Methodology for Flow and Salinity Estimates in the Sacramento-San Joaquin Delta and Suisun Marsh

33rd Annual Progress Report
June 2012

Chapter 1
Monitoring Station Locations

Authors: **Ralph Finch and Jane Schafer-Kramer**
Delta Modeling Section
Bay-Delta Office
California Department of Water Resources

Page left blank for two-sided printing

Contents

Figure

Tables

Page left blank for two-sided printing

1 Monitoring Station Locations

1.1 Introduction

In late 2010, the authors began to examine the DSM2 grid and Sacramento-San Joaquin River Delta bathymetry and geometry using ArcMap. Several deficiencies were noted in the current grid, among them the placement of nodes (channel junctions and ends) in the geo-referenced ArcMap. This led to the realization that we did not have accurate locations of the observed data monitoring stations. At that time, we had a single source of locations from the California Data Exchange Center (CDEC), and many were in error by several hundred meters or more. In 2011, both authors acquired lists of purported accurate station locations, and the authors conducted field measurements of some sites. We eventually had 10 lists of station locations including our own field measurements. We analyzed these lists using ESRI products and a script in ArcPython; this chapter describes and summarizes the analysis.

1.2 Procedure

We ended with 9 station lists, other than our measured stations, internally named[1]

- CDEC_Delta
- FlowStations
- NCRO_FlowStation_Oct2011
- Stations_SMayr
- SurfaceWater
- SurfaceWater_fr_GSmith
- USBR_Others
- WaterQuality
- waterquality_Stations_D1641

With so many lists of potentially good locations, two tasks became obvious:

1. estimate which stations and lists have accurate station location, and
2. estimate which stations are duplicated in other lists (in order to consolidate them).

We measured the locations of 10 stations in the North Delta to spot-check stations in other lists, and then found nearby stations—within 100 meters—in other lists using ArcPython and Arc Catalog (Table 1-1). An ArcPython script was written to perform this and other tasks described in this chapter and is reproduced as an appendix at the back of this chapter.

---

[1] The list names were made when the authors received each list and refer to type and source of data.

Table 1-1 Partial List of Stations and Station-Lists near Measured Stations

	OBJECTID *	IN_FID	NEAR_FID	NEAR_DIST	StationList	StaListName	BaseStaID
	20	1	0	0	None	None	Dlc
	1	2	22	11.3424413611733	SurfaceWater	B91120	Emm
	6	2	22	11.3423426810068	Stations_SMayr_Proj	B91120	Emm
	10	2	15	73.6189000975313	CDEC_Delta_Proj	EMM	Emm
	13	2	8	129.353842404674	USBR_Others_Proj	EMMATON	Emm
	15	2	56	11.3420830592777	SurfaceWater_fr_GSmith_Proj	B91120	Emm
	17	2	24	235.64324379616	waterquality_Stations_D1641_Proj	D22	Emm
	18	2	23	79.5521768792677	waterquality_Stations_D1641_Proj	D22A	Emm
	21	3	0	0	None	None	Fpt
	11	4	20	18.009518605283	CDEC_Delta_Proj	GES	Ges
	7	5	31	13.7022597024286	Stations_SMayr_Proj	B91651	Ggs
	2	6	32	20.2793525315143	SurfaceWater	B94100	Gsmr
	3	7	24	11.132401848978	SurfaceWater	B91212	Rvb
	8	7	24	11.1324547238763	Stations_SMayr_Proj	B91212	Rvb
	12	7	63	139.979675713814	CDEC_Delta_Proj	SRV	Rvb
	14	7	15	310.110796472058	USBR_Others_Proj	RIO VISTA	Rvb
	19	8	7	88.4415861668535	waterquality_Stations_D1641_Proj	C3A	Srh
	5	9	8	61.2312945804588	WaterQuality_Proj	Three Mile Slough near SJR	Tmsnsjr
	16	9	152	61.2318010725615	SurfaceWater_fr_GSmith_Proj	TMS	Tmsnsjr
	4	10	38	43.042055640107	SurfaceWater	B95060	Tsl
	9	10	39	43.0417589480545	Stations_SMayr_Proj	B95060	Tsl

Table note: the partial listing shown here does not contain all stations and station-lists.

The OBJECTID column is used as a primary key by ArcCatalog. The IN_FID and OUT_FID columns are the primary keys for the measured station and nearby station in another list, respectively. NEAR_DIST is the distance from the measured station to the nearby station in feet; 0 means no nearby station was found in any list. StationList is the name of the list that has the nearby station, and StaListName is the name of the station in the other list. A measured station that has no nearby stations in other lists has a value "None" under each of these columns. Finally, BaseStaID is the name of the measured station.

If a measured station is found to have more than one nearby station from the other lists, each station is listed in a new row. In Table 1-1, there are 10 measured stations. Two have no nearby stations in any lists (DLC and FPT) while EMM had matches in 6 other lists. Nearby stations were found in 7 other lists, with only lists FlowStations and NCRO_FlowStation_Oct2011 failing to have nearby stations. We also know that the CDEC list contains some stations with accurate locations and some inaccurate; our first task is to estimate if other lists might be accurate.

Table 1-2 shows the total number of stations and the number of stations within the legal Delta and Suisun Marsh for each station list we have. There are at least 98 stations in the Delta and Marsh (Figure 1-1), using the highest figure in Table 1-3 (CDEC_Delta). We field-measured the locations of 10 stations (GarminWaypoints), or about 10% of the total Delta stations. For any given list, assuming that its stations are scattered throughout the Delta, we might find the same proportion of stations that match the measured stations. This also assumes that the measured stations are scattered evenly throughout the Delta, when in fact we know they are in the North Delta.
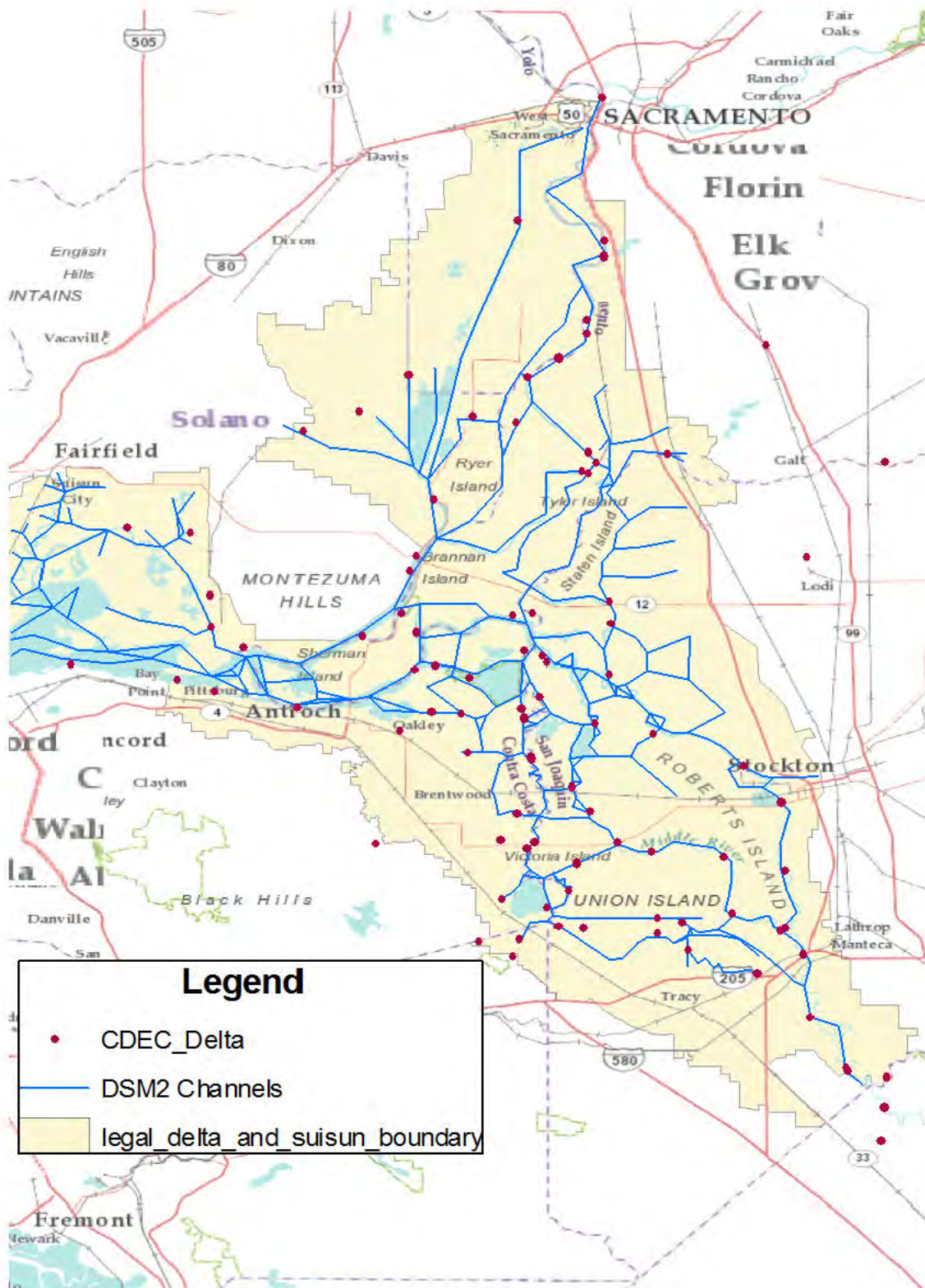
Figure 1-1 Sacramento-San Joaquin Legal Delta, DSM2 Channels, and CDEC Stations

Table 1-2 Station-Lists and Counts

Station-list name	Total number of stations	Number of stations within Delta+Marsh
CDEC_Delta	106	98
FlowStations	11	10
GarminWaypoints	10	10
NCRO_FlowStation_Oct2011	12	11
Stations_SMayr	77	57
SurfaceWater	75	55
SurfaceWater_fr_GSmith	269	60
USBR_Others	24	19
WaterQuality	32	32
waterquality_Stations_D1641	52	49

Table 1-3 is the same as Table 1-1, but sorted on the StationList column. This way it's easier to count how many stations in each list match the measured stations.

Table 1-3 Partial List of Stations and Station-Lists near Measured Stations, Sorted by Station-List

OBJECTID *	IN_FID	NEAR_FID	NEAR_DIST	StationList	StaListName	BaseStaID
10	2	15	73.6189000975313	CDEC_Delta_Proj	EMM	Emm
11	4	20	18.009518605283	CDEC_Delta_Proj	GES	Ges
12	7	63	139.979675713814	CDEC_Delta_Proj	SRV	Rvb
20	1	0	0	None	None	Dlc
21	3	0	0	None	None	Fpt
6	2	22	11.3423426810068	Stations_SMayr_Proj	B91120	Emm
7	5	31	13.7022597024286	Stations_SMayr_Proj	B91651	Ggs
8	7	24	11.1324547238763	Stations_SMayr_Proj	B91212	Rvb
9	10	39	43.0417589480545	Stations_SMayr_Proj	B95060	Tsl
1	2	22	11.3424413611733	SurfaceWater	B91120	Emm
2	6	32	20.2793525315143	SurfaceWater	B94100	Gsmr
3	7	24	11.132401848978	SurfaceWater	B91212	Rvb
4	10	38	43.042055640107	SurfaceWater	B95060	Tsl
15	2	56	11.3420830592777	SurfaceWater_fr_GSmith_Proj	B91120	Emm
16	9	152	61.2318010725615	SurfaceWater_fr_GSmith_Proj	TMS	Tmsnsjr
13	2	8	129.353842404674	USBR_Others_Proj	EMMATON	Emm
14	7	15	310.110796472058	USBR_Others_Proj	RIO VISTA	Rvb
5	9	8	61.2312945804588	WaterQuality_Proj	Three Mile Slough near SJR	Tmsnsjr
17	2	24	235.64324379616	waterquality_Stations_D1641_Proj	D22	Emm
18	2	23	79.5521768792677	waterquality_Stations_D1641_Proj	D22A	Emm
19	8	7	88.4415861668535	waterquality_Stations_D1641_Proj	C3A	Srh

Table 1-4 Proportion of Stations near Measured Stations

Station list name	Number of stations within Delta+Marsh	No. of stations near measured stations	Proportion of stations near measured stations
CDEC_Delta	98	3	$\frac{3}{98} * 100 = 3.1\%$
Stations_SMayr	57	4	$\frac{4}{57} * 100 = \textbf{7.0}\%$
SurfaceWater	55	4	$\frac{4}{55} * 100 = \textbf{7.3}\%$
SurfaceWater_fr_GSmith	60	2	$\frac{2}{60} * 100 = 3.3\%$
USBR_Others	19	2	$\frac{2}{19} * 100 = \textbf{10.5}\%$
WaterQuality	32	1	$\frac{1}{32} * 100 = 3.1\%$
waterquality_Stations_D1641	49	3	$\frac{3}{49} * 100 = \textbf{6.1}\%$

Finally we calculate the proportion of nearby stations to number of Delta and Marsh stations (Table 1-4).

Previously, we reasoned that if a given list had accurate station locations and its stations were uniformly scattered throughout the Delta+Marsh and if the measured stations were uniformly scattered throughout the Delta+Marsh, we would find about 10% of stations in a given list near the measured stations. But we know the measured stations are not uniformly located throughout the Delta and Marsh; they are concentrated in the North Delta. Therefore, we expect a given list with accurate locations to have somewhat less than 10% nearby stations.

From Table 1-4, four lists have 6% to 10% of their stations within 100 m of the measured stations; we conclude those lists have accurate locations unless shown otherwise (Stations_SMayr, SurfaceWater, USBR_Others, waterquality_Stations_D1641).

Table 1-5 is a portion of a larger table produced by the same script that produced Table 1-1 and Table 1-3. The snippet shown in Table 1-5 contains the stations in the SurfaceWater list (which seems to have accurate stations) that are near the WaterQuality list. Note the 4th column, which is the distance in feet between stations in both lists. Of the 6 close stations, 2 are probably from an identical measurement because their distance apart is a small fraction of a foot: numerical round-off. The other 4 nearby stations are within 10s of feet. So for the WaterQuality list, SurfaceWater has 6 close stations out of 55 in the Delta+Marsh, a proportion of 10.9%. In a similar analysis, 6 of 57 stations in the Stations_SMayr list are near, a proportion of 10.5%. We conclude the SurfaceWater list also has accurate stations.

Table 1-5 Partial Water Quality List and nearby Stations in Surface Water List

4	14	50	25.1335008014342	SurfaceWater	B95325	Doughty Cut above GLC
5	16	49	17.4214984379326	SurfaceWater	B95300	GLC at Tracy Road
6	21	64	1.38062553001212E-05	SurfaceWater	B95500	Middle River at Union Point
7	22	53	30.8798041703845	SurfaceWater	B95366	Old River at DMC - Above
8	23	52	4.38219294088364E-05	SurfaceWater	B95365	Old River at DMC - Below
9	25	55	18.10806906965	SurfaceWater	B95400	Old River near Head

In like manner we analyzed the SurfaceWater_fr_GSmith list. Forty-one different stations from the Stations_SMayr list out of 57 Delta+Marsh stations were near (71.9%); 31 of the 41 distances were within round-off error. Forty-two different stations of 55 (76.4%) from the SurfaceWater list were near, and 31 were within round-off error. We conclude that SurfaceWater_fr_GSmith probably has accurate locations, and many are duplicates of other lists.

This leaves only the 2 flow station lists as unknown accuracy (we know CDEC has inaccurate stations). Flow stations are often not located with stage or water quality stations so we will ignore those stations; their location is not needed for the task of re-locating DSM2 nodes.

1.3   Conclusion

We conclude that the station locations in the lists we were given are generally accurate. They will be assumed to be accurate until shown otherwise. The 2 flow station lists have not been checked, but will be in the future.

We know CDEC has a mixture of accurate and inaccurate locations, for reasons unknown. There may be a way of reliably predicting which stations in CDEC are accurately located (involving the number of significant digits in their latitude and longitude values).

The next major step in this project will be to consolidate duplicate stations and generate a list of common and alternate station names. We plan on using CDEC names as the base names, generating names similar to the CDEC style where necessary.

Appendix ArcPy Script

```
import arcpy, os, time
from arcpy.mapping import *
from arcpy.analysis import *
from arcpy.management import *
from arcpy import *
## For GPS-measured stations, find nearby stations in other lists

workspaceDir1 = "D:/delta/GIS/Observed/Monitoring Stations/"
workspaceDir2 = workspaceDir1 + "Jane 20110802/"
commonDir = "C:/Users/rfinch/Documents/ArcGIS/Packages/"
outDir = "Z:/Temp"
scratchWorkspace = "z:/Temp"
overwriteOutput = True


# the GPS-measured stations DB and layer
GPSMeasGDB = workspaceDir2 + "DeltaStationsGPS.gdb/"
GPS_lyr = GPSMeasGDB + "GarminWaypoints"
# the provided lists of station locations to check
StationListsNCRO = workspaceDir1 + "Stations_NCRO/BranchStations.mdb/"
StationListsGDB = workspaceDir2 + "DeltaStationLists.gdb/"
#
SMayr_lyr = StationListsGDB + "Stations_SMayr_Proj"
CDEC_lyr = StationListsGDB + "CDEC_Delta_Proj"
USBR_lyr = StationListsGDB + "USBR_Others_Proj"
SW_lyr = StationListsGDB + "SurfaceWater_fr_GSmith_Proj"
WQD1641_lyr = StationListsGDB + "waterquality_Stations_D1641_Proj"
NCRO_Oct2011_lyr = StationListsGDB + "NCRO_FlowStation_Oct2011_Proj"
NCRO_Flow_lyr = StationListsNCRO + "FlowStations"
NCRO_SW_lyr = StationListsNCRO + "SurfaceWater"
NCRO_WQ_lyr = StationListsNCRO + "WaterQuality_Proj"
# for each station list, which field is the primary station name field
NameFields = {GPS_lyr: "station", SMayr_lyr: 'STA_NO', CDEC_lyr: 'CDEC_ID', USBR_lyr: 'StationDescription', \
              SW_lyr: 'Site_ID', WQD1641_lyr: 'StationID', NCRO_Oct2011_lyr: "Internal_c", \
              NCRO_Flow_lyr: "Name", NCRO_SW_lyr: "Station_No", NCRO_WQ_lyr: "Station_Na"}
# A generic primary station name field
genericStaID = 'BaseStaID'
# which station lists to check
```

```
StationLyrs = [NCRO_Oct2011_lyr, NCRO_Flow_lyr, NCRO_SW_lyr, NCRO_WQ_lyr, \
               SMayr_lyr, CDEC_lyr, USBR_lyr, SW_lyr, WQD1641_lyr, GPS_lyr]
# where to put output tables
env.workspace = StationListsGDB
# find other stations with searchRadius of each Base list location
searchRadius = '100 Meters'
location = 'NO_LOCATION'
angle = 'NO_ANGLE'
closest = 'ALL'
closestCount = 5
tempTable = GPSMeasGDB + "temp"
for baseLyr in StationLyrs:
    temp = os.path.basename(baseLyr).replace('_lyr','')
    temp = temp.replace('_Proj','')
    outTable = 'Nearest_' + temp
    # get OID field of base table
    BaseOIDFld_lst = ListFields(baseLyr, '', 'OID')
    BaseOIDFld_nm = BaseOIDFld_lst[0].name
    print "Base List:", os.path.basename(baseLyr), "OID:", BaseOIDFld_nm
    # total number of stations in base layer, and nearest stations found
    nStasBase = int(GetCount(baseLyr).getOutput(0))
    nStasFound = 0
    # create nearest table...
    lyrCount = 0
    for lyr in StationLyrs:
        shortLyr = os.path.basename(lyr)
        if lyr == baseLyr:
            continue
        #print 'List: ',shortLyr
        try: Delete(tempTable)
        except: pass
        desc = Describe(lyr)
        if not desc.hasOID:
            print 'No ObjectID field, skipping layer', shortLyr
            continue
        # put the nearest table list into the temporary table...
        # we will add fields to it for the permanent table.
        GenerateNearTable(baseLyr, lyr, tempTable, searchRadius,location, angle, closest, closestCount)
        # now join the Base table to the nearest table...
        # get correct field delimiters
        delmField = arcpy.AddFieldDelimiters(tempTable, 'NEAR_FC')
```

```
        # ...join to get location names
        JoinField(tempTable, 'in_fid', baseLyr, BaseOIDFld_nm, NameFields[baseLyr])
        if lyrCount == 0:
            # Create permanent table
            try: Delete(GPSMeasGDB + outTable)
            except: pass
            CreateTable(GPSMeasGDB, outTable, tempTable, '')
            AddField(GPSMeasGDB + outTable, 'StationList', 'text', '', '', 50, '', '', '', '')
            AddField(GPSMeasGDB + outTable, 'StaListName', 'text', '', '', 254, '', '', '', '')
        # ...join the station name from the target station list
        JoinField(tempTable, 'near_fid', lyr, desc.OIDFieldName, NameFields[lyr])
        AddField(tempTable, 'StationList', 'text', '', '', 50, '', '', '', '')
        rows = UpdateCursor(tempTable, '', '', '', '')
        for row in rows:
            row.StationList = shortLyr[0:49]
            rows.updateRow(row)
        # Create FieldMappings object for append output fields
        fieldMappings = FieldMappings()
        # Add all fields from tempTable
        fieldMappings.addTable(tempTable)
        fldMap_staListName = fieldMappings.getFieldMap(fieldMappings.findFieldMapIndex(NameFields[lyr]))
        # Set name of permanent output field StaListName
        fld_staListName = fldMap_staListName.outputField
        fld_staListName.name = "StaListName"
        fldMap_staListName.outputField = fld_staListName
        fieldMappings.addFieldMap(fldMap_staListName)
        fieldMappings.removeFieldMap(fieldMappings.findFieldMapIndex(NameFields[lyr]))
        Append(tempTable, GPSMeasGDB + outTable, 'NO_TEST', fieldMappings, '')
        lyrCount += 1
try: arcpy.management.Delete(tempTable)
except: pass
# Checking against all station lists is done
# Check the output table for stations in the base layer that had no near matches
# and add the generic station id field, deleting the base list station id field
AddField(GPSMeasGDB + outTable, genericStaID, 'TEXT', '','', '50', '', 'NULLABLE', 'NON_REQUIRED', '')
staNumListPrev = 0
missingStas = []
rows = UpdateCursor(GPSMeasGDB + outTable, '', '', '', 'IN_FID A')
for row in rows:
    row.setValue(genericStaID,row.getValue(NameFields[baseLyr]))
    rows.updateRow(row)
```

```python
        staNumList = long(row.IN_FID)
        if staNumList == staNumListPrev:
            continue
        while staNumList-staNumListPrev > 1:
            missingStas += [staNumListPrev+1]
            staNumListPrev += 1
        staNumListPrev = staNumList
    DeleteField(GPSMeasGDB + outTable, NameFields[baseLyr])
    del row, rows
    rows = InsertCursor(GPSMeasGDB + outTable)
    nStasFound = nStasBase - len(missingStas)
    for sta in missingStas:
        row = rows.newRow()
        rowsBase = SearchCursor(baseLyr, BaseOIDFld_nm+" = "+str(sta), "", NameFields[baseLyr], "")
        for rowBase in rowsBase:    # should be only 1 row
            baseStaID = rowBase.getValue(NameFields[baseLyr])
            row.setValue(genericStaID,baseStaID)
        row.IN_FID = sta
        row.NEAR_FID = 0L
        row.NEAR_DIST = 0.0
        row.StationList = 'None'
        row.StaListName = 'None'
        rows.insertRow(row)
        #print "No near neighbor in", os.path.basename(baseLyr), "for station", baseStaID
    print 'Total stations', nStasBase, 'Stations Nearest', nStasFound
    print
print "Finished"
```