Processing Instructions for the 2014 Trend Analysis

Jason C. Fisher

September 12, 2016

If R (version $\geq$ 3.1) is not already installed on your computer, download and install the latest binary distribution from the Comprehensive R Archive Network (CRAN). Open an **R** session and install the required packages with the following commands:

```
repos <- c("http://cran.us.r-project.org", "https://jfisher-usgs.github.io/R")
install.packages("Trends", repos = repos, dependencies = TRUE, type = "both")
```

Support for merging Portable Document Format (PDF) files into a new file requires PDFtk Server, a cross-platform command-line tool for working with PDFs; download and install this software.

Load the Trends package into the current **R** session:

```
library(Trends)
```

Set the full path name to the directory containing input files; list all files in this directory:

```
list.files(path.in <- system.file("extdata", "SIR2014", package = "Trends"))
#  [1] "Config_Field.tsv"     "Config_NWQL.tsv"      "Config_RAD.tsv"
#  [4] "Config_VOC.tsv"       "Detection_Limits.tsv" "Observations.tsv"
#  [7] "Parameters.tsv"       "Site_Locations.dbf"   "Site_Locations.prj"
# [10] "Site_Locations.shp"   "Site_Locations.shx"   "Tables.xlsx"
# [13] "Water_Levels.tsv"
```

Set the full path name to the output directory; if it doesn't already exist, create this directory:

```
path.out <- file.path(getwd(), paste0("Trends_", format(Sys.time(), "%Y%m%d%H%M%S")))
dir.create(path = path.out, showWarnings = FALSE)
```

Set the graphics type for output figures to a PDF file format:

```
graphics.type <- "pdf"
```

Specify global arguments for reading table formatted data from a text file:

```
read.args <- list(header = TRUE, sep = "\t", colClasses = "character", na.strings = "",
                  fill = TRUE, strip.white = TRUE, comment.char = "", flush = TRUE,
                  stringsAsFactors = FALSE)
```

Read observational data from a text file:

```
file <- file.path(path.in, "Observations.tsv")
observations <- do.call(read.table, c(list(file), read.args))
```

Read parameter descriptions from a text file:

```
file <- file.path(path.in, "Parameters.tsv")
parameters <- do.call(read.table, c(list(file), read.args))
```

Read detection limits from a text file:

```
file <- file.path(path.in, "Detection_Limits.tsv")
detection.limits <- do.call(read.table, c(list(file), read.args))
```

Process observational data:

```
processed.obs <- ProcessObs(observations, parameters, detection.limits,
                            date.fmt = "%m/%d/%Y")
```

Read geo-referenced site locations from a point shapefile:

```
site.locations <- rgdal::readOGR(path.in, layer = "Site_Locations", verbose = FALSE)
```

Read water levels from a text file:

```
file <- file.path(path.in, "Water_Levels.tsv")
water.levels <- do.call(read.table, c(list(file), read.args))
```

Process water-level data:

```
processed.wl <- ProcessWL(water.levels, date.fmt = "%Y-%m-%d %H:%M")
```

Run trend analysis for National Water Quality Laboratory (NWQL) parameters from 1989 through 2012:

```
file <- file.path(path.in, "Config_NWQL.tsv")
config <- do.call(read.table, c(list(file), read.args))
processed.config <- ProcessConfig(config, processed.obs)
RunAnalysis(processed.obs, processed.config, path = path.out,
            id = "Stats_NWQL_1989-2012", sdate = "1989-01-01", edate = "2012-12-31",
            site.locations = site.locations, graphics.type = graphics.type)
```

Run trend analysis for field parameters (pH, Specific Conductance, and Temperature) from 1989 through 2012:

```
file <- file.path(path.in, "Config_Field.tsv")
config <- do.call(read.table, c(list(file), read.args))
processed.config <- ProcessConfig(config, processed.obs)
RunAnalysis(processed.obs, processed.config, path = path.out,
            id = "Stats_Field_1989-2012", sdate = "1989-01-01", edate = "2012-12-31",
            site.locations = site.locations, graphics.type = graphics.type)
```

Run trend analysis for radiation-related (RAD) parameters from 1981 through 2012:

```
file <- file.path(path.in, "Config_RAD.tsv")
config <- do.call(read.table, c(list(file), read.args))
processed.config <- ProcessConfig(config, processed.obs)
RunAnalysis(processed.obs, processed.config, path = path.out,
            id = "Stats_RAD_1981-2012", sdate = "1981-01-01", edate = "2012-12-31",
            site.locations = site.locations, graphics.type = graphics.type)
```

Run trend analysis for volatile organic compound (VOC) parameters from 1987 through 2012:

```
file <- file.path(path.in, "Config_VOC.tsv")
config <- do.call(read.table, c(list(file), read.args))
processed.config <- ProcessConfig(config, processed.obs)
RunAnalysis(processed.obs, processed.config, path = path.out,
            id = "Stats_VOC_1987-2012", sdate = "1987-01-01", edate = "2012-12-31",
            site.locations = site.locations, graphics.type = graphics.type)
```

Run trend analysis for Carbon Tetrachloride (P32102), a VOC parameter, at the RWMC Production well from 2005 through 2012:

```
rec <- with(processed.config, Parameter_id == "P32102" & Site_id == "433002113021701")
RunAnalysis(processed.obs, processed.config[rec, ], path = path.out,
            id = "Stats_RWMC_CCl4_2005-2012",sdate = "2005-01-01", edate = "2012-12-31",
            graphics.type = graphics.type)
```

Run trend analysis for Chloride (CL), a NWQL parameter, at the USGS 114, PW 8, USGS 87, USGS 35, and USGS 39 wells from 1989 through 2012:

```
file <- file.path(path.in, "Config_NWQL.tsv")
config <- do.call(read.table, c(list(file), read.args))
processed.config <- ProcessConfig(config, processed.obs)
site.ids <- c("433318112555001", "433456112572001", "433013113024201",
              "433339112565801", "433343112570001")
rec <- with(processed.config, Parameter_id == "CL" & Site_id %in% site.ids)
RunAnalysis(processed.obs, processed.config[rec, ], path = path.out,
            id = "Test",sdate = "1989-01-01", edate = "2012-12-31",
            graphics.type = graphics.type)
```

Account for residual water-level variability in the trend model.

```
RunAnalysis(processed.obs, processed.config[rec, ], path = path.out,
            id = "Test_resWLs",sdate = "1989-01-01", edate = "2012-12-31",
            graphics.type = graphics.type, explanatory.var = processed.wl,
            is.residual = TRUE)
```

Account for seasonal variability in the trend model.

```
RunAnalysis(processed.obs, processed.config[rec, ], path = path.out,
            id = "Test_Seas",sdate = "1989-01-01", edate = "2012-12-31",
            graphics.type = graphics.type, is.seasonality = TRUE)
```

Account for residual water-level variability and seasonal variability in the trend model.

```
RunAnalysis(processed.obs, processed.config[rec, ], path = path.out,
            id = "Test_resWLs_Seas",sdate = "1989-01-01", edate = "2012-12-31",
            graphics.type = graphics.type, is.seasonality = TRUE,
            explanatory.var = processed.wl, is.residual = TRUE)
```

To reprocess the 2014 water quality data, evaluate **R** code extracted from this vignette using the following command:

```
source(system.file("doc", "Trends-process.R", package = "Trends"), echo = TRUE)
list.files(path.out, full.names = TRUE, recursive = TRUE)  # path names of output files
```

Total processing time for this vignette was 2 minutes. Version information about **R** and loaded packages is as follows:

- R version 3.3.1 (2016-06-21), `x86_64-w64-mingw32`

- Base packages: base, datasets, graphics, grDevices, methods, stats, utils

- Other packages: Trends 1.0.1

- Loaded via a namespace (and not attached): devtools 1.12.0, digest 0.6.10, evaluate 0.9, formatR 1.4, grid 3.3.1, highr 0.6, knitr 1.14, lattice 0.20-34, magrittr 1.5, Matrix 1.2-7.1, memoise 1.0.0, rgdal 1.1-10, sp 1.2-3, splines 3.3.1, stringi 1.1.1, stringr 1.1.0, survival 2.39-5, tools 3.3.1, withr 1.0.2