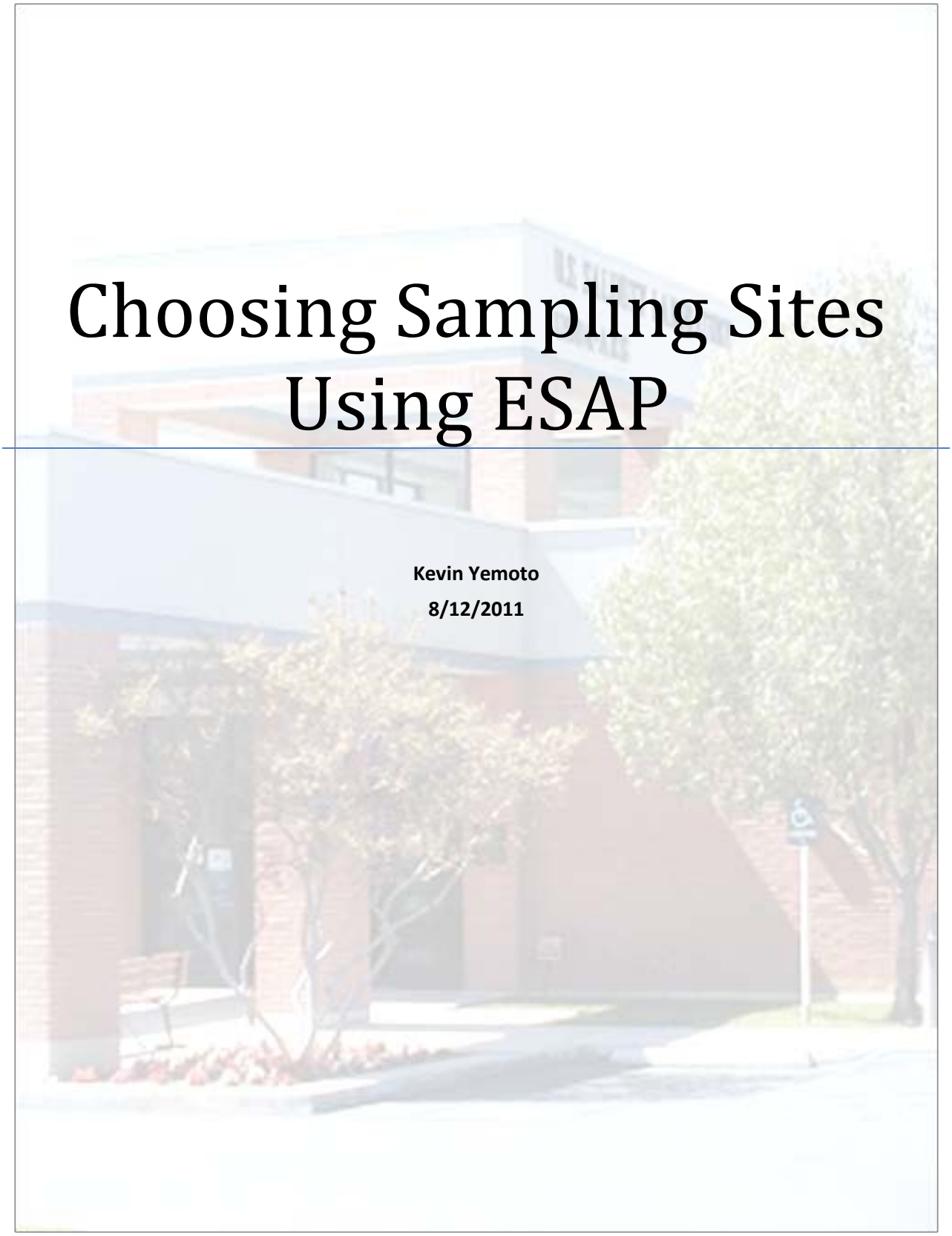
US SALINITY LABORATORY

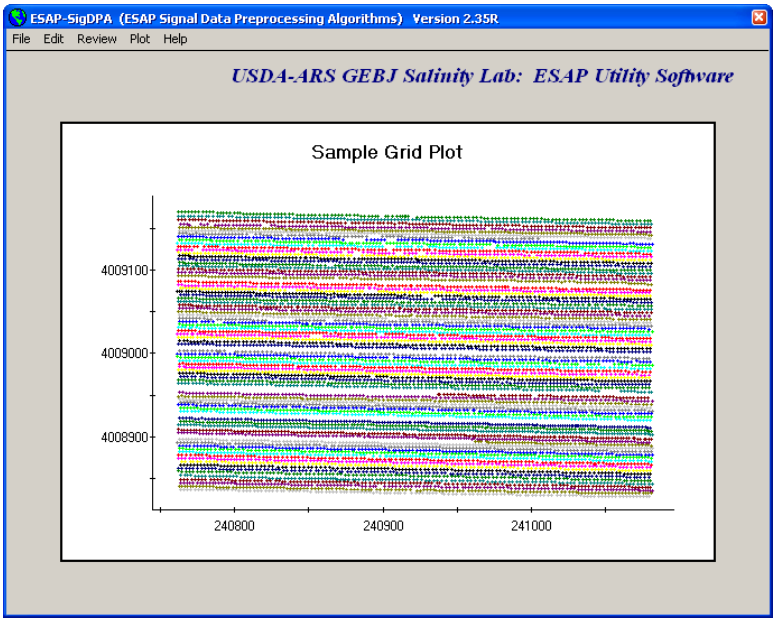
Choosing Sampling Sites Using ESAP



Kevin Yemoto

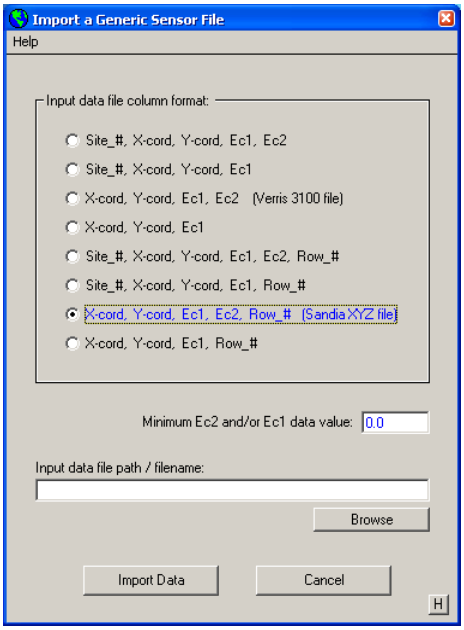
8/12/2011

SigDPA (Signal Data Pre-processing Algorithms)

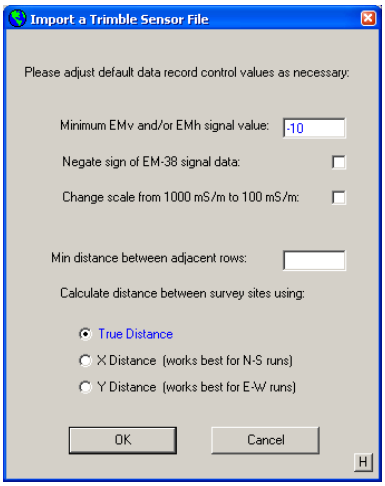


The SigDPA (Signal Data Pre-processing Algorithms) module of ESAP was developed as a tool in which to prepare and "clean up" raw survey information files for input into the RSSD module of ESAP. Import files should be in column aligned or comma-delimited text format containing both GPS and soil conductivity survey data. The file cannot contain more than 30000 sites and both SigDPA and RSSD are limited to 250 transects.

1. Import Data into SigDPA module of ESAP: Note the steps below pertain only to files that are in a column aligned ASCII text format. If survey data file is not in a column aligned ASCII text format and/or site numbers are going to be added to the file, the data file will need to be edited in spreadsheet software and imported into SigDPA as a generic sensor file. Assigning site numbers to the survey file would be a good idea as is allows the user to track which points are being removed. This also makes it easier if a range of points needs to be deleted. When running the RSSD module of program, temporarily id numbers will be assigned to the survey file unless a site id column is present. The temporary id numbers will not be saved to the survey file which may cause problems later on if changes to the survey file structure are made (for example if the order of the data points appear in the file was changed). The column order and information for files that will be imported as a generic sensor file will need to match one of the options seen in the image to the left.

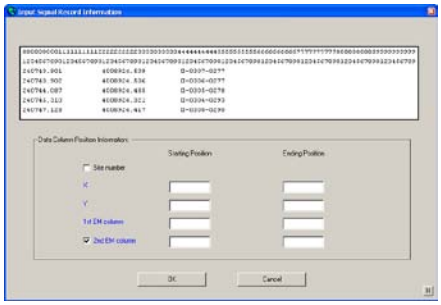


a. Open the SigDPA module of ESAP and import a Trimble/EM38 sensor file.
b. Negate the sign of the EM-38 signal data, specify the distance between rows and specify the how SigDPA should calculate the distance between survey sites.
    i. The polarity of the downloaded raw signal data from the EM-38 is reversed. To correct the polarity of the reading negate the sign of the EM-38 signal data.

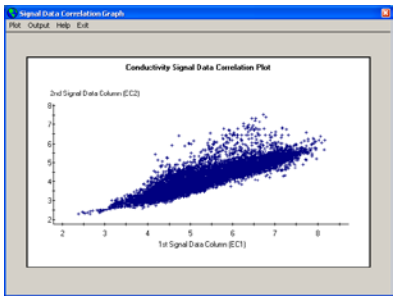


    ii. When surveying a field, readings are typically taken using one of two approaches; a stop-and-go, and a continuous. When taking reading in a stop-and-go fashion, the reading are typically laid out in a grid pattern; evenly spaced from one another. On the other hand, running a survey in a continuous manner usually leads to equally spaced row of readings or transects. SigDPA will calculate the row spacing using one of three options:
        1. True distance: This option uses the true distance between points. This option should be used for grid spaced surveys, or surveys whose transects don't run north south or east west.
        2. X distance: This option uses the difference between neighboring points in only the x direction. This option works best for surveys that have transects that run North-South.
        3. Y distance: This option uses the difference between neighboring points in only the y direction. This option works best for surveys that have transects that run East-West.

c. Indicate the starting and ending column positions for x,y coordinates, and EM signal data.
    i. The first row of numbers indicates the tenths place, and the second row of numbers indicates the ones place. For example the character - is located in number columns 42 and 47.

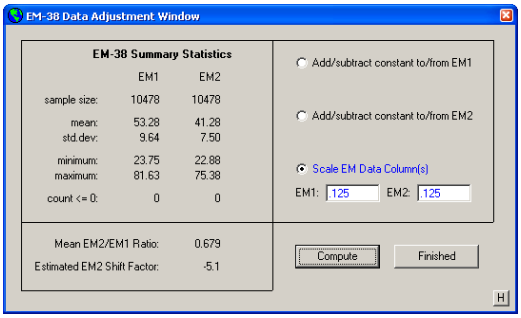


2. Check input data.
    a. The error log file found under the review menu will inform user of any invalid records as well as reasons why the records are invalid.
    b. Signal correlation plots can be viewed by selecting the option under the Plot menu. Typically there should be a correlation between the two signals.

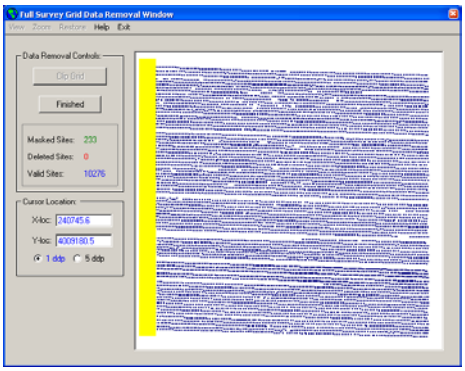


Having a signal correlation plot that doesn't show a correlation between the two signals usually indicates that the EM38 was not calibrated properly.

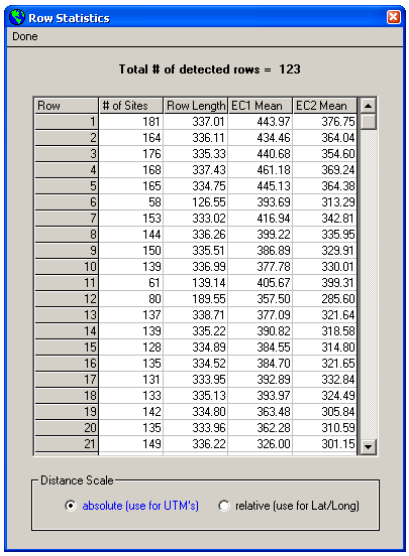
3. Edit Data
   a. Adjust EM signal: It is not recommended to add or subtract values from the raw survey readings. However the raw survey readings are multiplied by 8 when they are translated from the EM38 to the datalogger, therefore to get a correct reading one would need to divide each of the reading by 8 (or in this case multiply all the readings by 0.125).



   b. Clip row ends: (not usually performed on grid spaced surveys) When running a survey in a continuous manner and the end of a row one would either lift the instrument to make the turn or would make the turn with the instrument in contact with the ground. The data collected during the turn between rows in either method will need to be deleted. The data collected if the instrument is lifted when the turn is made results in negative numbers. The extra data collected in the turn, regardless if the instrument is lifted or not will mess up the row spacing calculations. Clip the row ends will help insure that rows are calculated correctly and that erroneous data is removed. Open the full survey grid data removal window for the whole grid by selecting interactive data editing and edit full survey grid under the edit menu. Plot the survey map by selecting one of the options in the view menu. Click clip grid button and draw a rectangle around the end of the rows. Everything that is highlighted in yellow will be deleted. Click finish and then click delete all masked points. Exit the window and return to the main window.



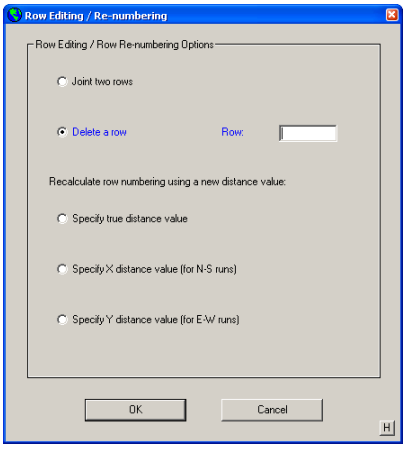
   c. Recalculate row spacing. (not usually performed on grid spaced surveys)
      i. Plot lag distance: The lag distance plot plots either the distance in one axis or the true distance between neighboring points depending on which option is selected. From these plots one can determine a good value to designate the spacing between rows.
      ii. Recalculate rows spacing: Select delete/join/recalculate rows from under the edit menu. Specify method and distance for calculated distance between rows.
      iii. Plot grid and review row statistics: When reviewing the row statistics, pay attention to row length and number of sites. Note if the window for reviewing row statistics is open, and changes to the survey file be it merging rows,



recalculating rows, or deleting row(s), the changes will not automatically be updated in the open window. For an updated table, close the window, and select review row statistics again. There is also no direct method within SigDPA for splitting rows. It is important therefore to set the row spacing to a level in which will split all the rows correctly. As mentioned before there is no direct method to spit specific rows however one can always merge rows that are split in too many pieces. Rows sometimes refuse to split. This usually happens in instances where the surveyor ran the same row twice. If the survey data points are sorted by chronologically, then either split the row into as many pieces as possible and delete the row that loops back on itself or, delete individual points by editing by row or, export the file and use software outside of SigDPA to edit the row and then import the data back in. If the survey data points are sorted by location the only method is to delete the entire row, unless one can differentiate between the points.

d. Delete and merge rows: (Not usually performed on grid style surveys.)

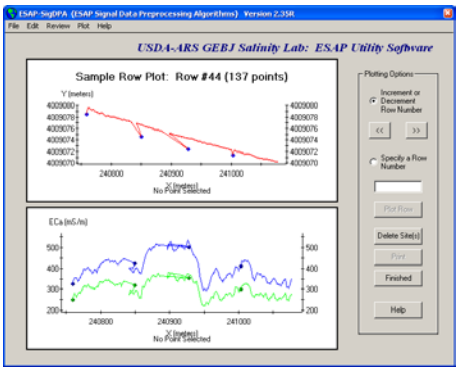
   i. Start by reviewing row statistics. On a piece of paper write down row numbers that have one data point for the entire row. Delete the rows by selecting delete/join/recalculate rows from the edit menu, and one at a time indicating a row number to delete. Note that the window will close after every row that is deleted, so the window will need to be opened up again every time a row is deleted. Also it is a good idea to delete the rows starting with the last row because the row numbers change every time a row is deleted.



   ii. Open the full survey grid data removal window for the whole grid. Plot the survey map, and one at a time select rows under the high-light row option. Note: If editing a large sample file, and if the accessibility tools have been installed on the computer, using the magnifier using inverted colors may help distinguish which row is highlighted. On a piece of paper write down row numbers that will need to be merged. Merge the rows by selecting delete/join/recalculate rows from the edit menu, and one at a time indicating a 2 row number to merge. Note that the window will close after every row that is

merged, so the window will need to be opened up again every time rows are merged. Once again it is a good idea to merge rows starting with the last row because the row numbers change every time a row is merged.

e. Remove points within survey and individual rows.

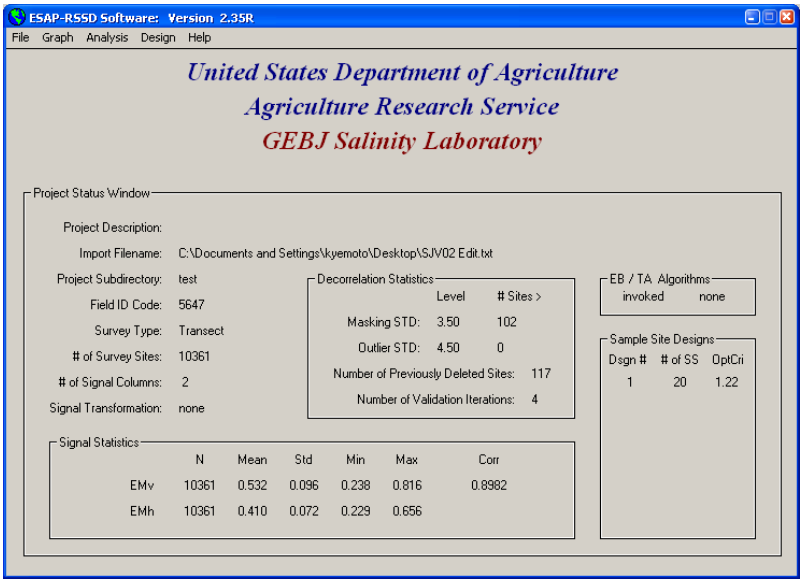


i. Points within the entire survey can be deleted by using the full survey grid data removal window again. This method is useful if one needs to delete blocks of points in the survey.

ii. Points within rows can be deleted by selecting edit individual rows option under interactive data editing under the edit menu. Points that need to be removed can be easily distinguished as a spike in the line. That point or points can be deleted by clicking on the points and then clicking delete site(s) button. Click finished when editing is done.

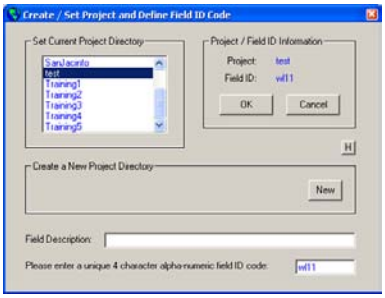
4. Export data: Export the data from the file menu. Select an output name and location for the file. If input file was a grid file select remove all row numbers. Finally select create and save the file button.

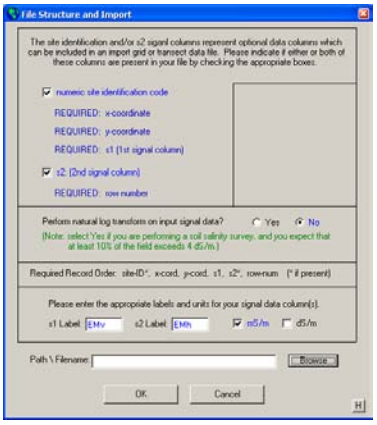
RSSD (Response Surface Sampling Design)



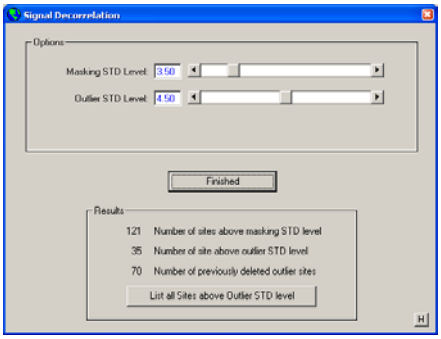
The RSSD (Response Surface Sampling Deign) module of ESAP is a statistical program that will generate an optimal soil sampling design from the survey information. Import files should be in tab delimited ASCII text format containing both GPS and soil conductivity survey data. Column heading should not be included in the file format. The file cannot contain more than 30000 sites and both SigDPA and RSSD are limited to 250 transects.

1. Create / Set Project and Define Field ID Code.
    a. Create or set a directory for which project files will be saved. Enter a unique 4 character alpha-numeric field id code for the file.
2. Import Survey data file
    a. Import the grid or transect survey data file.
    b. Indicate which fields are included in the data file. Also indicate an appropriate label and units for the different signal data column(s).
    c. Locate and select appropriate survey file name.
3. Graph the Histograms for signal data
    a. Open the interactive survey graphics mode window by opening the graphics window from the graph menu.
    b. Under the graphics window plot the histograms for input signal(s). The RSSD program works better with information that is normally distributed. By plotting the histograms will indicate whether the data should be log transformed.

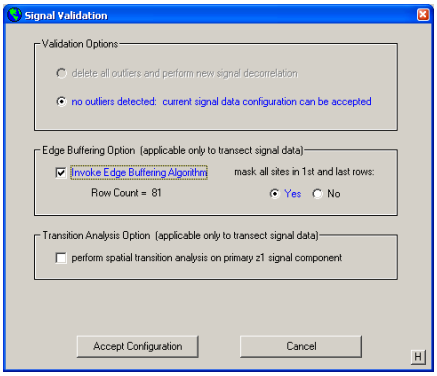




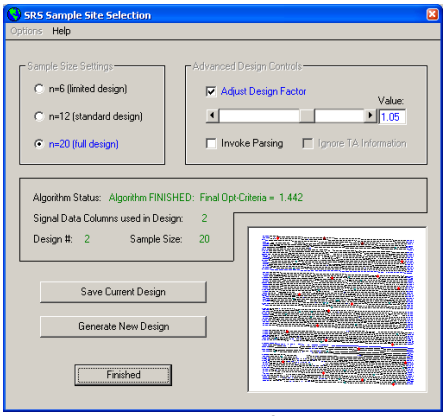
4. Log correct data: (if necessary) Apply or remove a natural log transformation by selecting basic statistics from the analysis menu. Applying or removing the natural log transformation can only be done to both input signals rather than each individual input signal.
5. Invoke Signal Decorrelation: Signal decorrelation will search for outliers. The default levels are set at 3.5 standard deviations for masking sites and 4.5 standard deviations for designating sites as outliers. Open the Signal decorrelation from the analysis window and perform a decorrelation. A list of sites that are above the outlier level can be viewed by clicking the list all sites about outlier STD level.



6. Invoke Signal Validation: At this time sites that are marked as outliers can simply be masked from being considered or can be deleted. Select signal validation from the analysis menu, and select option to delete all outliers or mask all outliers.
    a. If the option to delete all outlier is selected it will be necessary to perform a new signal decorrelation. Select the option to delete all outliers and click on invoke validation button. This completes one iteration cycle. Perform a new decorrelation then repeat step 6. (note: if one continues to delete all outliers by repeating step 6., eventually all possible outliers will have been masked and deleted in which case the option to delete all outliers will be blacked out. At this point continue with steps below.)

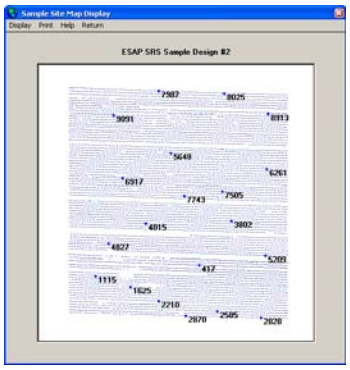


    b. If the option to mask all outliers is selected or the option to delete all outliers has been exhausted, select invoke edge buffering algorithm and mask all sites in first and last row (this will only be an option if survey file is a transect file). Performing a transition analysis option is best applied if GPS coverage or real-time accuracy of GPS equipment does not allow sub-meter accuracy. Select appropriate options and click on Accept configuration button to continue.
7. Calculate SRS Sample site selection: Select Calculate SRS Sample design from design menu. RSSD allows you 6 different sample size settings: 5,6,10,12,18, or 20 samples. (5, 10, and 18 sample size settings can be selected by selecting the option to exclude support sites from the options menu). Invoke SRSS auto-design selection algorithm and run a full ADS alg design by selecting the option from the options menu select. The program will then run simulations on 63 different sampling design specifications and will return a table indicating which setting will yield the best design. (note: anything below a score of 1.3 would be



considered adequate.) The F number at the top of the table indicates the Design Factor value. The different strata levels can be set by invoking parsing. Invoke SRSS Algorithm to generate the sampling design using the set values. Save current design and either generate a new design or click finished. Note when generating a new design after a design is saved will remove the site from the saved design from consideration. Therefore one would need to invoke SRSS auto-design selection algorithm again to get generate the best sampling design as the score will reflect the best design given the remaining sites. Up to 5 different designs may be saved.

8. Display and print sampling design map: select display sample site map under the design menu. Display and print the different SRS sampling designs. The map shows all the points in the survey and highlights and labels the sampling sites chosen.



9. View and print output files: The different output files for the saved sampling designs can be viewed by selecting view and print output files from the file menu. RSSD will create one file with some basic statistics as well as information on how many sites were masked and deleted. RSSD will also create 2 files for every saved sampling design. One of the files will indicate the settings used the other will give GPS coordinates, sample ids, and row numbers (if a transect file was used) for the sites chosen.