6-30-2003

Categorization of shellfish TMDL sites Final Report

Julie Herman
Virginia Institute of Marine Science

Carl Hershner
Virginia Institute of Marine Science

Howard Kator
Virginia Institute of Marine Science

Recommended Citation
Herman, J., Hershner, C., & Kator, H. (2003) Categorization of shellfish TMDL sites Final Report. Virginia Institute of Marine Science, College of William and Mary. https://doi.org/10.21220/V59908

Categorization of shellfish TMDL sites

Final Report

The Virginia Institute of Marine Science
College of William and Mary

Julie Herman and Carl Hershner
Center for Coastal Resources Management

Howard Kator
Department of Environmental and Aquatic Animal Health

June 30, 2003

Chapter 1. GIS database

There were two important goals for this project, including the assembly and compilation of digital data for the Coastal Plain and tidal waters of Virginia, and the categorization of the Division of Shellfish Sanitation (DSS) shellfish growing areas to determine if some are similar enough for water quality models to be effectively transferred from the modeled growing areas to other areas. This report summarizes the data and statistical analyses and discusses the results. Note: The digital data has a very fine resolution. The maps displayed in this report cannot convey this information on 8.5x11" paper. All digital data layers (shape files and grids), digital maps (jpgs), data files (excel files) and a digital copy of this report (pdf format) are included on 3 cdroms that accompany this report. The cdroms include a text file (called Readme.txt) that provides an explanation of each of the data layers and pertinent information in the tables contained in the data layers.

Data Collection

Using the Sanitary Survey maps from DSS as the source, the boundaries for each survey were digitized using Arc/Info with digital raster graphs (digital versions of USGS topographic maps) as the background. DSS typically refers to these survey boundaries as growing area boundaries. See Figure 1.

Various digital data layers were obtained for the region covered by the growing areas. These layers include some that were generated or modified by VIMS as well as layers from other government agencies. These layers (see Figures 2 to 8) include:

- NLCD (National Land Cover Data; U.S. Geological Survey, 1999). The digital version is nlcd on cdrom 2. Includes 15 land use/land cover types in raster format with 30x30m cells.

- NED (National Elevation Data; U.S. Geological Survey, 1999). The digital versions are ned and ned_int on cdrom 2. Elevations are in meters, in raster format with 30x30m cells.

- SSURGO and STATSGO (soils data from NRCS, 1995 to 2000). The digital version is soils.shp on cdrom 1. Polygons include soil type and depth-weighted average surface permeabilities in inches/hour.

- population density data (US Census Bureau, 2000). The digital version is popdens.shp on cdrom 1. Polygons include population densities in people/hectare.

- bathymetry (NOS Hydrographic Survey data; multiple years)–the shallow waters were used to calculate water volumes for each growing area. The digital versions are bathym and bathym_int on cdrom 2. Depths are in meters, in raster format with 30x30m cells.

- condemnation zones (DSS data; multiple years). The digital version is condemnzones.shp on cdrom 1. Polygons include growing area and condemnation zone numbers.

- Sanitary Survey deficiencies (DSS data; multiple years). The digital version is deficiencies.shp on cdrom 1. Points include growing area numbers, counties, deficiency labels, and pollution types and impact (i.e. all the information that is presented on DSS Sanitary Survey maps).

Statistical Analyses

For each growing area all the data from the first 5 layers listed above were calculated and input into a statistical software package (JMP; SAS Institute, 2001). Growing areas were grouped using hierarchical cluster analysis (with average linkage). Best professional judgement was used to determine which parameters from the data were critical for a reasonable set of clusters.

Six parameters were identified as the most important factors needed to cluster the growing areas.

1. Ratio of area of developed land:total area of growing area–this represents the input of fecal coliform from people and pets.
2. Ratio of area of forest and wetlands:total area of growing area–this represents the input of fecal coliform from wild animals.
3. Ratio of area of pasture land:total area of growing area–this represents the input of fecal coliform from farm animals.
4. Ratio of area of low elevation (0 to 3 meters) to total area in a growing area–this represents the percentage of low elevation in each growing area, and is one indicator of delivery of fecal coliform from sources to the water body. Lower elevations have lower slopes but may have more direct contact with water bodies.
5. Standardized area-weighted average surface permeability–this is another indicator of delivery of fecal coliform from sources to the water body. Intermediate permeabilities probably provide the most runoff, but poorly drained soils in low elevations (e.g. wetlands) may have direct contact with water bodies. To standardize, the data were divided by 20 (maximum possible permeability).
6. Standardized ratio of land area in a growing area to volume of its water body–this is an indicator of dilution of fecal coliform; the bigger the ratio the more potential sources there are relative to the amount of water available to receive them. To standardize, the data were divided by the maximum ratio.

Discussion

Numerous cluster analyses were run with many different parameters. The 6 final parameters (listed above) were used for 3 different cluster analyses, including an analysis using all 6 parameters, an analysis using the 3 source parameters, and an analysis using the 2 delivery parameters. The results are displayed as dendrograms (Figures 9 through 11) with the growing area numbers on the left and the branches of the dendrogram on the right. The shorter the branches, the more closely related adjacent numbers. Clusters are color-coded and match the colors on the growing area maps (Figures 12 through 14). Appendix 1 contains a table with the values for the 6 parameters and the cluster numbers for the 3 separate analyses.

All 6 parameters are shown clustered in Figures 9 and 12. A majority of the Eastern Shore growing areas and many of the low-lying growing areas on the western shore are clustered together, probably reflecting the coincidence of low elevation and low permeability soils (e.g. wetlands). The growing areas with the most developed land (and the highest population densities) cluster into several small groups. The source clusters (Figures 10 and 13) have two large clusters containing a majority of the growing areas, distinguished mainly by the amount of forest + wetland or the amount of pasture land. The high population densities again are separated from the other regions. The delivery clusters (Figures 11 and 14) show a slightly more even clustering of growing areas, which may be due to only using 2 parameters or to a relationship between elevation and soil permeability.

Conclusions

Most growing areas tended to group into just a few clusters. This pattern is probably due to the somewhat homogeneous nature of parameters in the Coastal Plain, e.g. low relief, mostly forested or agriculture with high population densities concentrated in a few areas. Some of the growing areas that fall into the same clusters in all three iterations may have enough similarities for one water quality model to be transferred between them. The rest of the growing areas require separate models. The digital data developed for this project are an integral part of the ongoing development of TMDLs for shellfish growing areas in Virginia, and are detailed enough to support modeling for individual growing areas.

Figures

Figure 1. Sanitary Survey boundaries, also called growing area boundaries (DSS data; multiple years). The digital version is boundmap.jpg on cdrom1.

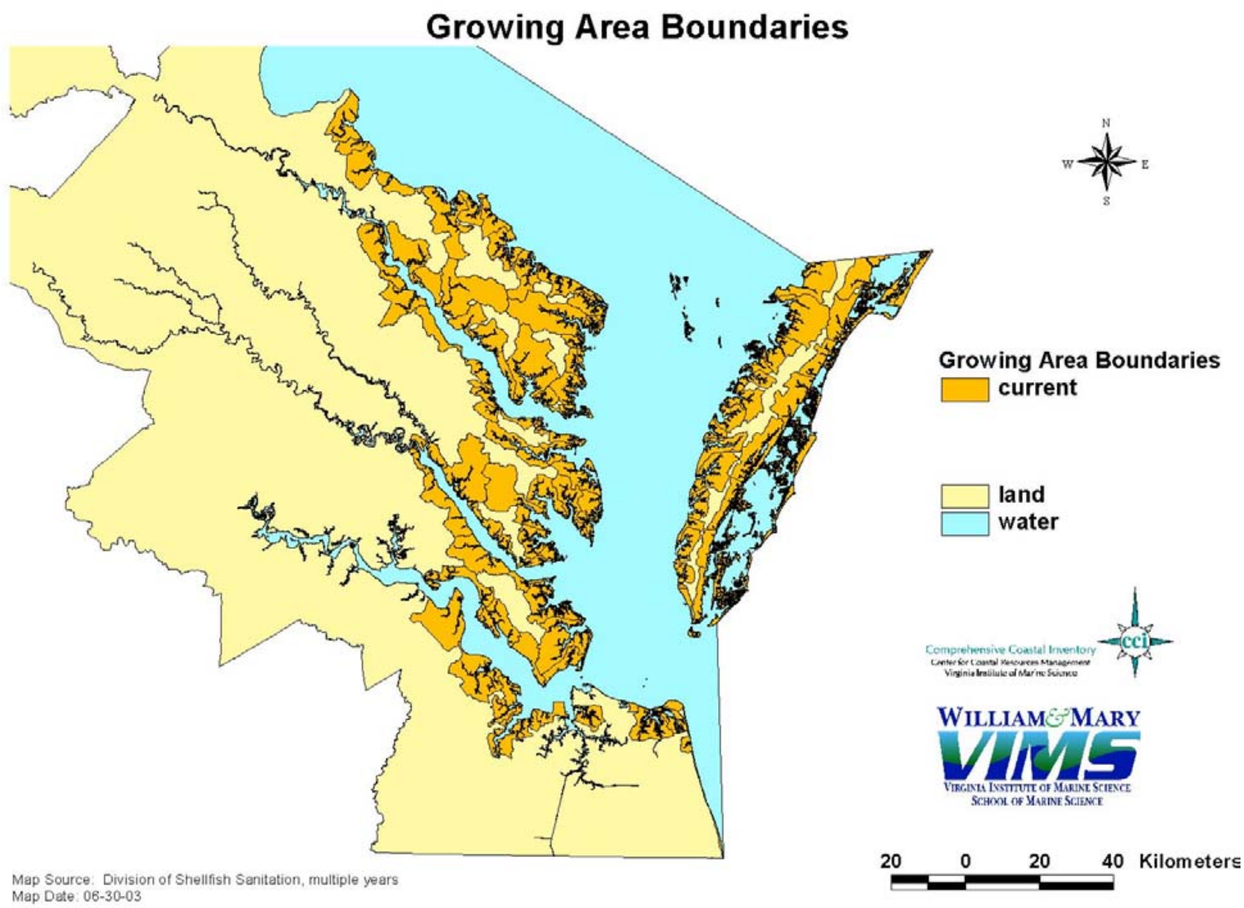


Figure 2. NLCD (National Land Cover Data; U.S. Geological Survey, 1999). The digital version is lumap.jpg on cdrom.1

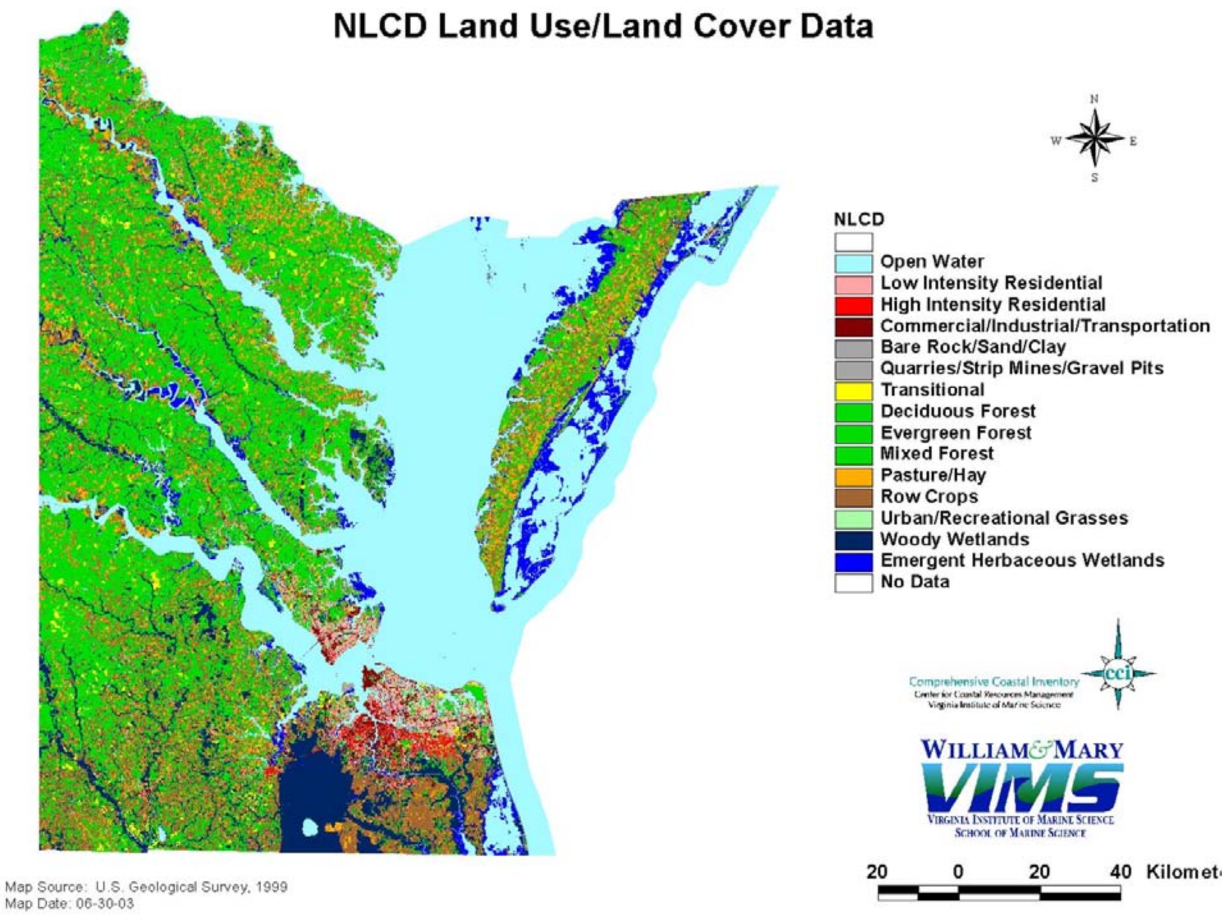


Figure 3. NED (National Elevation Data; U.S. Geological Survey, 1999). The digital version is elevmap.jpg on cdrom1.

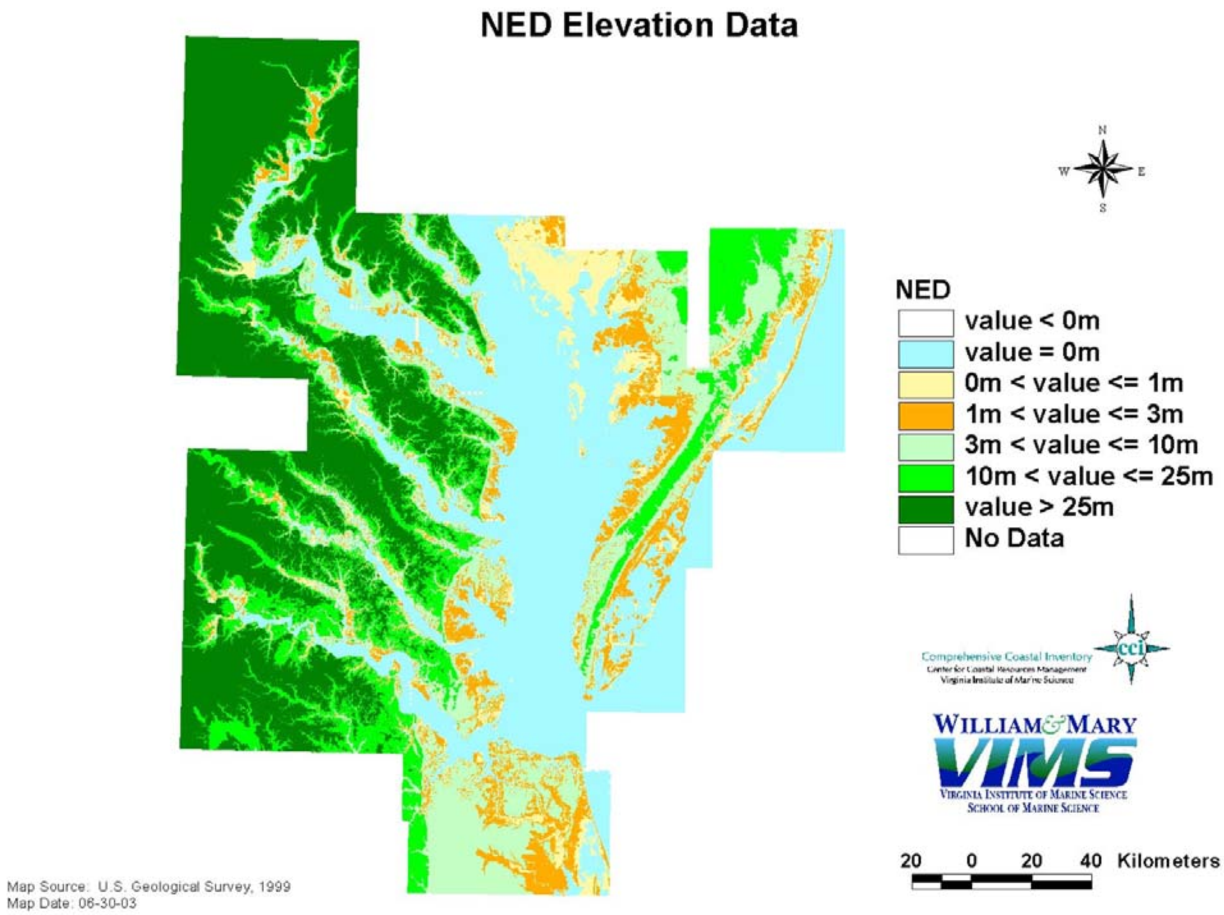


Figure 4. SSURGO and STATSGO (soils data from NRCS, 1995 to 2000). This depicts area-weighted average surface permeabilities. The digital version is soilpermmap.jpg on cdrom1.

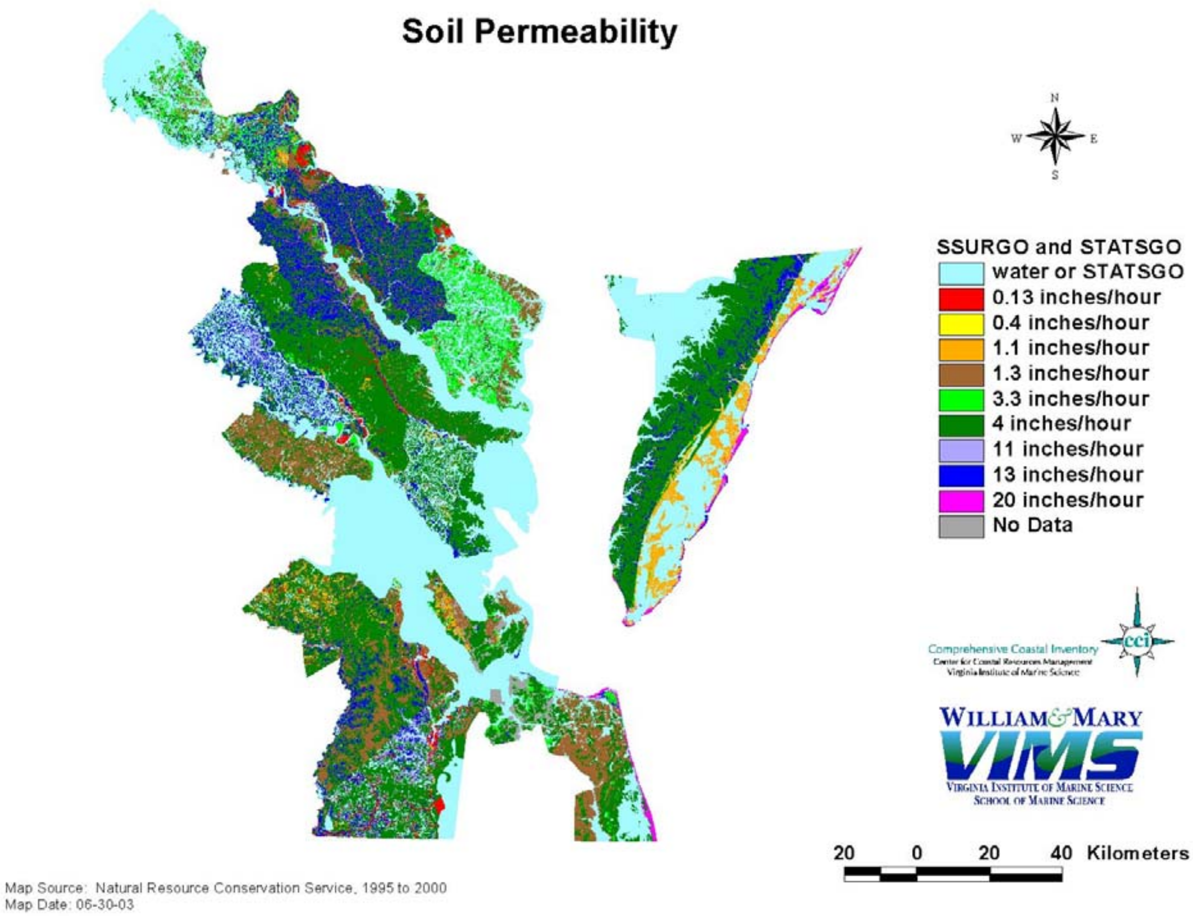


Figure 5. Population density data (US Census Bureau, 2000). The digital version is popdensmap.jpg on cdrom1.

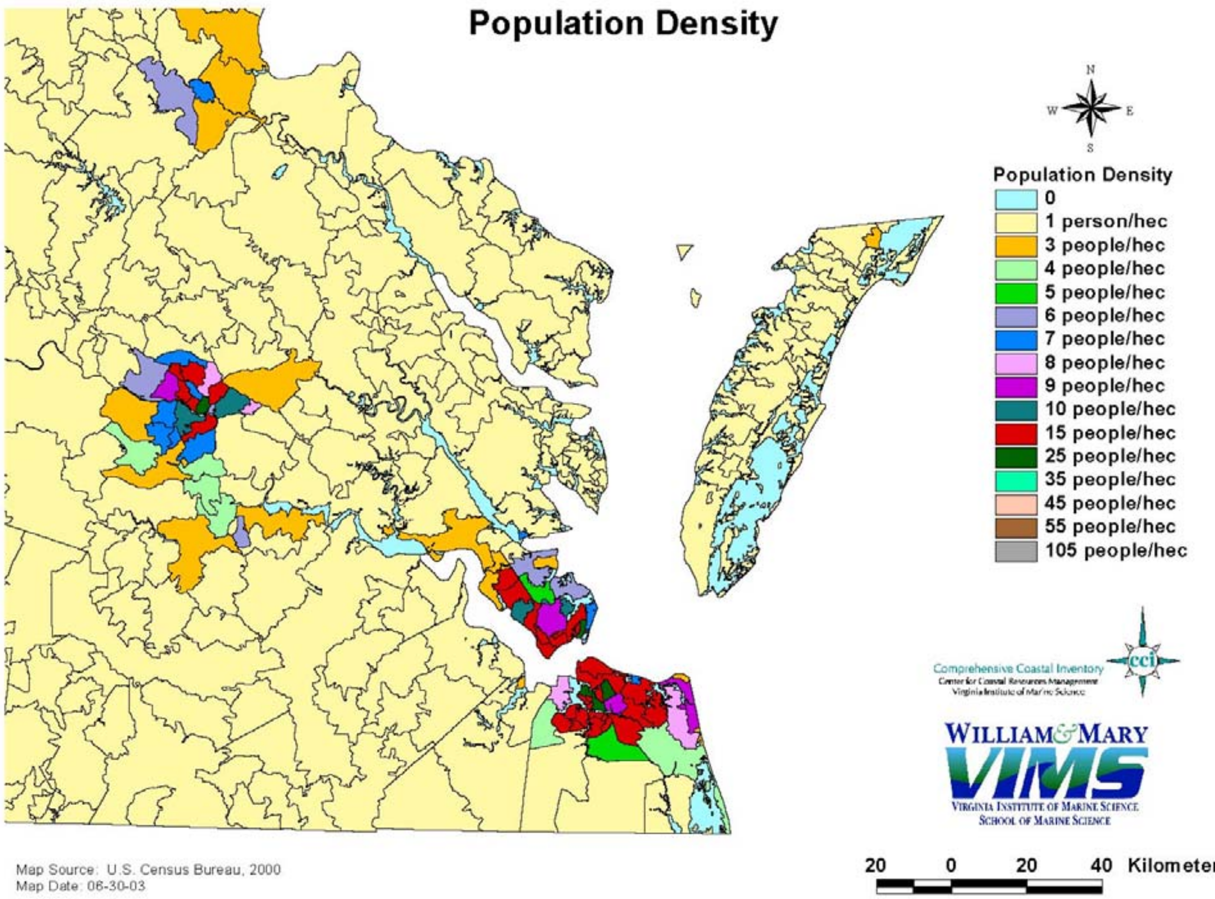


Figure 6. Bathymetry (NOS Hydrographic Survey data; multiple years)–the data were used to calculate water volumes for each growing area. The digital version is bathymdeq.jpg on cdrom1.

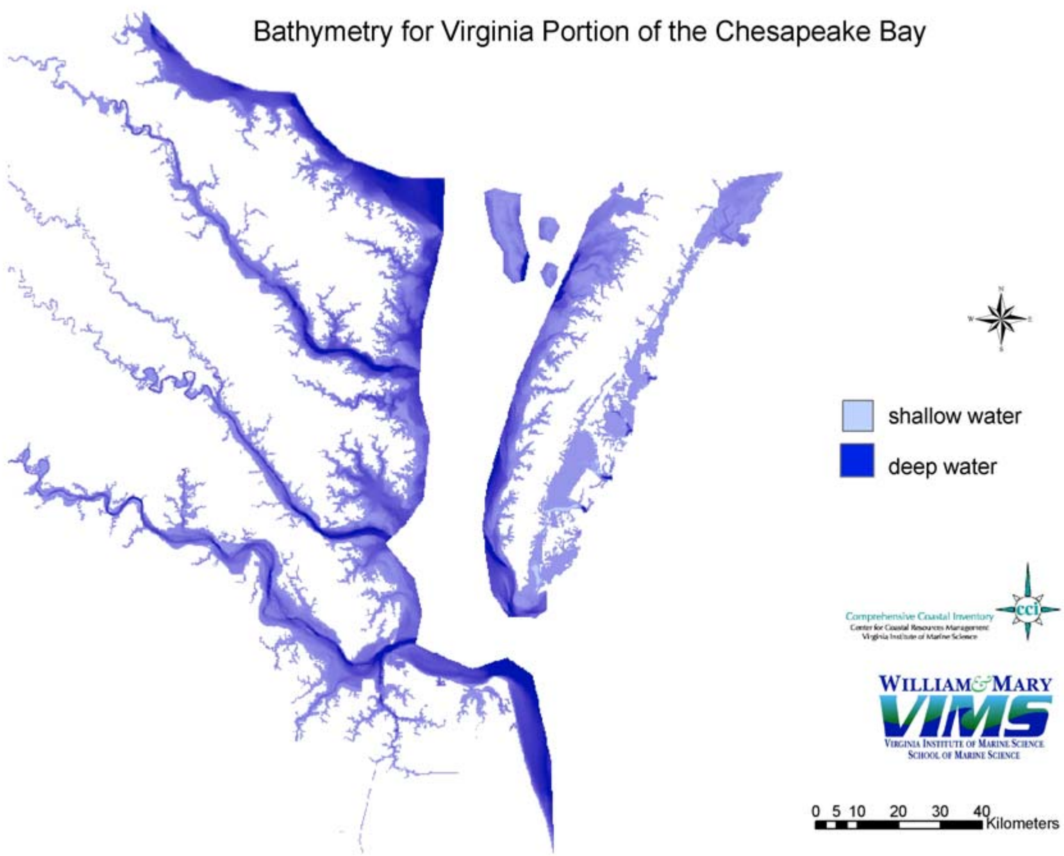


Figure 7. Condemnation zones for 1998 303d report (DSS data; multiple years). The digital version is condemnmap.jpg on cdrom 1.

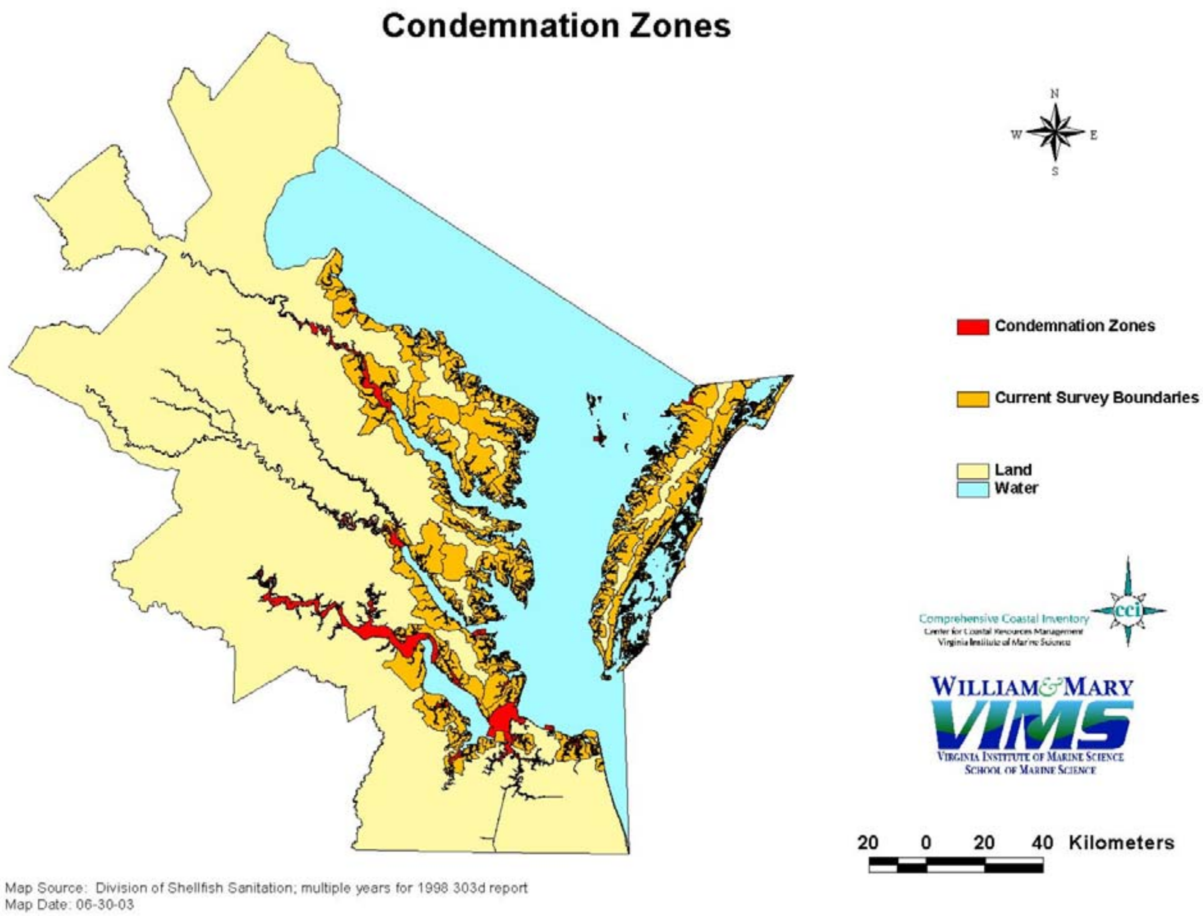
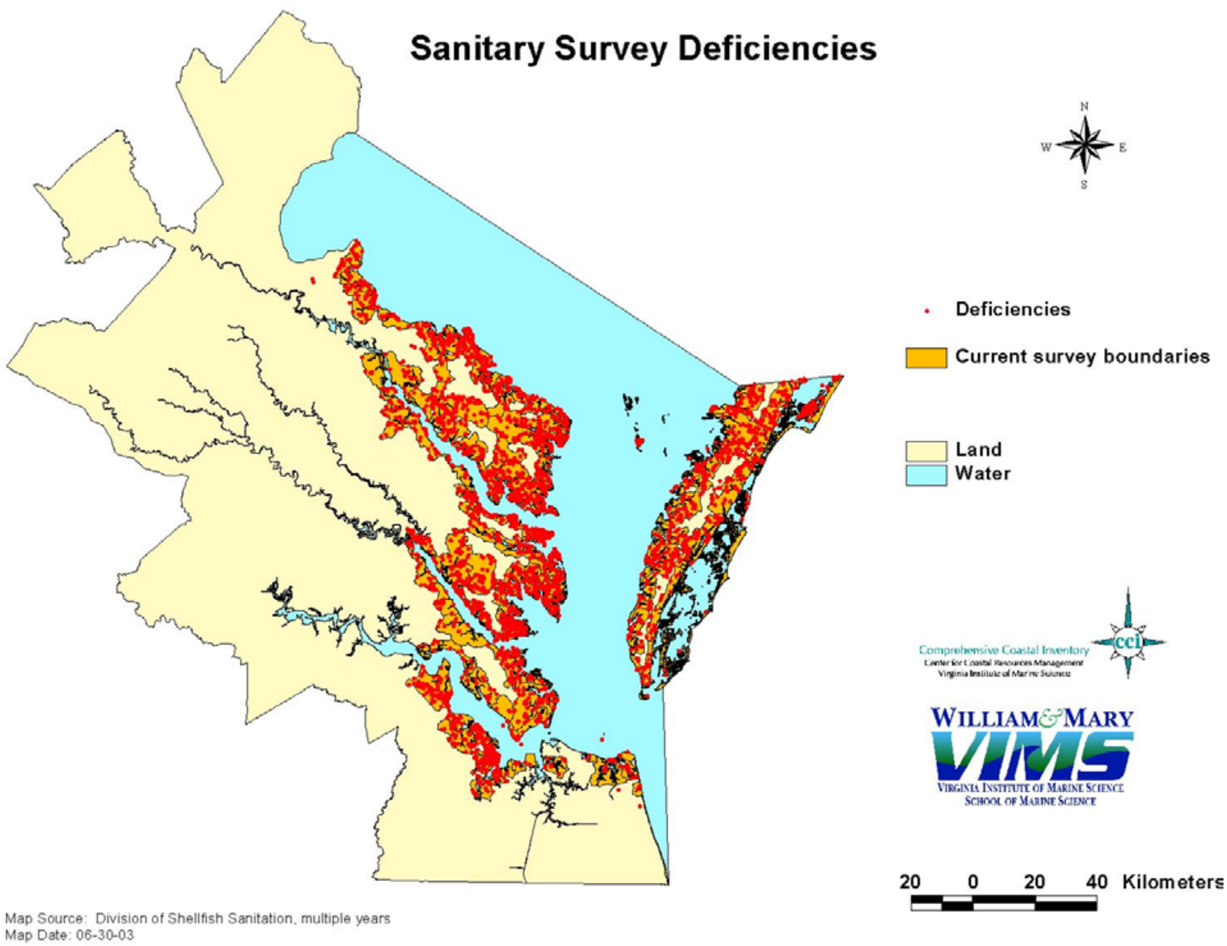


Figure 8. Sanitary Survey deficiencies (DSS data; multiple years). The digital version is deficiency.jpg on cdrom 1.



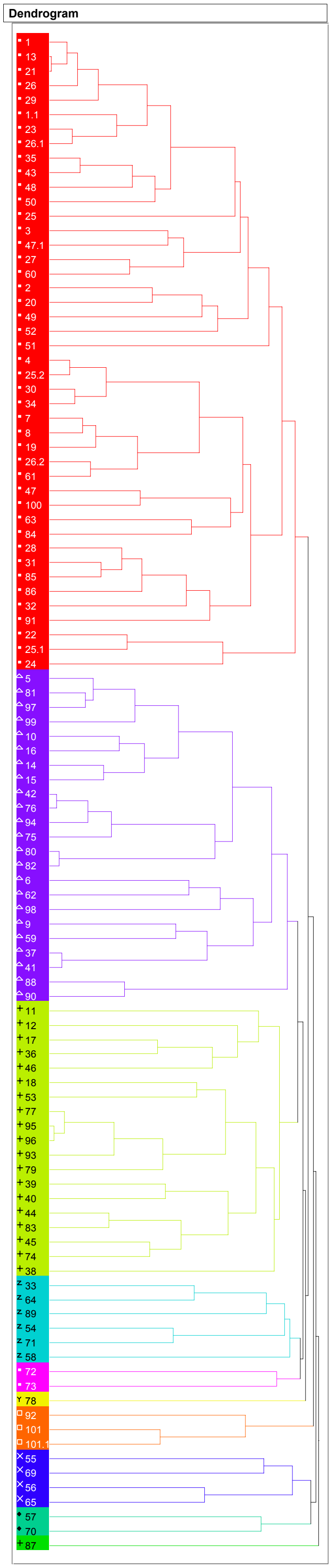


Figure 9. All parameters (developed area:total area; forested and wetland area:total area; pasture land:total area; low elevation:total area; standardized area-weighted surface permeabilities; standardized area:volume.). The digital version is alldendro.jpg on cdrom 1.

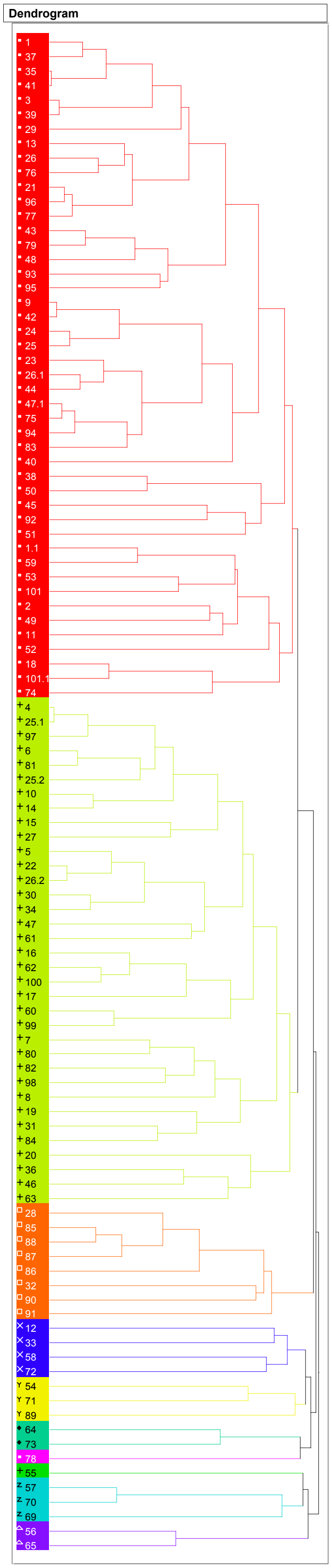


Figure 10. Source parameters (developed area:total area; forested and wetland area:total area; pasture land:total area). The digital version is sourcedendro.jpg on cdrom 1.

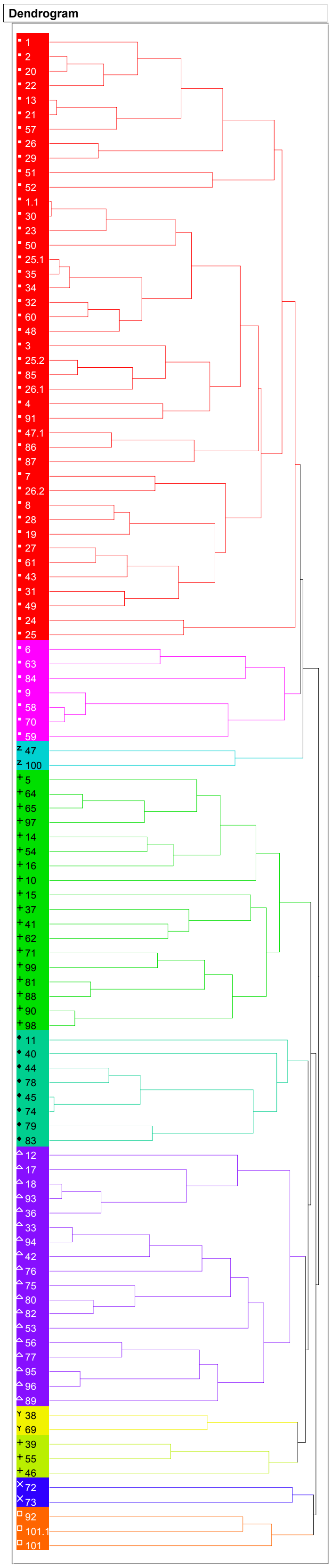


Figure 11. Delivery parameters (low elevation:total area; standardized area-weighted surface permeabilities). The digital version is deliverydendro.jpg on cdrom 1.

Figure 12. Growing areas clusters using all 6 parameters. Note: the colors on this map match the colors on the corresponding cluster diagram (except that pink on the map is red on the cluster diagram). The digital version is allmap.jpg on cdrom 1.

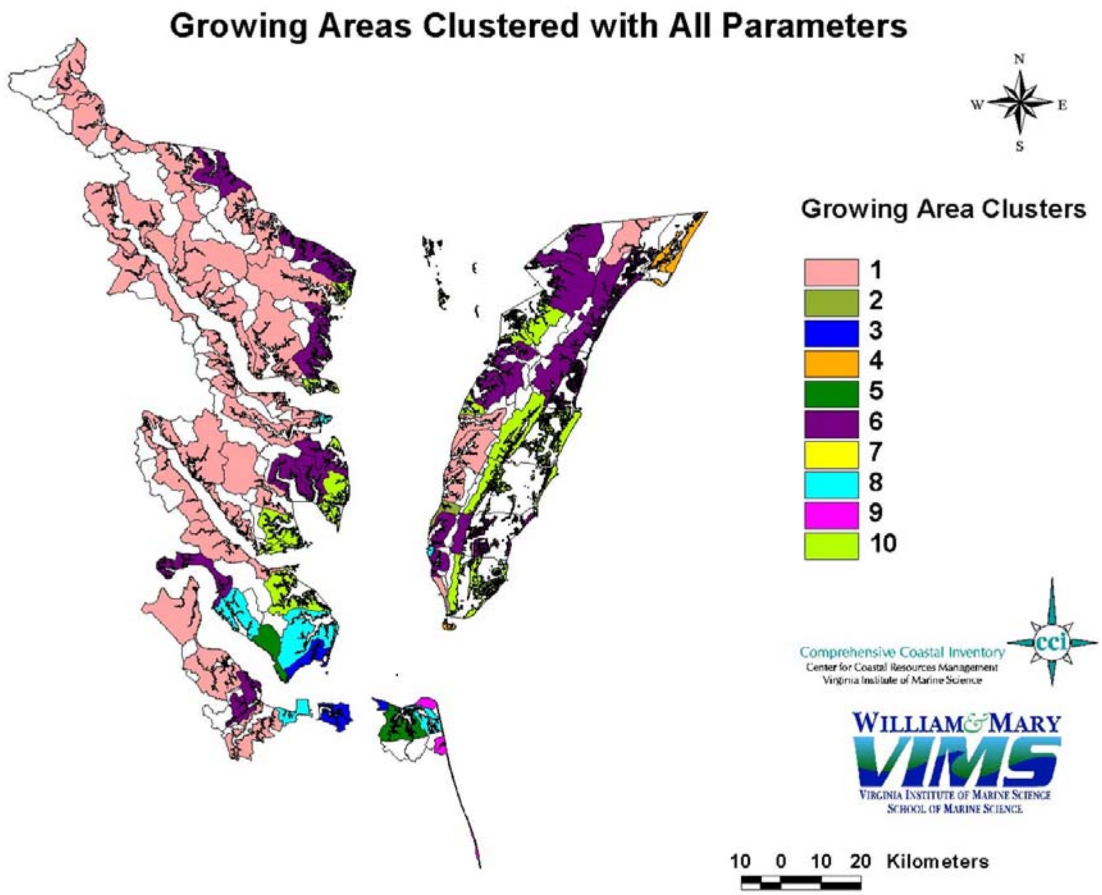


Figure 13. Growing areas clusters using 3 source parameters. Note: the colors on this map match the colors on the corresponding cluster diagram (except that pink on the map is red on the cluster diagram). The digital version is sourcemap.jpg on cdrom 1.

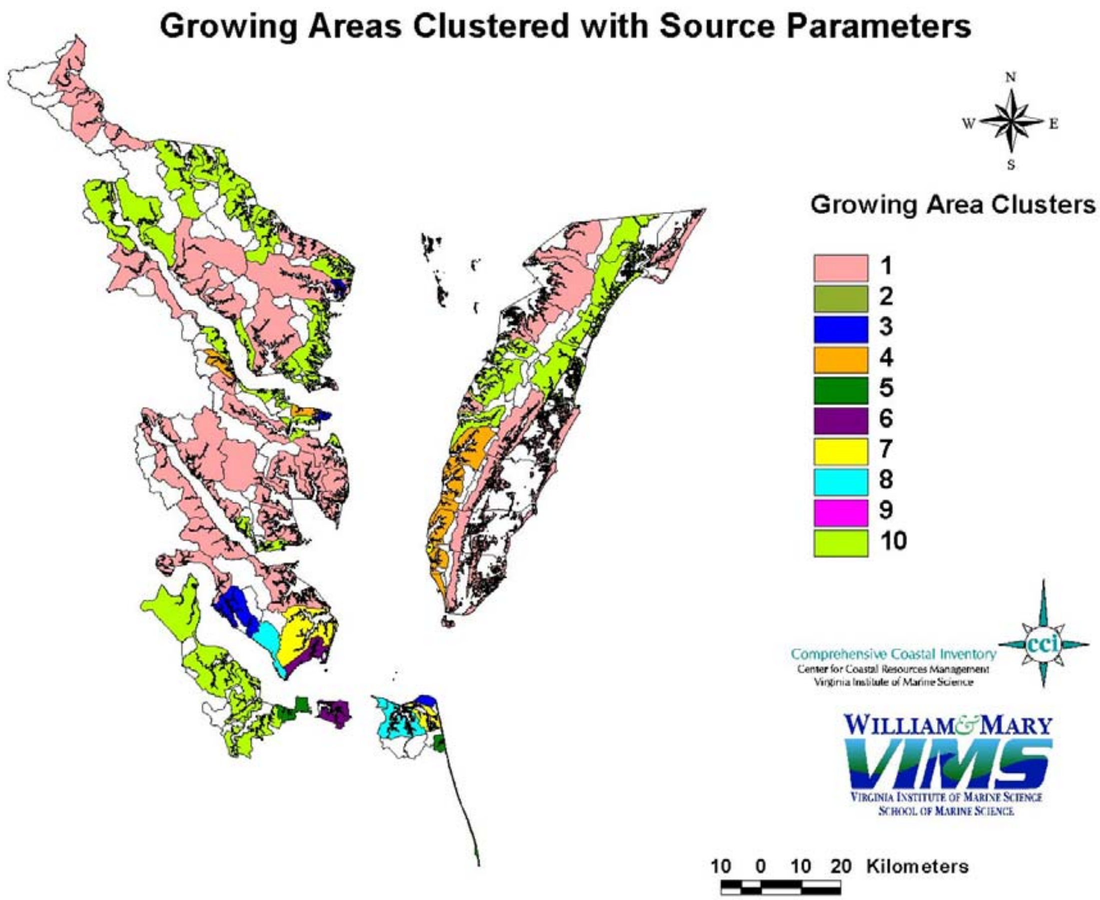
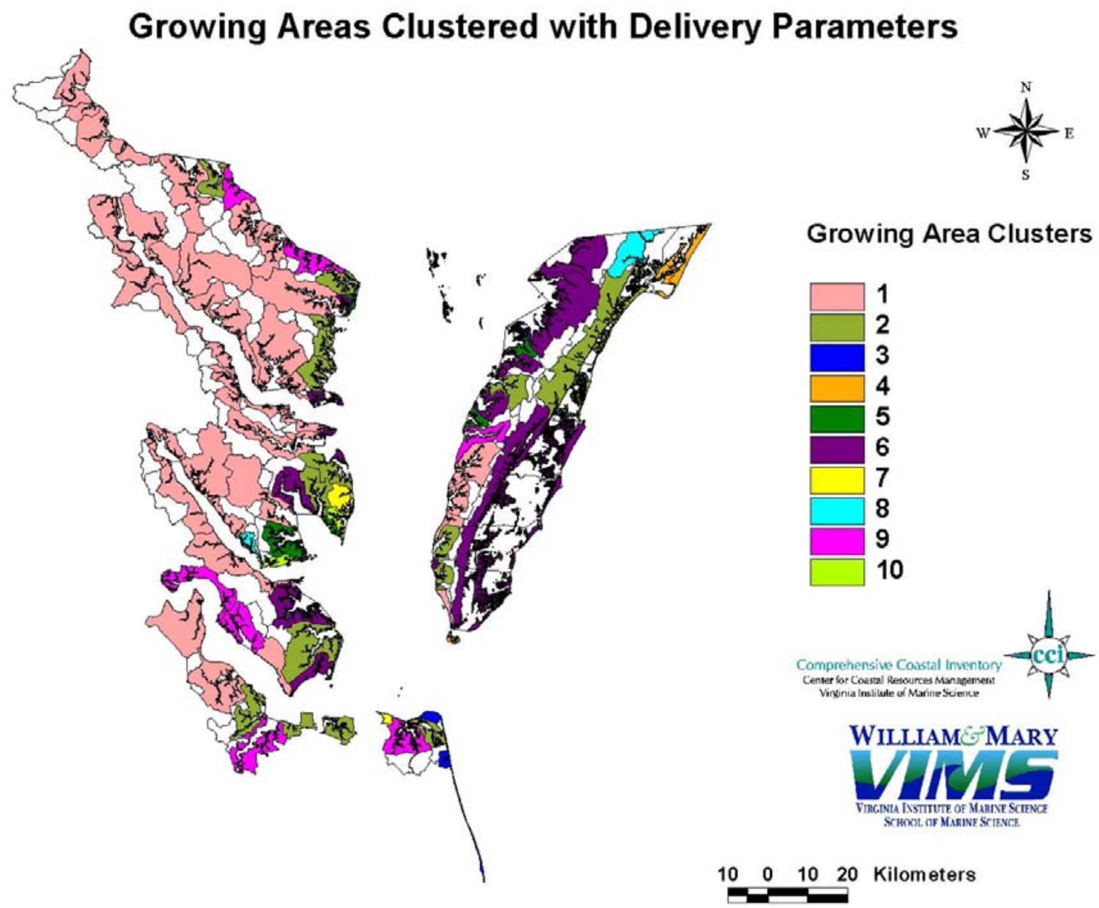


Figure 14. Growing areas clustered using 2 delivery parameters. Note: the colors on this map match the colors on the corresponding cluster diagram (except that pink on the map is red on the cluster diagram). The digital version is deliverymap.jpg on cdrom 1.



Chapter 2. DSS fecal coliform database

Fecal coliform most-probable-number (MPN) water quality data were obtained from the Virginia Department of Health, Division of Shellfish Sanitation at regular intervals over the course of this project. The most recent data set was obtained in May 2003. Given the time required to update the data files this was the last data set incorporated. The following describes how these data were manipulated to create a series of Excel files that can be used for analysis of fecal coliform contamination and potential linkage with a GIS database.

Data format and content of Excel files:

An Excel file was created for each growing area identified by numerical and alphabetical growing area name code as provided by DSS. Worksheets within each file were created as follows:

First worksheet.

Original data set as provided by DSS containing raw FC counts and ancillary field and meteorological data. Data were arrayed by sampling date.

Second worksheet.

1. Saved and copied each original DSS file into worksheet 2 within the Excel file .

2. Sorted and arranged all data by ascending date.

3. Retained column ‘Tide1Dir’, renamed it ‘’tide’

4. Eliminated all other tidal data columns

5. Retained column labeled ‘’MaxWindVel’, eliminate all other wind data

6. Created column labeled ‘Mean Temp’ as average of temperature data from stations sampled within growing area for each sampling date. Eliminated columns of temperature data collected at individual stations.

7. Created column labeled ‘Mean Sal’ as average of salinity data from stations sampled on a given date. Eliminated columns of salinity data for stations sampled. Stations showing no data were taken into consideration when average values calculated.

8. Eliminated all columns specifying stations at which sal/temp data are collected and actual values.

9. Created a column labeled ‘rain 3 days prior’ as the sum of rain in ‘amt’ columns for 24h,

date2 day, date3 day.

10. Labelled column 'week before' as 'rain 7 days prior'

11. Eliminated columns labeled Cloud Cover as well as all columns with Secchi readings.

Calculated running geometric mean and 90^{th} percentile values (90th percentile = antilog [(s)1.28 + x] where s = the standard deviation of the logarithms of the MPN values, x = the mean of the logarithms of the MPN values) from start of data set based on prior 30 data points for each sampling station listed. Missing data points or data points with "zeros" were considered as blanks for the purposes of geometric mean calculations.

Third worksheet. General descriptive statistics for each sampling station listed. Statistics calculated for all data points and the last 30 data points.

Fourth worksheet. Running geometric mean and 90^{th} percentile values calculated for last 5 years; geometric mean and 90^{th} percentile data also ranked and cumulative percentiles calculated. The latter tables can be used to evaluate the proportion of geometric mean counts which exceed the growing area standard (14 MPN/100 ml) and the 90^{th} percentile limit (43 FC MPN/100 ml).

Fifth worksheet. Chart 1 is a chart of the running geometric station means based on 30 data values from the inception of the database. Red dashed line is the growing area standard, i. e., 14 fecal coliform MPN/100 ml

Sixth worksheet. Chart 2 is a chart of the geometric mean for the last 30 data points. Red dashed line is the growing area standard,.

Seventh worksheet. Chart 3 is a chart of the 90^{th} percentile values for the last 30 data points. Red dashed line is the upper 90^{th} percentile value for a 3-tube MPN or 43 MPN/100 ml.

Eighth worksheet. Data used for charts 1, 2, and 3 are generally linked to this last worksheet.

Analysis

A variety of analyses were performed on the gross data files and GIS parameters. With a data set as large and complex as the DSS database a straightforward analysis of all data simultaneously was not useful. For the purpose of a more simplified analysis, we decided to "collapse" the fecal coliform data into an index of relative fecal coliform contamination. Accordingly, the last 30 geometric mean and 90^{th} percentile values were calculated for all stations within condemnation zones in a given growing area as shown on current water quality station location charts provided by DSS. These calculations are contained in the Excel file labeled '*Geo Mean 90^{th} Per Index.*' The first worksheet in this file (*'Avg sort across-avg-avg'*) calculates global index values for each growing area from the average geometric mean and 90^{th} percentile values for condemnation areas within the watershed. The second worksheet contains a color code that identifies station and station groupings used to calculate averages for each condemnation area in the worksheet labeled '*Data Sheet wAvgs*.' Calculated average values (shown in boxes) were then used in the first worksheet to obtain global geometric mean and 90th percentile values to assign to growing area.

Correlation of geometric mean with 90th percentile data was highly significant (Spearman's Rho = 0.89, $p<0.0001$). Correlation analysis of fecal coliform data with various GIS parameters were not statistically significant for GIS variables including: (a) ratio of area of growing area to volume of water body, (b) ratio of area of low elevation (<3m) to total area of growing area, (c) ratio of area of developed land to total area of growing area, (d) ratio of area of forest + wetlands to total area of growing area, (e) ratio of area of pasture to total area of growing area, and (f) ratio of area-weighted average surface permeability divided by 20 (to standardize). The absence of meaningful statistical relationships may indicate that watershed processes affecting fecal coliform densities in receiving waters are independent of these factors. However, GIS parameters are currently based on entire watersheds as are computer watershed and hydrographic models. Refinement of the GIS database to the level of condemnation zones may provide better relationships than when parameters reflecting the entire watershed are used.

Fecal coliform geometric mean and 90th percentile data were also clustered using a JMP (SAS) statistical software package. Results (Figures 1 and 2) show that both variables grouped as a function of index magnitude and thus may be valuable in classifying growing areas for remediation and/or specific attention. Cluster diagrams shown indicate higher values as one moves down the clusters. Using GIS the results of cluster analysis can be overlain on a map of all growing areas (Figures 3 and 4). The distribution of geometric mean and 90th percentile values can be examined to identify potential watershed characteristics responsible the observed distributions.

Successive worksheets in the '*Geo Mean 90th Per Index*' file contain geometric mean and 90th percentile data sorted numerically for all condemnation zones within each watershed. Data are sorted on both the geometric mean or the 90th percentile. Subsequent worksheets in the Excel file move consecutively from condemnation zones exhibiting the highest to those exhibiting the lowest geometric mean and 90th percentile values. Data sorted on geometric mean or 90th percentile identify the most impacted growing areas. When GIS for each condemnation area becomes available the fecal coliform data should be evaluated with respect to this more detailed database.

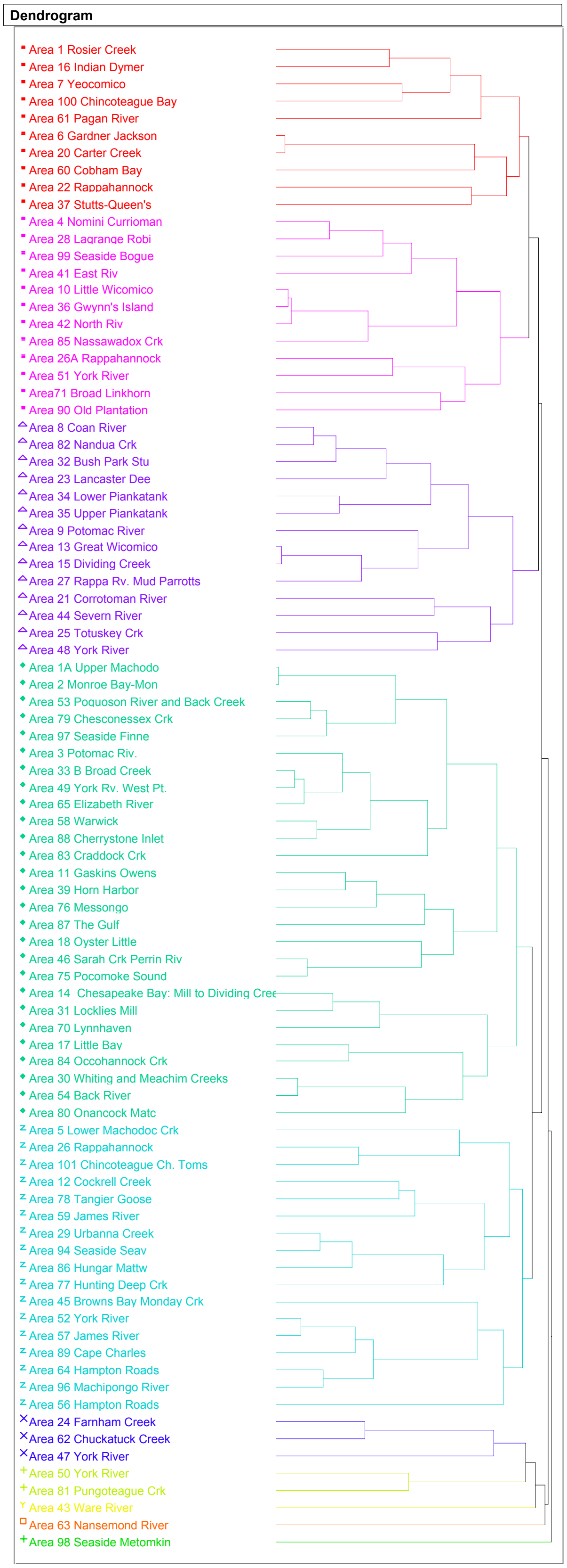


Figure 1. Fecal coliform index geometric mean values.

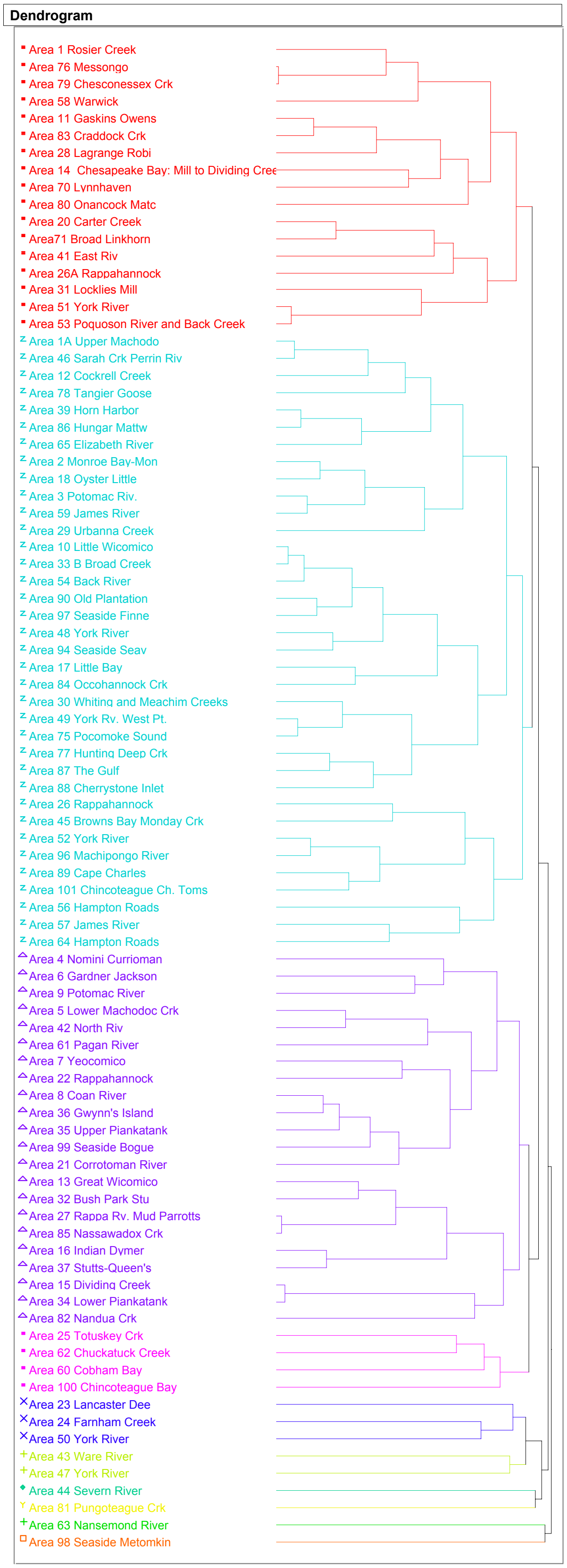


Figure 2. Fecal coliform index 90th percentile values.

Figure 3.

Geographic Distribution of Fecal Coliform Index Geometric Mean Values

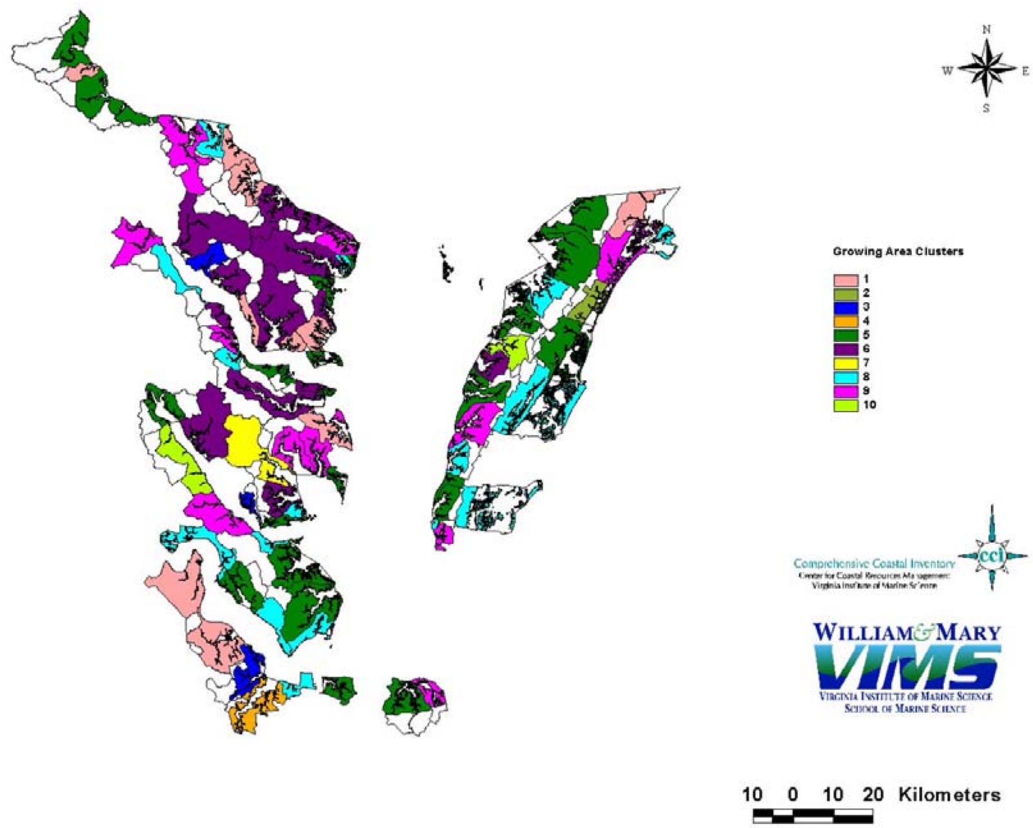
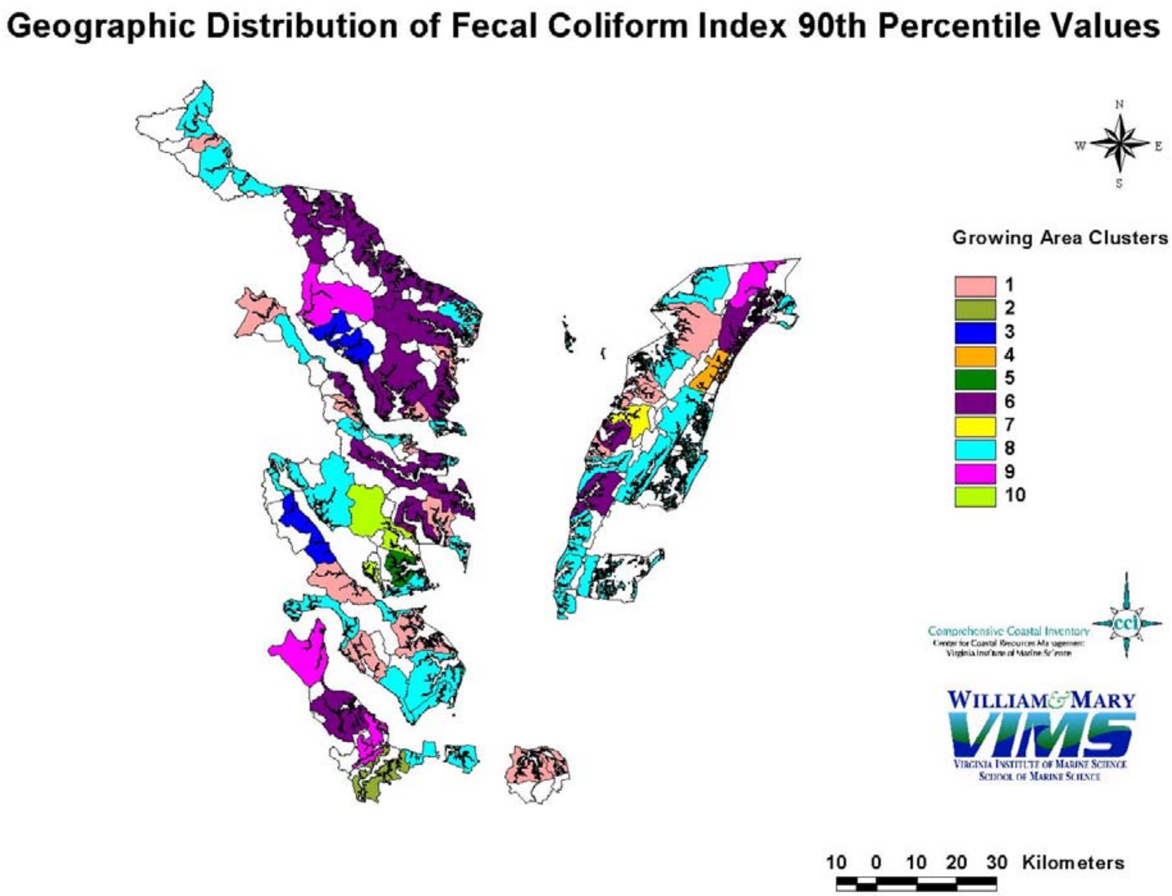


Figure 4.


Geographic Distribution of Fecal Coliform Index 90th Percentile Values
N
W
E
S
Growing Area Clusters
1
2
3
4
5
6
7
8
9
10
Comprehensive Coastal Inventory
cci
WILLIAM&MARY
VIMS
VIRGINIA INSTITUTE OF MARINE SCIENCE
SCHOOL OF MARINE SCIENCE
10 0 10 20 30 Kilometers

Appendix 1. Table containing values for final 6 parameters and the cluster numbers for the 3 analyses (clusterdata.xls on cdrom 1)

growing area number	low elevation:total area	developed land:total area	forest + wetland:total area	pasture:total area	surface permeability (standardized)	land area:water volume (standardized)	all 6 parameters (clusters)	3 source parameters (clusters)	2 delivery parameters (clusters)
1	0.076	0.018	0.734	0.064	0.183	0.016	1	1	1
1A	0.101	0.084	0.695	0.048	0.236	0.012	1	1	1
2	0.094	0.089	0.639	0.090	0.159	0.017	1	1	1
3	0.098	0.013	0.711	0.061	0.301	0.154	1	1	1
4	0.115	0.013	0.600	0.180	0.335	0.003	1	10	1
5	0.510	0.037	0.541	0.157	0.190	0.002	6	10	2
6	0.290	0.015	0.579	0.179	0.203	0.103	6	10	9
7	0.220	0.040	0.523	0.214	0.175	0.003	1	10	1
8	0.187	0.015	0.522	0.265	0.167	0.004	1	10	1
9	0.288	0.011	0.688	0.129	0.122	0.025	6	1	9
10	0.596	0.012	0.595	0.202	0.098	0.006	6	10	2
11	0.999	0.133	0.609	0.060	0.094	0.060	10	1	5
12	0.839	0.247	0.471	0.123	0.099	0.002	10	3	6
13	0.062	0.012	0.732	0.122	0.163	0.004	1	1	1
14	0.509	0.008	0.578	0.206	0.151	0.006	6	10	2
15	0.459	0.040	0.636	0.179	0.131	0.004	6	10	2
16	0.547	0.050	0.587	0.128	0.133	0.004	6	10	2
17	0.835	0.068	0.566	0.148	0.143	0.002	10	10	6
18	0.824	0.060	0.646	0.012	0.171	0.003	10	1	6
19	0.165	0.026	0.464	0.212	0.162	0.000	1	10	1
20	0.087	0.130	0.552	0.161	0.165	0.006	1	10	1
21	0.059	0.009	0.738	0.096	0.167	0.003	1	1	1
22	0.099	0.028	0.541	0.177	0.176	0.433	1	10	1
23	0.084	0.017	0.670	0.111	0.244	0.015	1	1	1
24	0.045	0.007	0.690	0.152	0.314	0.385	1	1	1
25	0.010	0.006	0.681	0.143	0.333	0.077	1	1	1
25A	0.140	0.018	0.604	0.177	0.250	0.399	1	10	1
25B	0.115	0.036	0.586	0.164	0.291	0.030	1	10	1
26	0.045	0.020	0.719	0.099	0.221	0.000	1	1	1
26A	0.127	0.029	0.673	0.102	0.268	0.049	1	1	1

26B	0.233	0.029	0.528	0.172	0.203	0.061	1	10	1
27	0.181	0.015	0.631	0.151	0.201	0.094	1	10	1
28	0.174	0.037	0.426	0.286	0.182	0.018	1	4	1
29	0.042	0.043	0.753	0.081	0.205	0.066	1	1	1
30	0.102	0.031	0.560	0.187	0.238	0.034	1	10	1
31	0.174	0.011	0.469	0.255	0.250	0.011	1	10	1
32	0.135	0.066	0.312	0.294	0.223	0.030	1	4	1
33	0.639	0.224	0.390	0.061	0.185	0.006	8	3	6
34	0.150	0.045	0.553	0.180	0.249	0.001	1	10	1
35	0.143	0.023	0.728	0.084	0.243	0.004	1	1	1
36	0.813	0.158	0.492	0.139	0.188	0.000	10	10	6
37	0.392	0.031	0.736	0.069	0.221	0.012	6	1	2
38	0.740	0.002	0.847	0.024	0.393	0.051	10	1	7
39	0.893	0.006	0.716	0.060	0.411	0.005	10	1	10
40	1.000	0.012	0.677	0.058	0.343	0.018	10	1	5
41	0.425	0.022	0.731	0.083	0.194	0.004	6	1	2
42	0.619	0.008	0.694	0.124	0.197	0.001	6	1	6
43	0.165	0.015	0.779	0.080	0.212	0.004	1	1	1
44	0.984	0.025	0.664	0.092	0.229	0.002	10	1	5
45	1.000	0.049	0.771	0.026	0.209	0.012	10	1	5
46	0.847	0.144	0.511	0.105	0.317	0.017	10	10	10
47	0.295	0.017	0.516	0.149	0.398	0.045	1	10	8
47A	0.209	0.003	0.655	0.107	0.318	0.170	1	1	1
48	0.122	0.002	0.799	0.078	0.243	0.094	1	1	1
49	0.191	0.138	0.662	0.084	0.236	0.001	1	1	1
50	0.114	0.010	0.859	0.050	0.208	0.000	1	1	1
51	0.110	0.100	0.781	0.017	0.073	0.170	1	1	1
52	0.164	0.200	0.622	0.075	0.087	0.088	1	1	1
53	0.687	0.133	0.665	0.036	0.148	0.003	10	1	6
54	0.535	0.487	0.391	0.030	0.161	0.008	8	7	2
55	0.887	0.541	0.006	0.000	0.446	0.000	3	2	10
56	0.720	0.894	0.074	0.000	0.214	0.027	3	6	6
57	0.042	0.672	0.258	0.005	0.158	0.000	5	8	1
58	0.303	0.311	0.528	0.036	0.118	0.013	8	3	9
59	0.350	0.104	0.675	0.048	0.081	0.021	6	1	9

60	0.122	0.015	0.623	0.099	0.228	0.169	1	10	1
61	0.190	0.056	0.526	0.136	0.214	0.026	1	10	1
62	0.394	0.046	0.582	0.103	0.180	0.021	6	10	2
63	0.279	0.113	0.461	0.142	0.233	0.003	1	10	9
64	0.548	0.200	0.263	0.043	0.193	0.000	8	5	2
65	0.561	0.865	0.102	0.000	0.192	0.012	3	6	2
69	0.690	0.601	0.147	0.000	0.371	0.000	3	8	7
70	0.295	0.650	0.258	0.000	0.115	0.015	5	8	9
71	0.505	0.554	0.370	0.000	0.274	0.003	8	7	2
72	0.538	0.217	0.532	0.000	0.685	0.000	9	3	3
73	0.558	0.238	0.287	0.006	0.520	0.022	9	5	3
74	1.000	0.001	0.665	0.000	0.211	0.000	10	1	5
75	0.692	0.011	0.652	0.103	0.244	0.005	6	1	6
76	0.615	0.010	0.714	0.114	0.231	0.005	6	1	6
77	0.736	0.011	0.737	0.106	0.228	0.002	10	1	6
78	1.000	0.008	0.202	0.004	0.237	0.000	7	9	5
79	0.892	0.021	0.786	0.094	0.207	0.002	10	1	5
80	0.670	0.041	0.549	0.227	0.245	0.005	6	10	6
81	0.497	0.011	0.569	0.168	0.249	0.013	6	10	2
82	0.669	0.011	0.547	0.202	0.229	0.005	6	10	6
83	0.902	0.003	0.647	0.122	0.235	0.011	10	1	5
84	0.325	0.013	0.494	0.237	0.288	0.008	1	10	9
85	0.128	0.015	0.415	0.276	0.292	0.016	1	4	1
86	0.216	0.004	0.398	0.317	0.299	0.011	1	4	1
87	0.182	0.024	0.395	0.291	0.280	1.000	2	4	1
88	0.488	0.017	0.398	0.274	0.236	0.007	6	4	2
89	0.744	0.369	0.356	0.146	0.269	0.004	8	7	6
90	0.438	0.008	0.340	0.341	0.258	0.044	6	4	2
91	0.083	0.019	0.329	0.225	0.334	0.000	1	4	1
92	0.999	0.008	0.789	0.000	0.563	0.000	4	1	4
93	0.817	0.006	0.777	0.050	0.176	0.003	10	1	6
94	0.650	0.004	0.662	0.095	0.184	0.001	6	1	6
95	0.765	0.002	0.760	0.077	0.213	0.002	10	1	6
96	0.778	0.013	0.730	0.092	0.211	0.005	10	1	6
97	0.546	0.007	0.603	0.164	0.218	0.010	6	10	2

98	0.431	0.001	0.544	0.236	0.267	0.120	6	10	2
99	0.498	0.012	0.606	0.114	0.304	0.005	6	10	2
100	0.224	0.060	0.575	0.113	0.378	0.002	1	10	8
101	0.981	0.141	0.636	0.009	0.718	0.001	4	1	4
101A	0.972	0.044	0.655	0.004	0.637	0.001	4	1	4