

## **Introduction to Statistical Analysis and Machine Learning in R.**

Dan Schmutz, M.S., Chief Environmental Scientist, Greenman-Pedersen, Inc.

### **Description:**

This workshop will provide a gentle introduction to the arcane world of R, covering topics useful for those with no prior exposure to the software as well as those interested in learning more advanced applications. This free, open source system for statistical computation and graphics has become the de facto standard for ecologists with more than half of published scientific papers now using the software. Given its flexibility and widespread adoption by the scientific community, R training has broad applicability for environmental scientists and lake managers working in all contexts requiring data management, statistical analysis, and graphical presentation of data. R allows precise documentation of steps taken in data processing in order to provide reproducible results in alignment with the Open Science movement. Graphical user interfaces and integrated development environments now make R much more user friendly for beginners (i.e., more point-and-click and less command line). Users have contributed over 15,000 packages to R, greatly extending its capabilities. Have your laptop handy, if possible, as the workshop will include a hands-on portion.

Note: This workshop may be held over one full day or spaced out into two half days.

### **Topics:**

Overview of R Capabilities

How to Not Panic: Getting Help and Typical Problems

Installing R and R Studio

Project Organization

Getting to Know R and R Studio

R the language: Data Structures, Scripting, Control Statements, Functions

Must Have Packages: Get 'Em While They're Free

Fun with Data: Importing, Cleaning, Filtering, Reshaping, Summarizing, Exporting

Basic Graphing and Statistical Analysis

Interactive Graphing for Rapid Exploratory Analyses

Introduction to Machine Learning

### **About the Instructor:**

Dan Schmutz is a Vice President and Chief Environmental Scientist for Greenman-Pedersen, Inc. (GPI), a 1,200-person multidisciplinary consulting firm. He holds a Master's degree in Zoology from the University of South Florida. He has over 20 years of professional experience focused on the development and application of ecological and hydrological field assessments, GIS analyses, and appropriate statistical techniques for addressing questions of interest to water supply managers.