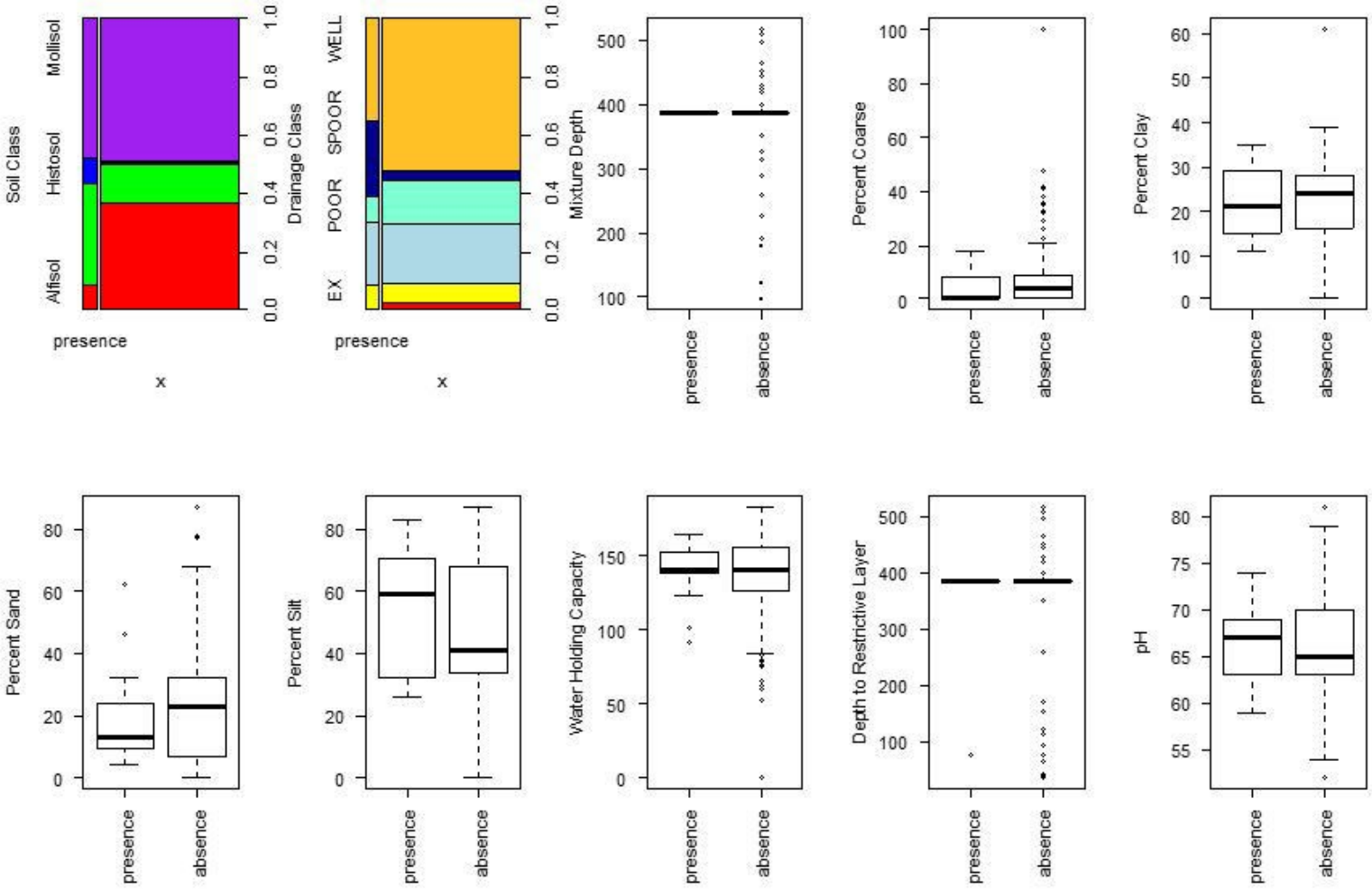

Percent Alfisols
Percent Entisols
Percent Histosols
Percent Inceptisols
Percent Mollisols
0
20
40
60
80
100
presence
absence

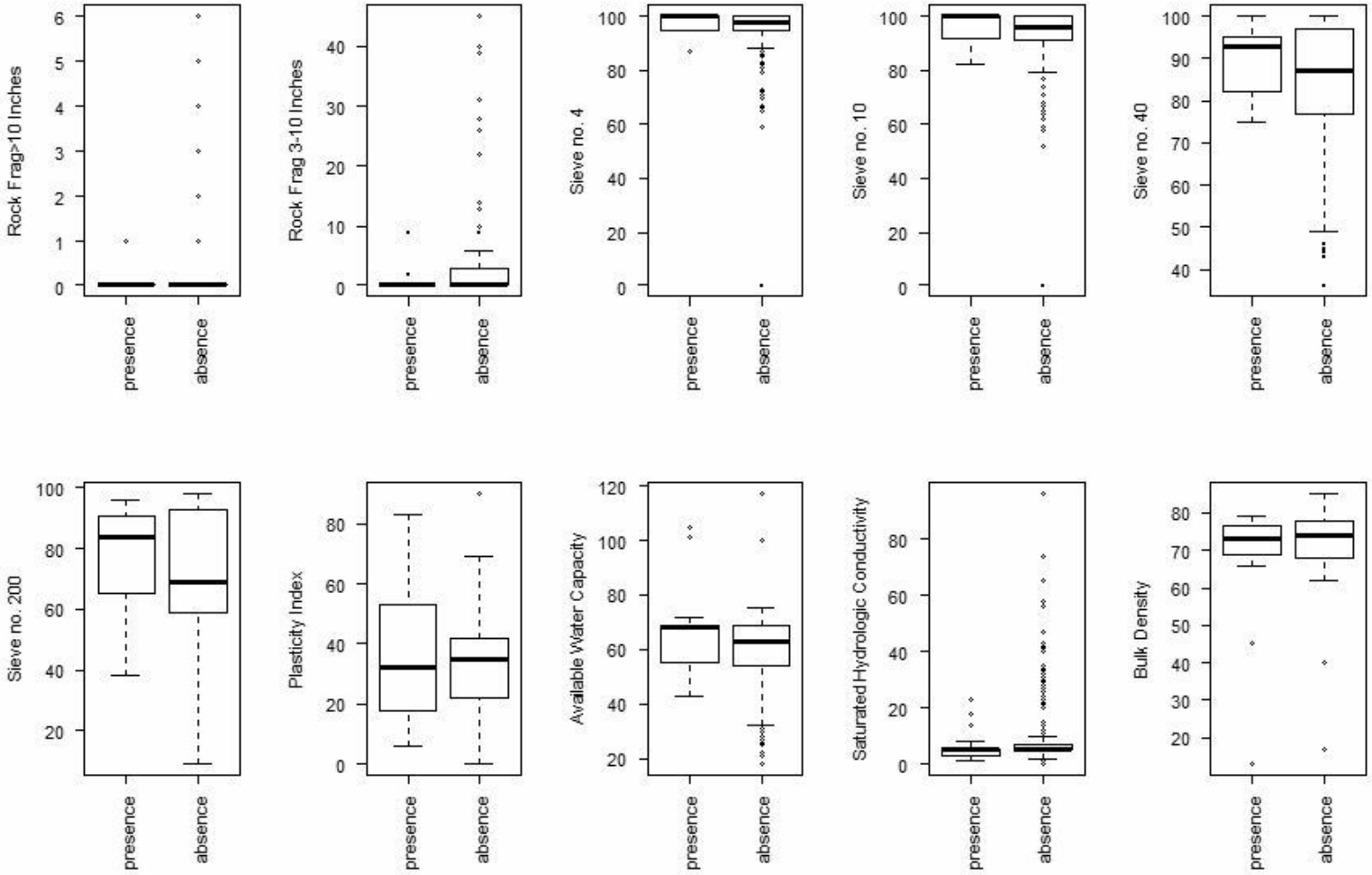
```
Call:
rpart(formula = as.factor(data$group) ~ ., data = data)
 n= 261

   CP nsplit rel error xerror     xstd
1 0.10     0      1.0   1.0 0.2148688
2 0.01     2      0.8   0.9 0.2046865

Variable importance
SIEVENO200_R       AWC_R        SILT
          15          13          13
         WHC        LL_R        SAND
          13          12          12
       slope  P_MOLLISOLS   SOIL_CLASS
           8           4           4
  SANDMED_R   SIEVENO4_R  SIEVENO40_R
           3           3           3

Node number 1: 261 observations,   complexity param=0.1
  predicted class=absence  expected loss=0.07662835 P(node) =1
    class counts:   20  241
   probabilities: 0.077 0.923
  left son=2 (21 obs) right son=3 (240 obs)
  Primary splits:
     SIEVENO200_R < 25.5    to the left,  improve=6.031410, (1 mis
     LL_R         < 41.5    to the left,  improve=6.031410, (1 missing)
     elev         < 280.5   to the left,  improve=5.654515, (0 missing)
     SILT         < 16.5    to the left,  improve=5.311055, (0 missing)
     AWC_R        < 32.5    to the left,  improve=4.987652, (1 missing
  Surrogate splits:
     SILT  < 15.5    to the left,  agree=0.988, adj=0.85, (1 split)
     WHC   < 83.5    to the left,  agree=0.988, adj=0.85, (0 split)
     AWC_R < 32.5    to the left,  agree=0.988, adj=0.85, (0 split)
     SAND  < 60.5    to the right, agree=0.985, adj=0.80, (0 split)
     LL_R  < 41.5    to the left,  agree=0.985, adj=0.80, (0 split)
```

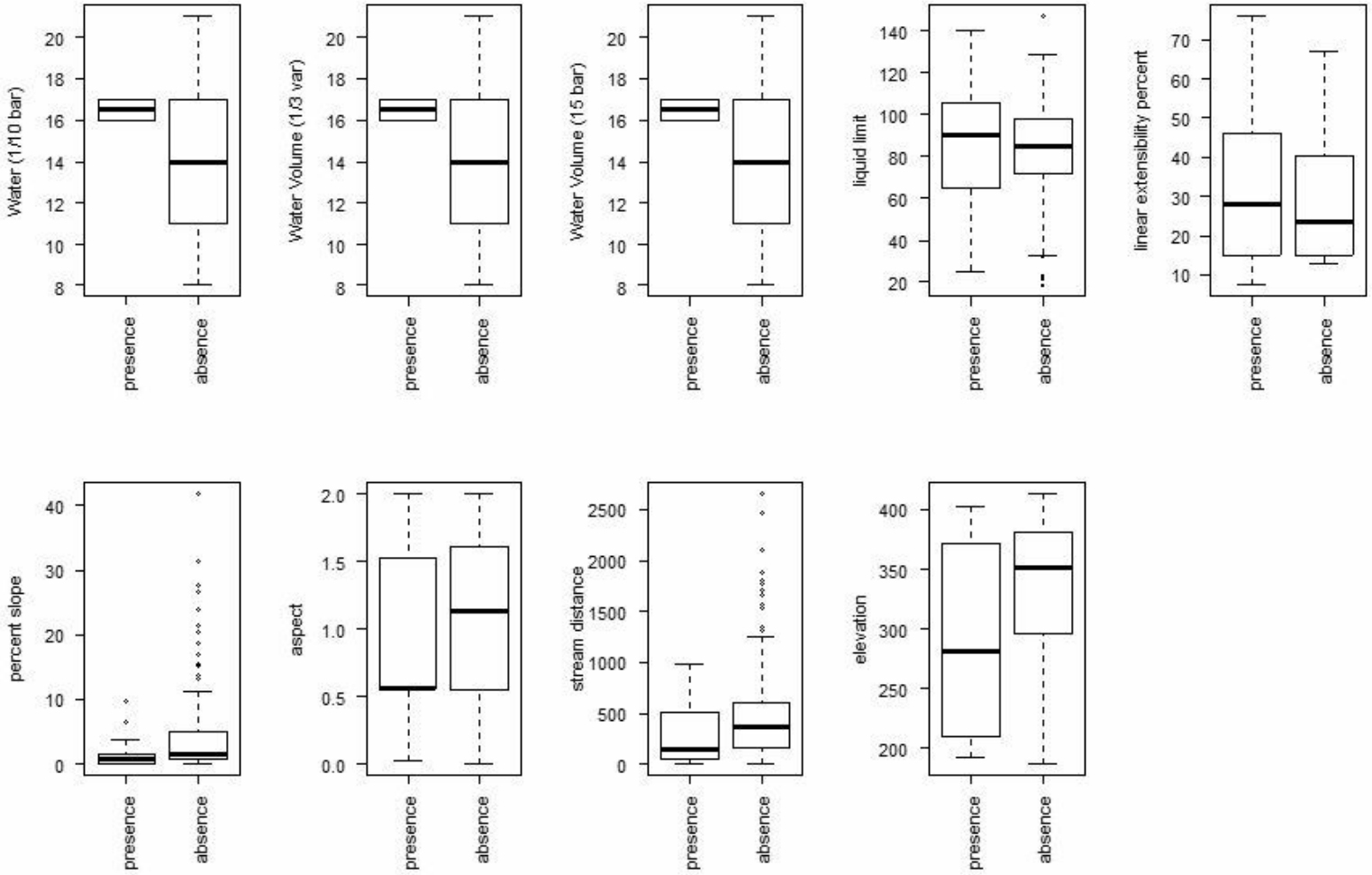


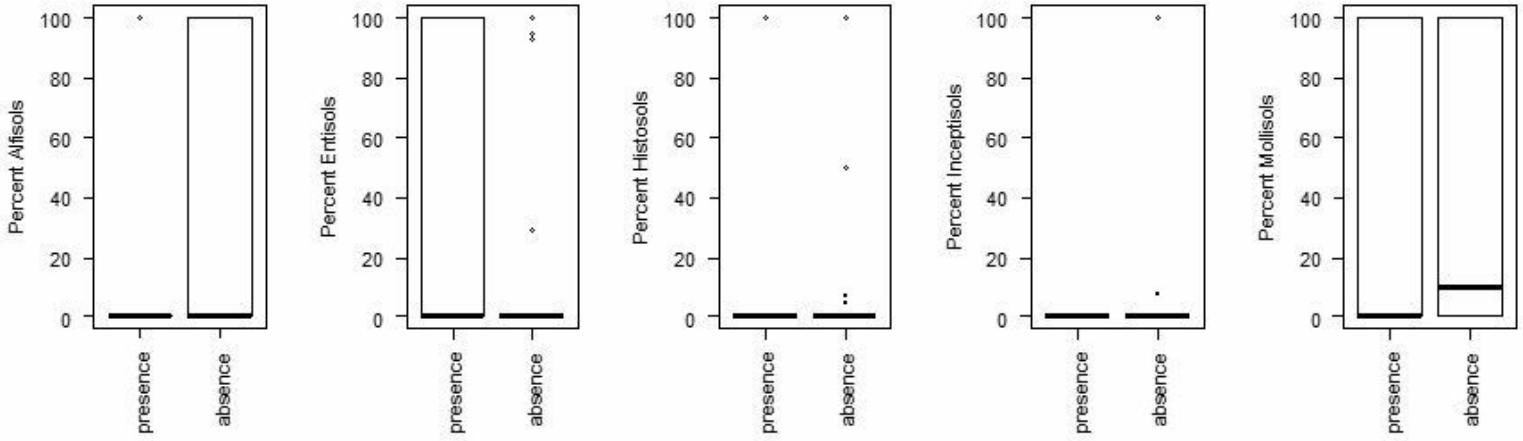
# NCI Wet Meadow

N presence = 23
N absence = 241

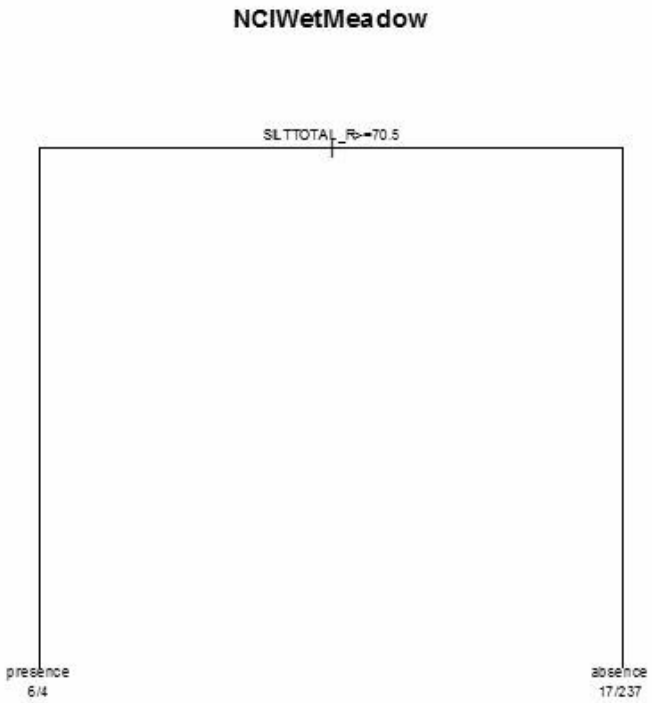



Rock Frag>10 Inches
Rock Frag 3-10 Inches
Sieve no. 4
Sieve no. 10
Sieve no. 40
Sieve no. 200
Plasticity Index
Available Water Capacity
Saturated Hydrologic Conductivity
Bulk Density
presence
absence


Water (1/10 bar)
Water Volume (1/3 var)
Water Volume (15 bar)
liquid limit
linear extensibility percent
percent slope
aspect
stream distance
elevation
presence
absence

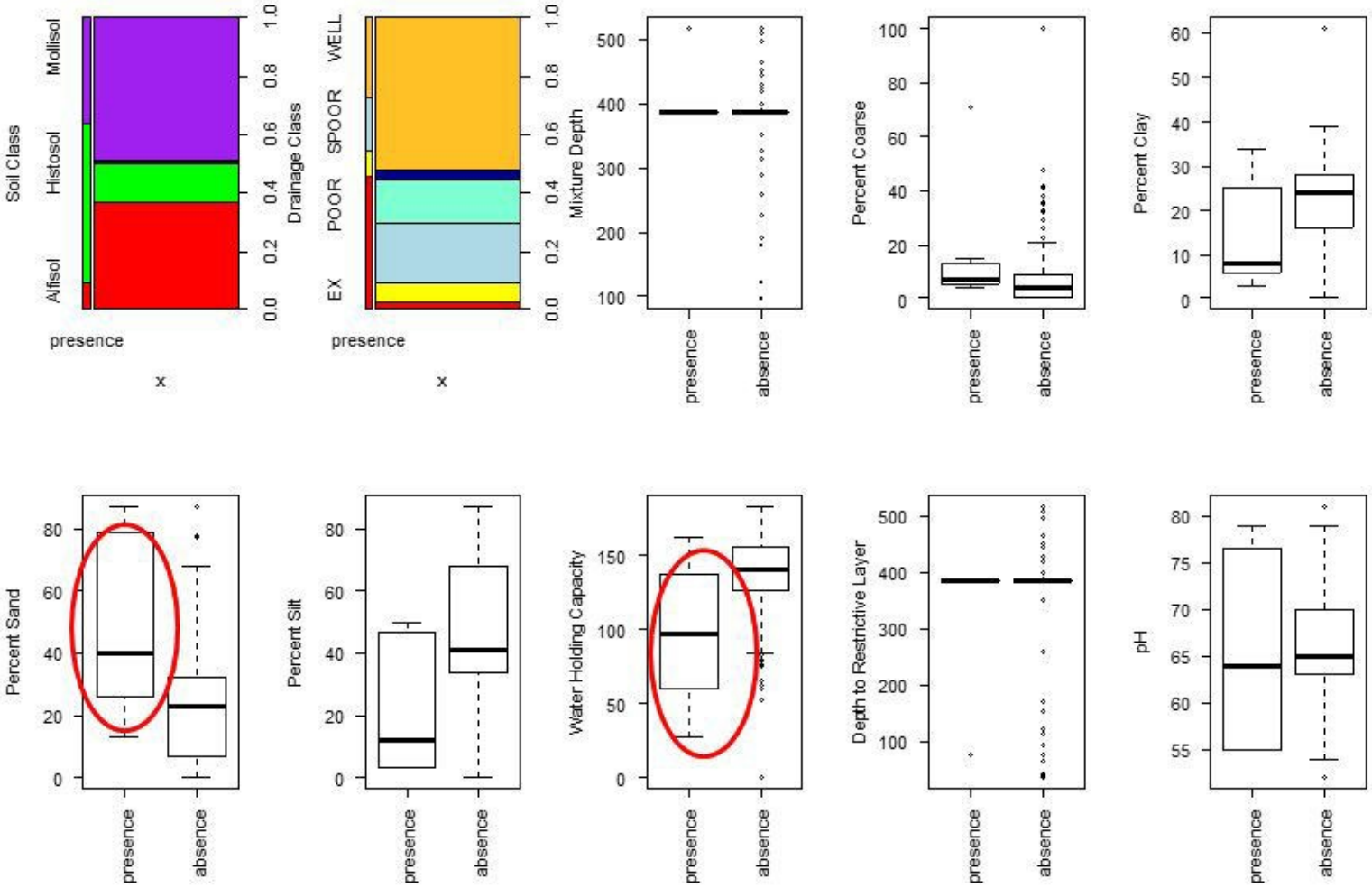

Percent Alfisols
Percent Entisols
Percent Histosols
Percent Inceptisols
Percent Mollisols
0
20
40
60
80
100
presence
absence

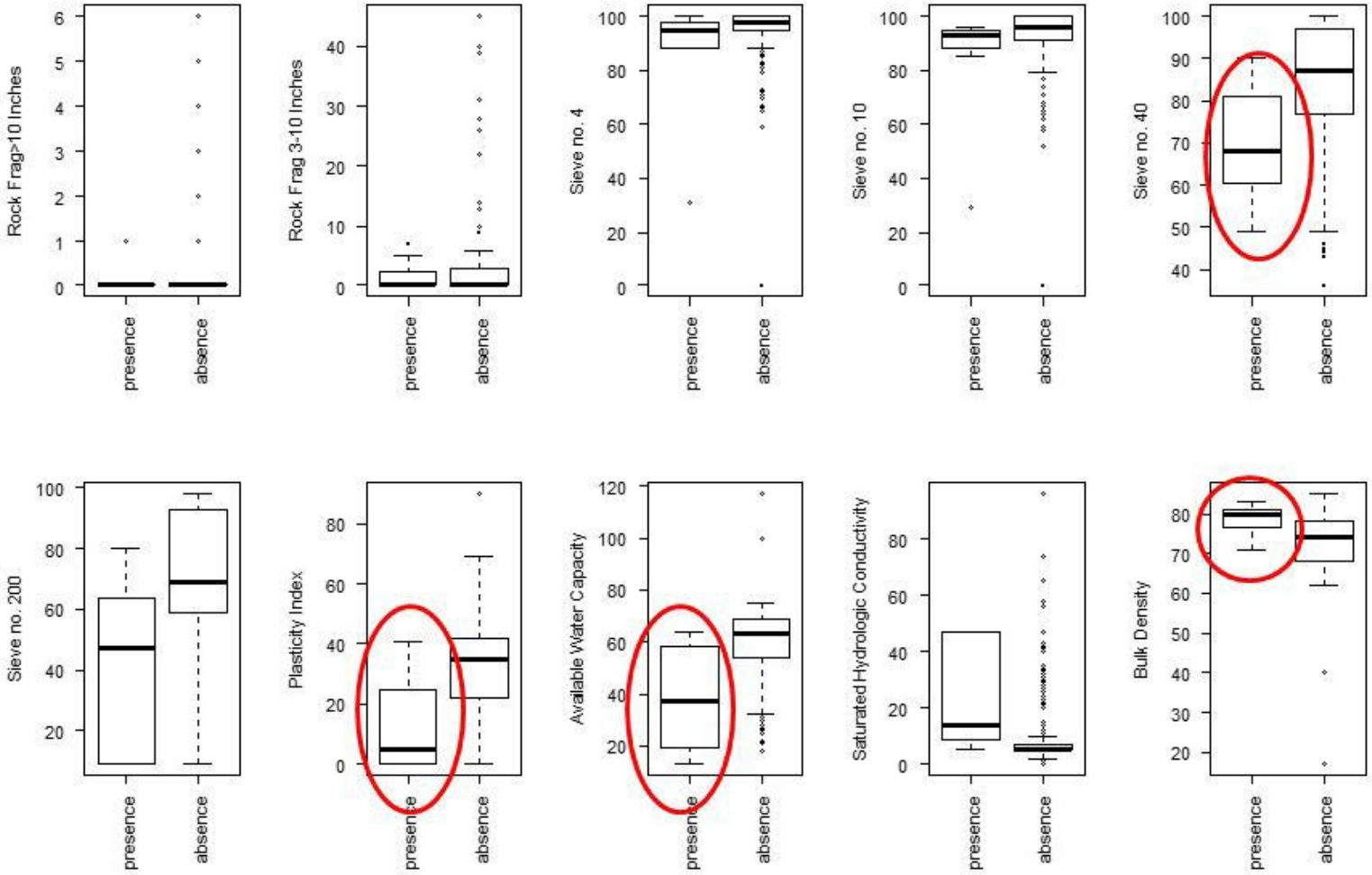
```
Call:
rpart(formula = as.factor(data$group) ~ ., data = data)
 n= 264

        CP nsplit rel error   xerror
1 0.08695652      0 1.0000000 1.000000
2 0.01000000      1 0.9130435 1.173913
     xstd
1 0.1992245
2 0.2140554

Variable importance
 SILTTOTAL_R        SILT       AWC_R
          36          21          11
DBTHIRDBAR_R   P_HISTOSOLS   SOIL_CLASS
          11          11          11

Node number 1: 264 observations,   complexity param=0.08695652
  predicted class=absence  expected loss=0.08712121  P(node) =1
    class counts:   23   241
   probabilities: 0.087 0.913
  left son=2 (10 obs) right son=3 (254 obs)
  Primary splits:
     SILTTOTAL_R    < 70.5 to the right, improve=5.468015, (0 missing)
     WFIFTEENBAR_R  < 22.5 to the right, improve=4.083923, (1 missing)
     DC_VPOOR       < 81.5 to the right, improve=3.833675, (0 missing)
     DRAIN_CLASS    splits as  RRR-RLR,  improve=3.447281, (0 missing)
     CLAYTOTAL_R    < 35.5 to the right, improve=3.447281, (0 missing)
  Surrogate splits:
     SILT        < 75.5 to the right, agree=0.985, adj=0.6, (0 split)
     SOIL_CLASS    splits as  RRLLR,    agree=0.973, adj=0.3, (0 split)
     P_HISTOSOLS  < 28.5 to the right, agree=0.973,  adj=0.3, (0 split)
     DBTHIRDBAR_R  < 53.5 to the left,  agree=0.973, adj=0.3, (0 split)
     AWC_R        < 87.5 to the right, agree=0.973,  adj=0.3, (0 split)
```

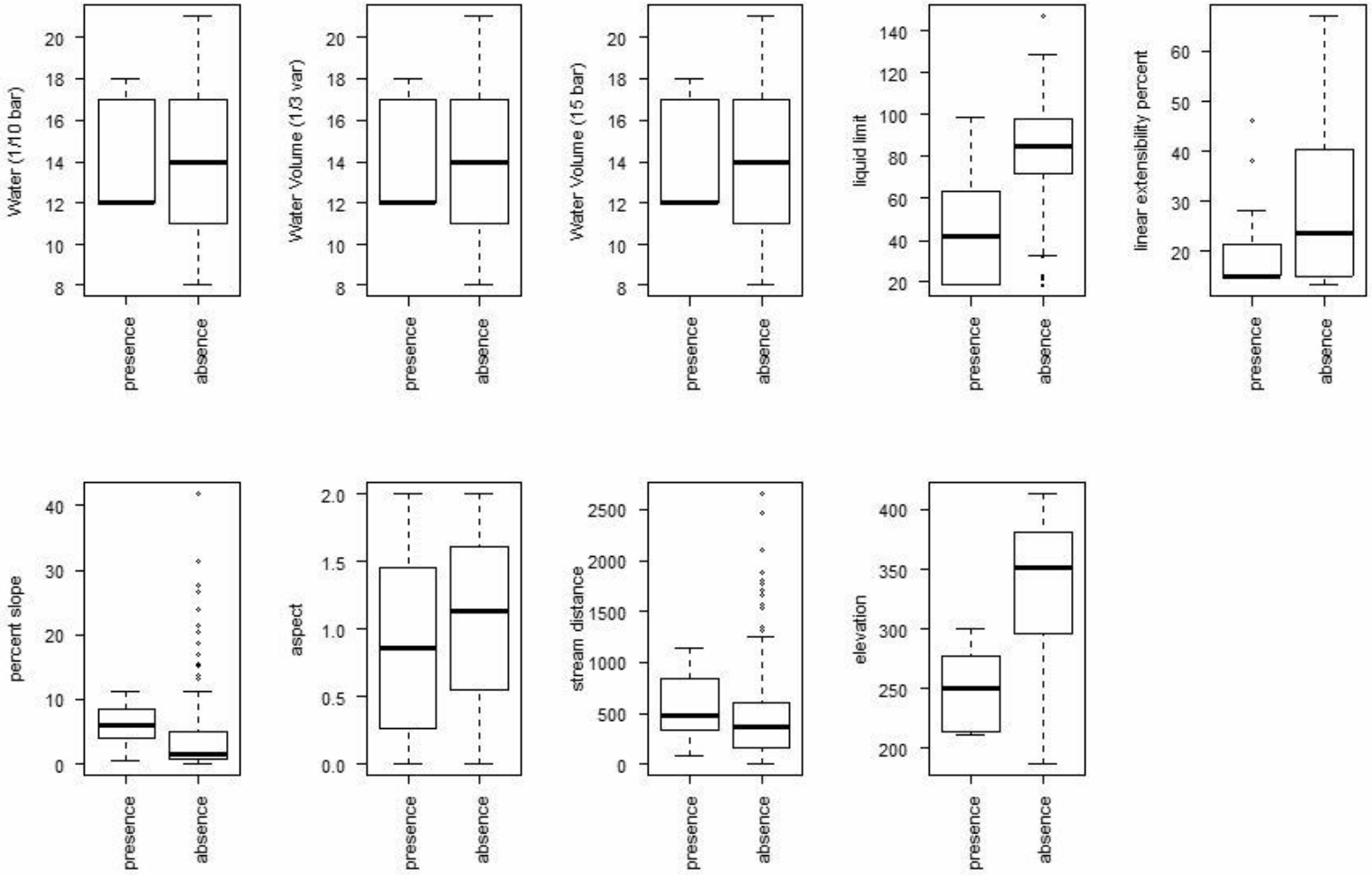


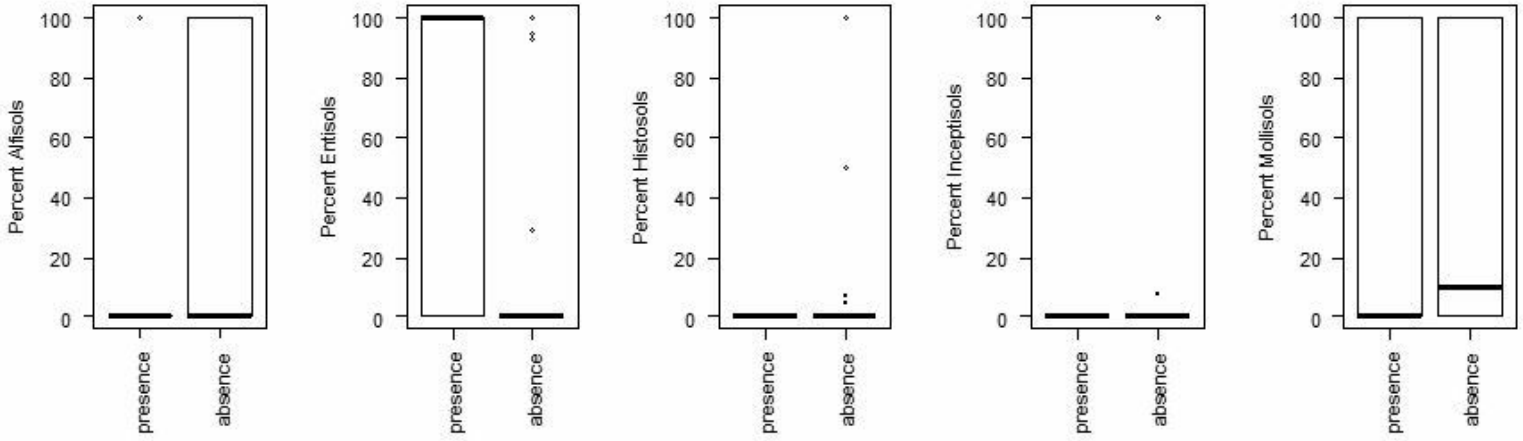
# NCI Oak Barrens

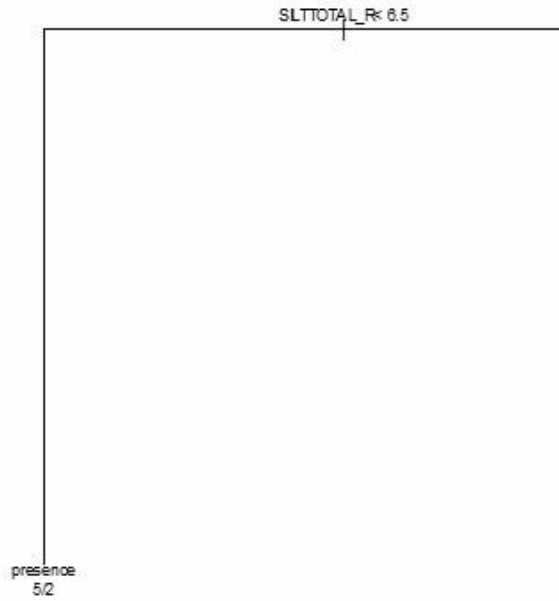
N presence = 11
N absence = 241




Rock Frag>10 Inches
Rock Frag 3-10 Inches
Sieve no. 4
Sieve no. 10
Sieve no. 40
Sieve no. 200
Plasticity Index
Available Water Capacity
Saturated Hydrologic Conductivity
Bulk Density
presence
absence


Water (1/10 bar)
Water Volume (1/3 var)
Water Volume (15 bar)
liquid limit
linear extensibility percent
percent slope
aspect
stream distance
elevation
presence
absence


Percent Alfisols
Percent Entisols
Percent Histosols
Percent Inceptisols
Percent Mollisols
0
20
40
60
80
100
presence
absence

```
Call:
rpart(formula = as.factor(data$group) ~ ., data = data)
 n= 252

       CP nsplit rel error  xerror     xstd
1 0.2727273      0 1.0000000 1.000000 0.2948573
2 0.0100000      1 0.7272727 1.363636 0.3414498

Variable importance
SILTTOTAL_R CLAYTOTAL_R  SANDFINE_R   SANDMED_R SANDTOTAL_R
       22         16         16         16         16
     SILT
       16

Node number 1: 252 observations,    complexity param=0.2727273
  predicted class=absence   expected loss=0.04365079  P(node) =1
    class counts:    11   241
   probabilities: 0.044 0.956
  left son=2 (7 obs) right son=3 (245 obs)
  Primary splits:
     SILTTOTAL_R < 6.5  to the left,  improve=6.476417, (0 missing)
     SANDFINE_R  < 35   to the right, improve=6.476417, (0 missing)
     PI_R        < 0.5  to the left,  improve=6.473796, (1 missing)
     SANDTOTAL_R < 84.5 to the right, improve=5.792606, (0 missing)
     SILT        < 6.5  to the left,  improve=5.584765, (0 missing)
  Surrogate splits:
     SILT        < 4.5  to the left,  agree=0.992, adj=0.714, (0 split)
     SANDTOTAL_R < 91   to the right, agree=0.992, adj=0.714, (0 split)
     SANDMED_R   < 29.5 to the right, agree=0.992, adj=0.714, (0 split)
     SANDFINE_R  < 35   to the right, agree=0.992, adj=0.714, (0 split)
     CLAYTOTAL_R < 3.5  to the left,  agree=0.992, adj=0.714, (0 split)

Node number 2: 7 observations
  predicted class=presence  expected loss=0.2857143  P(node) =0.02777778
    class counts:     5     2
   probabilities: 0.714 0.286

Node number 3: 245 observations
  predicted class=absence   expected loss=0.0244898  P(node) =0.9722222
    class counts:     6   239
   probabilities: 0.024 0.976
```

**North-Central Oak Barrens**