Surface Elevation Table- Marker Horizon (SET-MH) Data Analaysis in R

Adam Starke

Thursday, January 15, 2015

This document is intended to document the process used for database storage and data analysis of our SET-MH data. Our current workflow uses Access2010 to store the data and RStudio (running the R language) for data handling.

## Data Storage

The SET-MH method of marsh monitoring generates a fairly large set of data that needs to be carefully maintained and catalogued. Fortunately, there has been a substaintial effort put forth by the National Park Service's National Capital Region Network Inventory and Monitoring Program to create a user-friendly Access database. This database stores and organizes the data to allow for easy access to summarise and export data to other software if desired. Although the statistical analyses needed to calculate trends in marsh elevation are somewhat basic the restructuring of the data to allow for analysis can be somewhat complex. With this limtation of the data we decided to use the statistical programming language R. R is one of the most widely used programming languages and contains an enormous amount of tools and scripts (known as packages in R) that make these restructuring of data possible. The ability to create 'scripts' allow for quick updates to analysis after data entry. A number of other features also exist that make R a good choice, particulary the graphics capabilities and the support for creation of 'data-driven documents' which allows for reproducible research. This document is actually written with R code embedded within to allow for documenting the exact steps taken for data organization, restructuring and analysis while performing the actual analysis.

As mentioned the data are currently stored in an Access database located at:

[1] "C:\\Users\\astarke\\Documents\\Marsh-Wetlands\\SET\_Monitoring\_Database\\Database\_storage\\SET\_DB\_BE\_ver\_2.94\_TNC\_Master.mdb"

This file is referred to as the 'backend database' which holds all the raw SET-MH data.The 'front-end' of the database contains the user interface (UI) and is linked to the backend allowing multiple 'front ends' all pointing to one dataset. It's suggested that the front-end UI be used for quick checks on data

## Accessing Data through R

Using the power of R's scripting language capabilities the quickest, most direct, way to update SET-MH analyses is to simply navigate to the folder containing the RStudio project file (SET-MH\_Analysis.Rproj) currently located at: C:/Users/astarke/Documents/R\_Code/SET-MH\_Analysis. Opening this file in RStudio will load all the code associated with this project. Run the script 'MarshProject.R' by simply opening it and clicking the *'Source'* button in the top-right of the window. This will pull all the code from the file and run it. You will see activity in the "Console" window