

# Project: Statistical Design, Analysis and Guidance on the Pajaro and Lower Peninsula Watershed Assessments

TASK 3: GRTS Survey Designs and Sample Draws Memorandum  
Pajaro and Lower Peninsula Watersheds

Report prepared for the Santa Clara Valley Water District  
Ecological Data Collection and Analysis Project (Priority D.5)

Agreement # A3765F

SFEI-ASC Project # 4092 Task 003

Submitted by: San Francisco Estuary Institute

Authors: Sarah Lowe

THIS REPORT SHOULD BE CITED AS:

Lowe, S. (2015). Project: Statistical Design, Analysis and Guidance on the Pajaro and Lower Peninsula Watershed Assessments. TASK 3: GRTS Survey Designs and Sample Draws Memorandum – Pajaro and Lower Peninsula Watersheds Report prepared for the Santa Clara Valley Water District Ecological Data Collection and Analysis Project (Priority D.5). San Francisco Estuary Institute & Aquatic Science Center, Richmond, CA. Contribution # 763.

## Table of Contents

Introduction .....	1
Pajaro River Watershed Stream Survey Design and Sample Draw.....	1
1. Determine the Number of Sites to Target .....	2
2. Develop the Final Survey Design and Sample Draw .....	4
Lower Peninsula Watershed Stream Survey Design and Sample Draw.....	7
1. Determine the Number of Sites to Target .....	7
2. Develop the Final Survey Design and Sample Draw .....	8
References .....	10
APPENDIX A: Annotated R Code Example.....	11
Overview .....	11
Pajaro River Watershed GRTS Sample Draw R Code .....	11



## Introduction

The watershed wide, probability based, stream condition survey designs and sample draws for the Santa Clara County Pajaro and Lower Peninsula watershed assessments (2015 and 2016 respectively) employ the Generalized Random Tessellation Stratified (GRTS) survey design tools, developed by the U. S. Environmental Protection Agency (EPA) for the National Environmental Monitoring and Assessment Program (EMAP; Messer et al. 1991; Stevens and Olsen 2003; Stevens and Olsen 2004).

In a GRTS probability sample for streams, CRAM assessment areas are randomly selected from a GIS based stream network (or linear sample frame - described in *Task 2: Basis of Assessment Memorandum*), while accounting for the proportion of the stream resource that each assessment area represents. Survey results can be analyzed to estimate the proportion of the total resource in the study area that is likely to have any particular condition as assessed using CRAM. The efficiency of the GRTS survey design can be increased (but, still maintain its unbiased nature) by stratifying the sample draw. Increasing the relative proportion of samples in the District's Primary Areas of Interest (PAIs) and/or higher stream orders can deliberately increase the number of samples in areas of most interest to the District, and increasing the confidence levels around the means in those areas of specific interest, while preserving the ability to evaluate the conditions in the watershed as a whole. The GRTS sample draws for both the Pajaro and Lower Peninsula stream surveys were stratified by PAI and stream order to increase the confidence levels around the means in areas of most interest to the District.

This memorandum outlines the survey design decisions and programming steps for the GRTS sample draws for the Pajaro River and Lower Peninsula watershed wide stream condition surveys based on CRAM assessed in 2015 and 2016 respectively.

The GRTS survey design and analysis package "spsurvey"<sup>1</sup> was created in R (a statistical computing and graphics programming language) and is a design and analysis tool for developing environmental survey designs and sample draws for linear and area resources (e.g. streams, wetlands, lakes etc.).

## Pajaro River Watershed Stream Survey Design and Sample Draw

There is no 'right' way to estimate the number of assessment sites to target in a new ambient watershed survey, and in general the limiting factors are often the sampling budget and time. The more variable the resource, the more samples will be needed to characterize it with a reasonable level of confidence. A power analyses of previous assessments in similar watersheds and best professional judgment were used to inform the study design for the District's watershed wide stream assessments based on CRAM.

SFEI performed a power analyses based on the results of the Coyote Creek and Guadalupe River watershed assessments conducted by the District in 2010 and 2012, sought advice from the EPA GRTS developers, and experimented with sample stratification between PAIs and stream order. Results were presented and discussed with District staff and together the final survey design and sample draw was determined. Below is a summary of those efforts.

---

<sup>1</sup> Thomas M. Kincaid and Anthony R. Olsen (2011). spsurvey: Spatial Survey Design and Analysis. R package version 2.2.

## 1. Determine the Number of Sites to Target

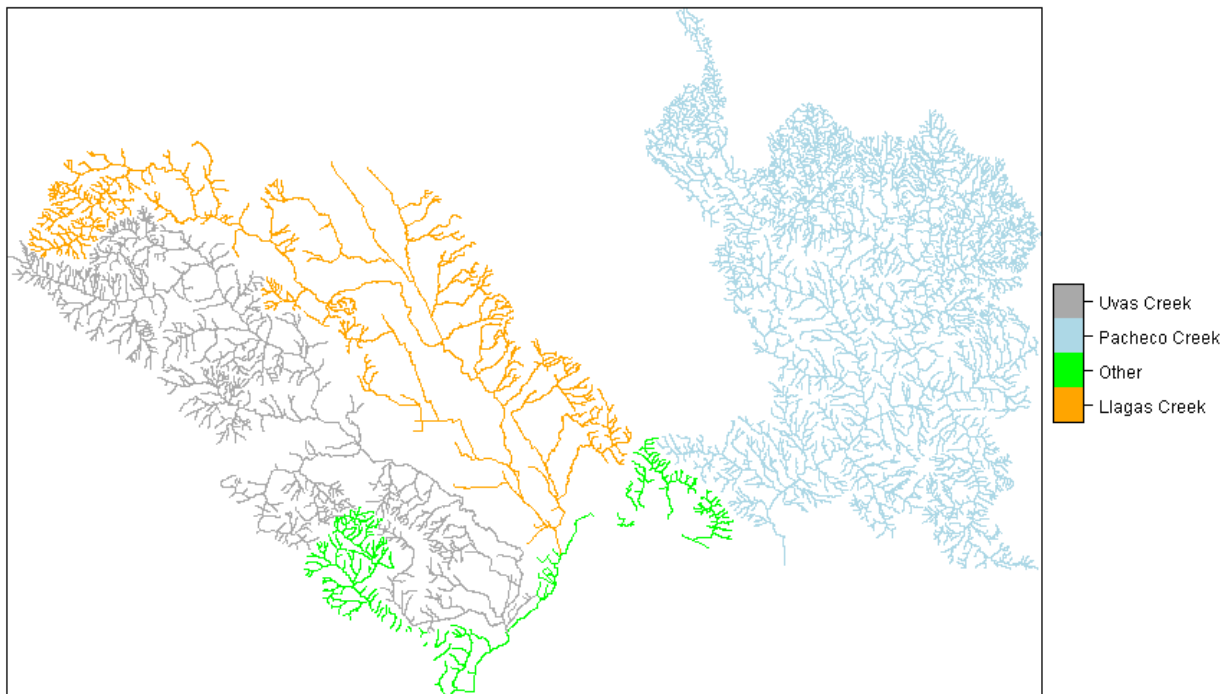
### A. Conduct a power analysis of similar previous surveys

To get a sense of the number of sites to target per PAI, a power analysis was performed using CRAM results from the District's previous CRAM stream surveys in Coyote Creek and Guadalupe River watersheds.

The mean standard deviations of the Overall CRAM score cumulative distribution function plots (CDFs) from those assessments were evaluated with the goal of being able to tell a difference in scores by 10 CRAM points. At a 90% confidence level the power analysis suggested that at least 24 sites were needed in each PAI to detect a 10 point difference in scores. While the Pajaro River watershed is different than the previous surveys, the power analysis provided some empirical context for estimating the number of samples to allocate to each PAI - if we want to discern a 10 point difference in scores. Additionally, statisticians generally say that one wants to take as many samples as can be afforded because it is usually funding that limits the sample size in most field studies.

### B. Become familiar with the distribution of the stream network in the study area

Based on the District's *Creeks* stream layer, the portion of the Pajaro River watershed that falls within Santa Clara County consists of about 1,500 miles of streams that include three PAIs and small portions of other tributaries that flow into the main-stem of the Pajaro River. The District's stream assessment focused on the three PAIs and separately included the 'other' portions of stream reaches in other subwatersheds in order to assess the whole study area within the County (Figure 1).



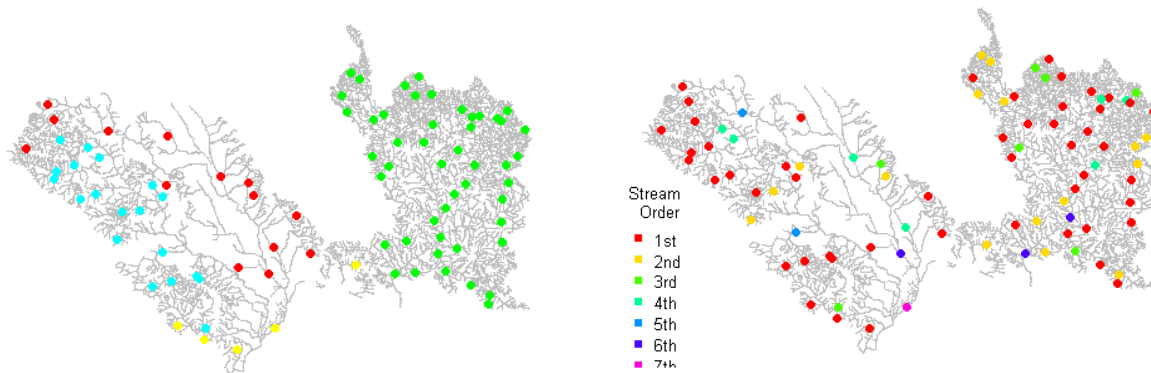
**Figure 1.** Map of the Pajaro Watershed stream condition assessment sample frame showing the three Primary Areas of Interest (PAIs) and the 'Other tributaries' - including portions of Pescadero Creek, the Pajaro River mainstem and other small watershed tributaries.

The Pacheco Creek subwatershed includes about 55% of the total stream lengths in the entire Pajaro watershed study area, while Uvas and Llagas Creek subwatersheds each contribute about 20% of the total stream length and the 'other' small watershed stream reaches, including the Pajaro main-stem, makeup about 5%. First and second order streams within the watershed make up about 75% of the total stream length, while third and fourth order streams make up almost 20%. Fifth and sixth order streams make up about 5%, and the Pajaro River main-stem (seventh order stream reaches) is about 1% of the stream length within the study area.

C. Develop an unstratified, trial sample draw

In an unstratified GRTS sample draw, sites are allocated to a stream resource proportional to the stream lengths across the whole study area (or sample frame). That means that sites are allocated proportional to the stream lengths in the PAIs and by Strahler stream order. Therefore, it is expected that low number Strahler stream order reaches, high up in the watershed, which comprise most of the stream length will get the majority of the sampling sites as will the largest subwatershed (or PAI). By running an unstratified sample draw, one gets a sense of the sample distribution across the entire resource.

A trial, unstratified, sample draw of 86 sites for the Pajaro watershed (Figure 2) indicated that almost half the sites were allocated to the Pacheco Creek watershed (48 sites) and Uvas and Llagas Creek watersheds had less than 20 sites each (19 and 14 respectively; Figure 2.A). Additionally, almost 80% of the sites were allocated to first and second order streams in the upper watershed (Figure 2.B) in areas where the District has less authority to effectively manage and make improvements.



2.A. Sites colored by Primary Area of Interest (PAI).

2.B. Sites colored by Strahler stream order.

**Figure 2.** Unstratified GRTS sample draw of 86 sites for the Pajaro watershed. Figure 2.A shows that almost half the sites were allocated to the Pacheco Creek watershed (48 sites shown in green), 19 sites (light blue) in and Uvas, and 14 sites (red) in Llagas Creek watersheds. Figure 2.B shows that the majority of the sites were allocated to first and second order streams in the upper watershed (red and yellow dots, which comprise almost 80% of the sites).

After review of the unstratified sample draw, the design team determined that 86 targeted sites for the whole watershed was an adequate sample size based on budget, time, and spatial coverage across the

watershed. However, the team decided that an unstratified sample draw was unacceptable as it allocated less than 20 sites in Uvas and Llagas PAIs respectively and too many sites in the upper watershed. The team decided to stratify the GRTS sample draw on both PAIs and Strahler stream order and recommended the following stratifying measures:

1. Total number of target CRAM assessment areas recommended = 86
2. Assign specific numbers of target sites to each PAI:
  - 24 sites each to Uvas and Llagas PAIs,
  - 8 sites in the “other” small watershed - including 4 sites along the Pajaro main-stem, and
  - 30 sites to the Pacheco Creek watershed.
3. Further stratify on stream order to force more sites into the higher order streams (or the lower elevations of the watershed). The team felt it was more important to allocate sites in the lower elevations of the watershed for three reasons:
  - The District’s main interests, and ability to impact change, lie within the downstream portions of the watershed.
  - If left alone, the probability sample will proportionally allocate the majority of sites in low order streams and under-sample the higher order streams where the District has more ability to make management changes.
  - Based on previous experience in the Coyote and Guadalupe watersheds it may be very difficult to gain permission from landowners to access CRAM sites in the upper elevations of the watersheds.

## 2. Develop the Final Survey Design and Sample Draw

Based on discussion with the developers of EPA’s GRTS design tools (2/19/2015 Tom Kincaid), the best way to direct the number of sites in each PAI and force more sites into the higher order streams (in the lower watershed) is to treat each PAI as a stratum (where the exact number of sites can be defined), and *request* the number of sites per stream order using the ‘spsurvey’ program’s multi-density category (mdcaty). This design treats each PAI as its own, stand-alone, sample draw yet they can be combined to characterize the ecological condition of streams across the Pajaro watershed as a whole (this is done by adjusting the sample weights - at the data analysis phase- to be proportional to the stream lengths of the entire watershed).

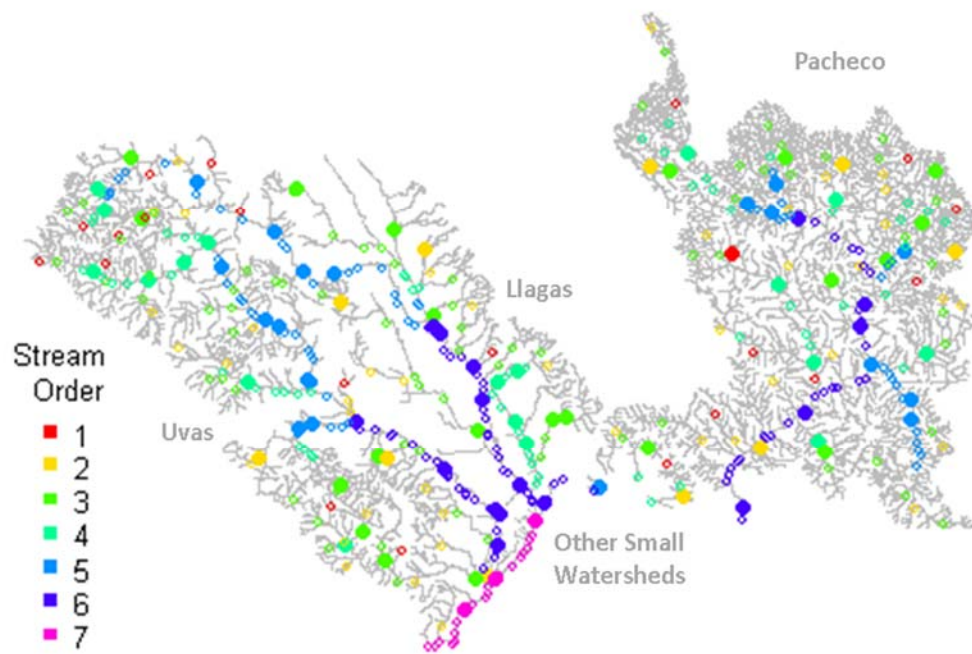
In the final Pajaro River watershed stream survey sample draw, assessment areas were distributed across the District’s updated *Creeks* dataset for all Strahler stream orders (1-7) in four strata: three PAIs (the entire Uvas and Llagas watersheds, and the Pacheco Creek watershed within Santa Clara County), and ‘Other’ small tributaries within Santa Clara County (including parts of the Pajaro River that falls within the County, and a few additional small portions of the other tributaries).

A total of 88 CRAM assessment areas were targeted with an oversample of three times the targeted number of sites. Table 1 summarizes the target and oversample sites by stratum (the three PAIs and the ‘Other’ small tributaries), Figure 3 shows a map of the distribution of sites by stream order across the study area, and Table 2 lists the design parameters for the GRTS sample draw.



**Table 1.** Number of targeted and oversample sites in each PAI in the Pajaro River Watershed CRAM stream condition assessment survey 2015.

Primary Area of Interest (PAI)	Target Sites	Oversample Sites	Total Sites
Uvas Creek watershed	24	72	96
Llagas Creek watershed	24	72	96
Pacheco Creek watershed	30	90	120
Other tributaries (not a PAI)	10	30	40
<b>Total</b>	<b>88</b>	<b>264</b>	<b>352</b>



**Figure 3.** Final distribution of target (solid dots) and oversample (tiny open circles) sites for the Pajaro River Watershed CRAM stream condition assessment survey 2015. Sites are color coded by Strahler stream order.

**Table 2.** List of the Pajaro watershed condition assessment design parameters for the GRTS sample draw.

<b>Study Area</b>	Pajaro watershed within Santa Clara County, CA
<b>Sample Frame</b>	'Santa Clara County <i>Creeks</i> ' GIS data layer updated with flow direction and Strahler stream orders (1 through 7) in the Pajaro watershed.
<b>Sample Strata</b>	Three Primary Areas of Interest ( PAIs): Uvas , Llagas and Pacheco Creek watersheds within Santa Clara County, and  'Other' small tributaries within Santa Clara County
<b>Survey Design</b>	Generalized Random Tessellation Stratified Design (GRTS)
<b>Resource Type</b>	Streams (a linear resource)
<b>Target Sample Size</b>	88 AAs
<b>Over sample</b>	3x Target Sample Size
<b>Panels</b>	None

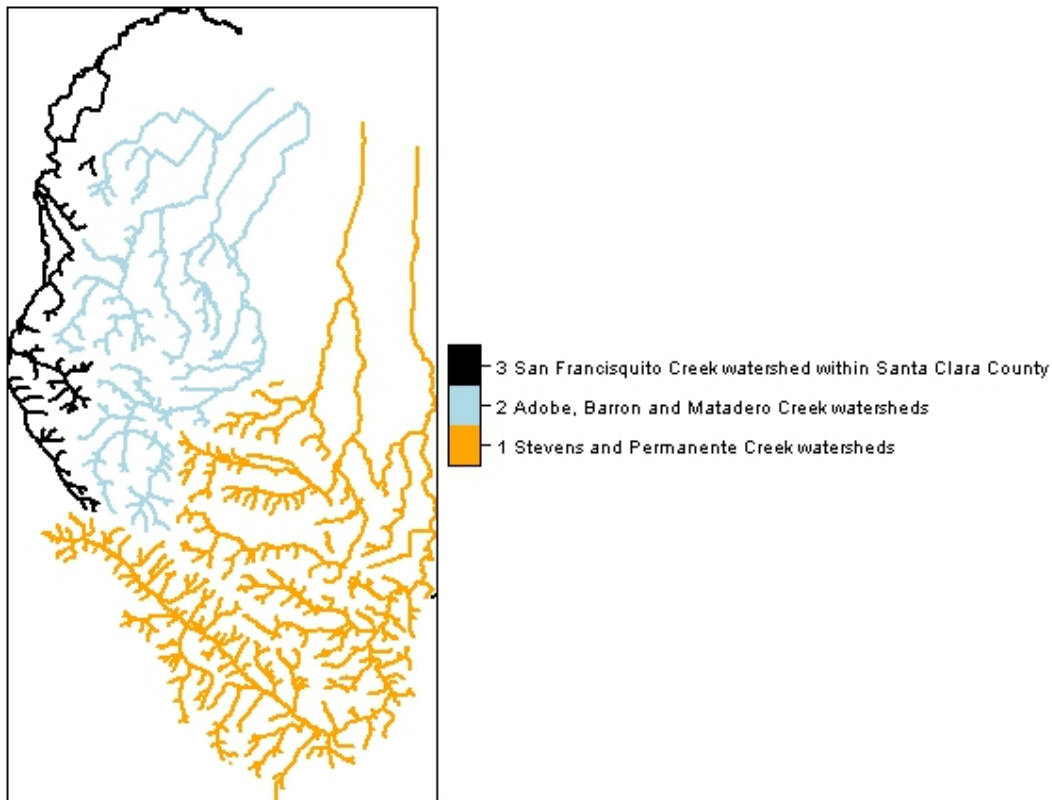
Notice that the final number of target sites was increased to 88 from the initial 86 sites recommended in the design request described in section 1.C (above). The update was made while developing the final sample draw in R. When 86 sites were targeted (with 8 sites requested in the 'Other tributaries' stratum) only one or two target sites were allocated to the Pajaro River mainstem (stream order 7 in the 'Other' category). The District project leads requested about 4 sites in the mainstem. By adding two extra samples into the 'Other' stratum, and increasing the total sample draw to 88, the final sample draw targeted 3 sites in the Pajaro River mainstem. Using the multi-density category (*mdcaty*) to *request* a specific number of sites by stream order, doesn't necessarily result in exactly the number of sites requested - likely because the statistical draw must spatially balance the request across the stratum.

# Lower Peninsula Watershed Stream Survey Design and Sample Draw

## 1. Determine the Number of Sites to Target

### A. Become familiar with the distribution of the stream network in the study area

The three PAIs in the Lower Peninsula watershed study area included the freshwater reaches of: 1) Stevens and Permanente Creek watersheds, 2) Adobe, Barron and Matadero Creek watersheds, and 3) San Francisquito Creek watershed within Santa Clara County (Figure 4).



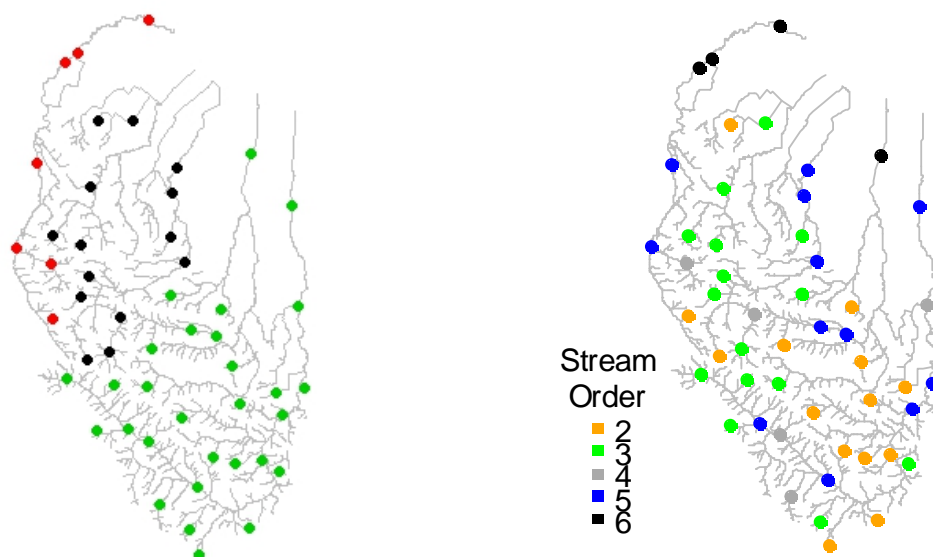
**Figure 4.** Map of the Lower Peninsula Watershed stream condition assessment sample frame showing the three Primary Areas of Interest (PAIs). Includes Strahler stream orders 2-6 (BAARI v.1).

The Lower Peninsula watershed study area encompasses about 80 square miles and includes about 500 miles of streams based on the BAARIV.1 streams layer (which includes all Strahler stream orders 1 through 6). However, the actual stream condition survey's sample frame does not include the BAARI 1<sup>st</sup> order streams and that stream network is about 250 miles in length. The Stevens and Permanente Creek streams includes over 55% of the total stream length (stream orders 2-6), while Adobe, Barron and Matadero Creeks watershed streams contribute almost 30%. The San Francisquito Creek watershed within Santa Clara County contributes the remaining 15% of the stream lengths.

Second order streams make up about 40% of the total stream length in the sample frame, and third and fourth order streams make up almost 40%. Fifth and sixth order streams make up a little over 20% of the total sample frame stream lengths.

## B. Develop an unstratified, trial sample draw

An unstratified GRTS sample draw, was developed for the Lower Peninsula study area to evaluate the unstratified allocation of assessment sites proportional to the stream resources across the whole sample frame. The trial sample draw of 50 sites indicated that almost half the sites (Figure 5.A) were allocated to the Stevens & Permanente Creek watershed (29 sites shown in green), 14 sites (black) in Adobe, Barron and Matadero Creeks watershed, and 7 sites (red) in San Francisquito Creek watershed that falls within Santa Clara County. Figure 5.B shows that more than half of the sites were allocated to second and third order streams at the top of the watersheds (orange and green dots, which comprise about 60% of the sites).



5.A. Sites colored by Primary Area of Interest (PAI).

5.B. Sites colored by Strahler stream order.

**Figure 5.** Unstratified GRTS sample draw of 50 sites for the Lower Peninsula watershed study area. Figure 5.A shows that almost half the sites were allocated to the Stevens & Permanente Creek watershed (29 sites shown in green), 14 sites (black) in Adobe, Barron and Matadero Creeks watershed, and 7 sites (red) in San Francisquito Creek watershed that falls within Santa Clara County. Figure 5.B shows that more than half of the sites were allocated to second and third order streams in the upper watershed (orange and green dots, which comprise about 60% of the sites).

## 2. Develop the Final Survey Design and Sample Draw

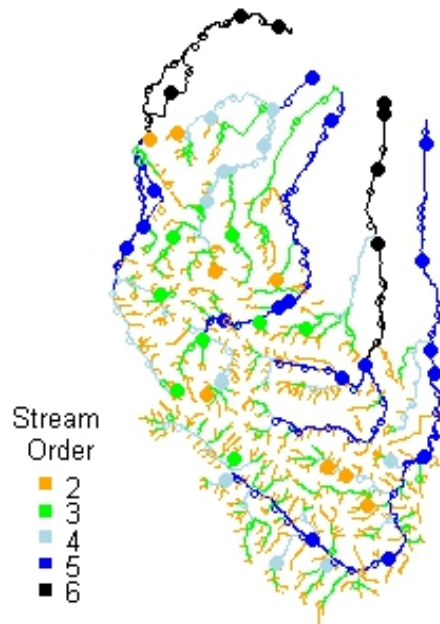
In the final Lower Peninsula watershed stream condition survey sample draw, assessment areas were distributed across the BAARI v.1 streams dataset for Strahler stream orders 2-6 (BAARI stream order 1 was dropped) in three strata: 1) Stevens and Permanente Creek watersheds, 2) Adobe, Barron and Matadero Creek watersheds, and 3) San Francisquito Creek watershed within Santa Clara County.

A total of 50 CRAM assessment areas were targeted with an oversample of three times the targeted number of sites. Table 3 summarizes the number of target and oversample sites by stratum (the three

PAIs). Figure 6 shows a map of the sites by stream order across the study area, and Table 4 lists the design parameters for the GRTS sample draw.

**Table 3.** Number of targeted and oversample sites in each PAI in the Lower Peninsulas Watershed CRAM stream condition assessment survey 2016.

Primary Area of Interest (PAI)	Target Sites	Oversample Sites	Total Sites
Stevens and Permanente Creeks	22	54	72
Adobe, Barron and Matadero Creeks	18	30	40
San Francisquito Creek in SC County	10	66	88
<b>Total</b>	<b>50</b>	<b>150</b>	<b>200</b>



**Figure 6.** Final distribution of target (solid dots) and oversample (tiny open circles) sites for the Lower Peninsula Watershed CRAM stream condition assessment survey 2016. Sites are color coded by Strahler stream order.

**Table 4.** Summary of Lower Peninsula watershed Assessment GRTS parameters for the sample draw.

<b>Study Area</b>	Lower Peninsula watershed within Santa Clara County, CA
<b>Sample Frame</b>	Bay Area Aquatic Resources Inventory (BAARI V.1) fluvial streams GIS layer included Strahler stream orders (2-6) in the Lower Peninsula watershed
<b>Sample Strata</b>	Three Primary Areas of Interest (PAIs): 1 Stevens and Permanente Creek watersheds 2 Adobe, Barron and Matadero Creek watersheds 3 San Francisquito Creek watershed within Santa Clara County
<b>Survey Design</b>	Generalized Random Tessellation Stratified Design (GRTS)
<b>Resource Type</b>	Streams (a linear resource)
<b>Target Sample Size</b>	50 AAs
<b>Over sample</b>	3x Target Sample Size
<b>Panels</b>	None

## References

Messer, J.J ., R. A. Linthurst, and W. S. Overton. 1991. An EPA program for monitoring ecological status and trends. *Environmental Monitoring and Assessment* 17:67-78.

Stevens, D. L. and A. R. Olsen. 2003. Variance estimation for spatially balanced samples of environmental resources. *EnvironMetrics*, 14: 593-610.

Stevens, D. L. and A. R. Olsen. 2004. Spatially balanced sampling of natural resources. *Journal of the American Statistical Association*, 99: 262-278.

## APPENDIX A: Annotated R Code Example

### Overview

Below is the R code for the Pajaro River Watershed Condition Assessment sample draw using the 'spsurvey' design package. It is assumed that the user is familiar with R and RStudio. This document does not explain how to install or navigate R or RStudio. Additional support for R and RStudio is available online including general documentation and online tutorials and specific spsurvey package support.

The EPA's Generalized Random Tessellation Stratified (GRTS) survey design guidance documentation and example code are available online: [http://archive.epa.gov/nheerl/arm/web/html/design\\_intro.html](http://archive.epa.gov/nheerl/arm/web/html/design_intro.html). That website includes guidance documentation and examples in R (complete with shape files and code) that can be downloaded and run in order to become familiar with the program and GRTS. A 2015 GRTS stream survey design and sample draw guidance document developed from Tom Kincaid of EPA is available from the CRAN.R.Project site: [http://cran.r-project.org/web/packages/spsurvey/vignettes/Linear\\_Design.pdf](http://cran.r-project.org/web/packages/spsurvey/vignettes/Linear_Design.pdf).

To develop a GRTS sample draw you must install R on your computer<sup>2</sup>. Additionally there are several R interface platforms including Tinn-R and RStudio. SFEI uses the RStudio interactive platform that can also be downloaded online<sup>3</sup>.

Once R is open, the code below can be copied into an R script window and run, however the formatting is not optimal (can also use the Pajaro R code file supplied). Before you begin, however, you should create a work folder on your C drive: The code below expects a folder here: "C:\\R\_Wrk\\PajaroGRTS" with two subfolders "indata" and "outdata". The GIS shape file (sample frame<sup>4</sup>) is placed in the indata folder.

The code below develops two sample draws and plots them as an overlay of the stream sample frame:

1. an unstratified sample of 86 target sites and
2. a stratified sample by PAI (using stratum) and Strahler stream order (using mdcaty).

### Pajaro River Watershed GRTS Sample Draw R Code

```
 #(Note: The "#" symbol is used in R to preface all comments in the code.)
#####
# GRTS Line sample draw (Sarah Lowe, SFEI, Feb.2015)
# # Description of Sample Design
# -----
# Target Population = Riverine wetlands
#
# Study Area: SCVWD Pajaro River Watershed study area within Santa Clara County
# Sample Frame: Santa Clara County Creeks GIS dataset updated with stream order and flow direction.
```

<sup>2</sup> Download R from: <https://www.r-project.org/>

<sup>3</sup> <https://www.rstudio.com/products/rstudio/download/>

<sup>4</sup> Sample\_Frame\_v6\_revised\_20141217 (shape file)

```

# Includes stream orders 1-7 (see Agreement #A3765F Memo 2 & 3 for more information)
# Survey Design = GRTS
# Resource Type = Linear streams (in feet) using SCVWD Creeks ESRI ArcGIS data-layer, accessed - #Sept.2014 and updated
#
# Sample Size = 88
# Over sample = 264 (3X target sample size)
# Total samples = 352
#
# Panels = 1
# Strata = 3 Primary Areas of Interest (or PAIs)
# Multi-density Categories = Yes (Strahler stream order
# to allocate more sites in lower watershed - see code for details)
#
# GIS data Input Projection = UTM zone 10N NAD83 units meters
#####

# Install and/or load R spsurvey package
install.packages("spsurvey") # < this downloads and install the package on your computer
# (only needs to be installed once)
library(spsurvey) #< this loads the library into your open R session

##### SET YOUR WORKING DIRECTORY - for input and output files
setwd("C:\\R_Wrk\\PajaroGRTS")
# Make sure you have a folder on your C drive that includes subfolders ("indata" and "outdata")
# Indata - place GIS dataset (the stream sample frame) here
# outdata - is where the output files will go (e.g. final sample frame shape file)

##### IMPORT GIS DATA SET
# Set file path: inlayer file path is GIS shapefile name w/o extensions (e.g. ".shp")
# prefaced by the folder name where the file is (i.e. existing folder name within the working directory defined above)
inlayer <- "indata/Sample_Frame_v6_revised_20141217" #complete SCVWD Creeks layer

# Import (add to memory) the ArcGIS shapefile for grts and assign it to the variable named "shp" (shape file)
shp <- read.shape(inlayer)

# Import (add to memory) the ArcGIS data-file (.dbf) for grts and assign it to the data-frame named "att" (attribute table)
att <- read.dbf(inlayer)

##### DATA REVIEW AND FORMATTING
# Explore what is in the GIS attribute table
names(att) # get names of headers on att file (note: Shape_Leng is in FEET)
summary(att) # summarize the data within the att file
str(att) # See the internal structure of the data frame

# sum stream lengths for Sample Frame and by SubWatershed
sum(att$Shape_Leng,na.rm = FALSE) # total sample frame length
aggregate(Shape_Leng ~ Subwatersh, data=att, FUN=sum) # sum by Subwatersh
# ouput of sums:
# Subwatersh Shape_Leng (ft.)
#1 Llagas Creek 1326485.0
#2 Other small subwatershed 205785.5
#3 Pacheco Creek 4292736.7
#4 Pajaro River 56627.9

```



```

#5    Pescadero Creek    238115.4
#6    Uvas Creek        1649676.6
#TOTAL length all streams in sample frame = 7,769,427 ft.
#
#Create a new variable for the 4 PAIs in the 'att' table and
# aggregate the small tributary reaches into the "Other" PAI.
# First make Subwatersh a Character variable rather than Factor,
# then create a new variable named PAI from Subwatersh, and
# then rename specific values - in this case we rename 3 subwatersheds to 'Other'.
att$Subwatersh<-as.character(att$Subwatersh)
att$PAI <- att$Subwatersh
att[which(att$PAI=='Pajaro River'),'PAI']<-'Other'
att[which(att$PAI=='Pescadero Creek'),'PAI']<-'Other'
att[which(att$PAI=='Other small subwatershed'),'PAI']<-'Other'
att[which(att$PAI=='Pacheco Creek'),'PAI']<-'Pacheco'
att[which(att$PAI=='Uvas Creek'),'PAI']<-'Uvas'
att[which(att$PAI=='Llagas Creek'),'PAI']<-'Llagas'

#Finally confirm update: check the names in PAI column and review summed stream lengths:
names(att)
aggregate(Shape_Leng ~ PAI, data=att, FUN=sum) # sum by PAI

#####
###SAMPLE DRAW - V1 TRIAL UNSTRATIFIED EQUAL PROBABILITY SAMPLE DRAW
#####
# set design - 86 target sites with equal probability and no oversample sites
Equaldsgn <- list(
  None=list(panel=c(Panel1=86), seltype="Equal",
    over=0))
# define the output shape file (will show up in the outdata folder)
outLayerEqual <- "outdata/V1UnSt"

set.seed(5406611) # this is any random number. It defines a starting point so the draw can be repeated
# run the sample draw - equal probability NO strata
Equalsites <- grts(design=Equaldsgn,
  DesignID="Equal",
  type.frame="linear",
  src.frame="shapefile",
  in.shape=inlayer,
  att.frame=att,
  shapefile= "TRUE",
  out.shape= outLayerEqual)

dsgnsum(Equalsites, auxvar = NULL) # optional - use this to initially review number of sites

# PLOT of RESULTS
# read in the GRTS sample draw out.shape file
outShp <- read.shape(outLayerEqual)
#define the projection of source data as UTM zone10n NAD83
proj4string(outShp)=CRS("+init=epsg:26910")
#CREATE SIMPLE PLOT of inputted stream lines (sample frame (inlayer above) named 'shp')
# and overlay GRTS output points (colored by PAI)
plot(shp, col="lightblue") +

```

```

points(outShp[outShp@data == "PAI"], pch=16, col=outShp@data$PAI)

# end V1 Unstratified sample
#####

#####
# V2a49 - Stratified, Unequal probability, GRTS survey design w/ an oversample
#####
Stratdsgn2 <- list(Pacheco=list(panel=c(Panel1=30),
  seltype="Unequal",
  caty.n=c("1"=2, "2"=4, "3"=6, "4"=6, "5"=6,
    "6"=6),
  over=120),
  Llagas=list(panel=c(Panel1=24),
  seltype="Unequal",
  caty.n=c("1"=1, "2"=2, "3"=5, "4"=5, "5"=6,
    "6"=5),
  over=96),
  Uvas=list(panel=c(Panel1=24),
  seltype="Unequal",
  caty.n=c("1"=1, "2"=3, "3"=5, "4"=5, "5"=5,
    "6"=5),
  over=96),
  Other=list(panel=c(Panel1=10),
  seltype="Unequal",
  caty.n=c("1"=1, "2"=1, "3"=1, "4"=1, "5"=1, "6"=1,
    "7"=4),
  over=40))

Stratdsgn2 #<- review design

#Assign grts sample draw OUTPUT shape file name (will be placed in outdata folder)
outLayerUnequal <- "outdata/StratUnequal_v2a49test"

#set.seed
set.seed(495371)

Stratsites <- grts(design=Stratdsgn2,
  DesignID="StratUnEq_v2a",
  type.frame="linear",
  src.frame="shapefile",
  in.shape=inlayer,
  att.frame=att,
  stratum="PAI",
  mdcaty="Strahler",
  shapefile= "TRUE",
  out.shape= outLayerUnequal)

dsgnsum(Stratsites)

#### Plot the sites ####
outShp <- read.shape(outLayerUnequal) # read in out shape file
proj4string(outShp)=CRS("+init=epsg:26910")

```

```

# set color palette
col.list <- c("red", "blue", "green", "yellow", "orange", "dark gray", "black")
palette(col.list) # make the col.list palette the default for now
# or alternative color palette:
#col.rainbow <- rainbow(7) #assign 7 colors from the col.rainbow palette
# palette(col.rainbow) # make the col.rainbow palette the default for now

symbol <- c(176, 19) #symbols for Overample and Panel1 sites (alphabetically assigned)
plot(shp, col="light grey") +
  points(outShp[outShp@data == "mdcaty"], pch=symbol[outShp@data$panel], cex = 1.2,
         col=outShp@data$mdcaty)+
  legend("bottomleft", inset=c(0.02,0.1), title="Stream \n Order",
         legend = levels(outShp@data$mdcaty), bty= "n",
         col = 1:7, cex = 0.8, pch = 15)

# end V2a49 final sample
#####

```