

Importance of Data Mining in the Assessment of Anoxic Problem from Estuarine Monitoring Data

Ping Wang, Richard A. Batiuk, Lewis C. Linker, and Mary Ellen Ley

Abstract--A data retrieval procedure is used to process observed dissolved oxygen (DO) concentrations in the Chesapeake Bay to obtain information on the intensity or degree of anoxia, and to explore optimal DO-thresholds to define the degree of anoxia. The method consists of a series of processes, such as data interpolation, pattern mining, feature transformation, uncertainty management, and verification. A new metric, "anoxic intensity", is an outcome from the data mining, which addresses the degrees of anoxia and improves over the traditional way using "anoxic volume" as the key metric of anoxia. The data retrieval method described overcomes the current methodology shortfalls that may lead to data misinterpretation. This paper emphasizes the importance of data mining for data utilization.

I. INTRODUCTION

Anoxia is the absence of or an extremely low supply of oxygen. In the context of the Chesapeake Bay water quality it describes bottom waters with dissolved oxygen levels damaging or fatal to living resources. This paper describes data mining of Chesapeake Bay dissolved oxygen (DO) monitoring data for use in water quality management.

In a large stratified eutrophic water body such as the Chesapeake, anoxic water is usually found in deep waters, especially in the summer, due to the oxygen consuming decay of organic materials in bottom waters and sediments, and the lack of reaeration or oxygen exchange with the upper water layers caused by temperature and salinity induced density stratification. Anoxic volume, i.e., the volume of water that has oxygen concentrations lower than a specified DO-threshold, has been used as a metric of the degree of

anoxia. Different DO-thresholds have been used to define anoxia for different purposes, for example, concentrations of $DO=0.0$ mg/l, or <0.07 mg/l [1], or ≤ 0.2 mg/l [2], or <1 mg/l have been used to define anoxia [3], [4]. The selection of different anoxia thresholds may be due to the consideration of minimum DO required for various aquatic species on certain life stages.

The anoxic volume is an appropriate DO metric when we want to know the volume of water unavailable to certain usages. However, anoxic volume may inadequately represent the relative intensity of anoxia among years [5]. This paper discusses, through data mining of observed DO data from the Chesapeake Bay, how to calculate a more appropriate DO metric, the anoxic intensity, that can adequately represent relative DO problems.

II. METHODS

The data utilization in this work involves a series of processes, including data interpolation, downsizing/grouping, ranking, feature transformation, pattern mining, uncertainty management, and verification. The core of this work is to analyze features of anoxia and apply feature transformation to better represent anoxia problems.

A. Basic data preparation

There are more than 50 long-term monitoring stations in the mainstem of the Chesapeake Bay (Fig. 1), and each has been sampled 1-2 times each month since 1985. The mainstem Bay has a surface area of 7,000 km² with an average depth of 7-meters and a maximum depth of 40 meters. The depth interval of sampling is usually 1 meter, but is sometimes 2 meters.

Six sampling cruises in the summer (mid-June to mid-September) were selected to assess annual summer anoxia. The observed DO values from discrete monitoring stations in the mainstem Bay were interpolated over the whole volume of the mainstem Bay. The year 2001 version of the three-dimensional Chesapeake Bay Volumetric Interpolator [6] grid is adopted, which is generally 1 km long by 1 km wide by 1 meter deep in the mainstem but varies substantially along the shoreline and in the tributaries.

The DO in the cells without observed data is

P. Wang is with the Chesapeake Bay Office, University of Maryland Center for Environmental Science, Annapolis, MD 21403 USA (phone: 410-267-5744; fax: 410-267-5777; e-mail: pwang@chesapeakebay.net).

R.A. Batiuk is with the Chesapeake Bay Program, USEPA, Annapolis, MD 21403 USA (phone: 410-267-5731, e-mail: batiuk.richard@epa.gov).

L.C. Linker is with the Chesapeake Bay Program, USEPA, Annapolis, MD 21403 USA (phone: 410-267-5744, e-mail: linker.lewis@epa.gov).

M.E. Ley is with the Chesapeake Bay Office, USGS, Annapolis, MD 21403 USA (phone: 410-267-5750, e-mail: mley@usgs.gov).

interpolated from observed DO at the similar depth from three nearest stations that are weighed by the inverse of the squared distance. The variation of DO concentrations in the Bay is insignificant horizontally but significant vertically, because dissolved oxygen in the stratified Bay is highly dependent on the “oxygen source” of surface reaeration and the “oxygen sink” of bottom sediment oxygen demand. The vertical sampling interval is close to the height of grid cells. Such interpolated DO concentrations were tested to be acceptable [5], [6]. The interpolated values have one more decimal point than the measured values to avoid false flattening of the values.

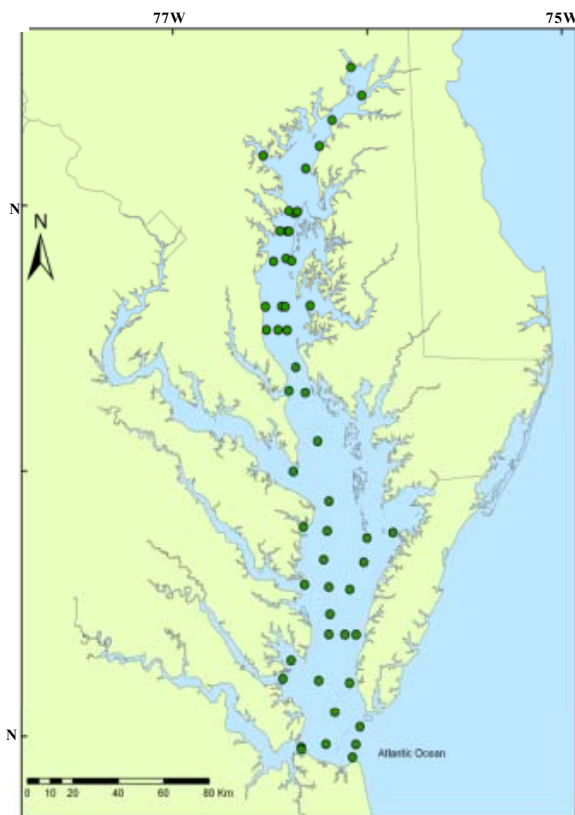


Fig. 1. The Chesapeake Bay and its adjacent ocean. The dots are long term monitoring stations in the mainstem Bay, which were used to interpolate DO concentrations.

This work uses the improved procedures in assigning the observed DO concentrations to grid cells and handling missing values [5], as described in the following.

The location of vertical cast of the dissolved oxygen probe at a given fixed monitoring station is determined by GPS instrumentation, however, the depth to bottom often varies by two or three meters because of wind and tidal effects or a drift of the DO probe relative to the

bed. This situation results in a mismatch between the observed depth for a sampling cruise and the interpolator grid depth. If the mismatch is less than 2 meters, the mismatch is considered due to different stages of tide, bottom bathymetry, boat location, freshwater discharge, winds, and the following methods are used to align the observations with the interpolator grid. If the observed total depth is greater than the interpolator grid column, the extra depths over 2 meters in the observation are ignored. The rest of the observation profile is adjusted to match the interpolator grid column, and DO in the observation profile is recalculated for the depths matching each cell center in the grid column. The cells are assigned with DO values accordingly.

If the observed total depth is shorter than the interpolator grid column, then the observation profile is stretched up to two meters longer. Dissolved oxygen in the observation profile are recalculated then assigned to the grid cells of corresponding depths. Any remaining deep cells in the interpolator grid column are assigned the value of “missing”.

In examining whether a cell meets a certain DO-threshold, for the cells having a “missing” value, we first look the deepest cell in a grid column that has a DO value (non-missing) and use a specified DO-threshold to check whether it is within the threshold of anoxia, we then assign the same condition to all cells underneath.

B. Features and feature selection for anoxia

Anoxic volume (AV). Anoxic volume is a common metric in water quality management. It is determined by adding up the volume (V) of water from each cell i that meets the threshold of anoxia:

$$AV = \sum V_i \quad (1)$$

Anoxic volume is a useful metric indicating the volume of water that violates a DO criterion. For the anoxia defined as low DO but non-zero value, a water body with a larger anoxic volume may not have a greater degree of anoxia than that of another with a smaller anoxic volume [5]. An example is given here, using hypothetical but realistic observations from the Chesapeake Bay on two separate cruises. In one, cruise A, the lower layer of water in a square kilometer was observed to be 5-meters deep and containing water of $DO=0.1$ mg/l. Above this layer $DO > 0.2$ mg/l. In the second cruise, B, the lower layer of water in a square kilometer was 7-meters deep, containing water with $DO=0.2$ mg/l. The water above 7 meters has $DO > 0.2$ mg/l. Using a threshold of $DO \leq 0.2$ mg/l, the anoxic volumes in waters A and B are 0.005 km^3 and 0.007 km^3 , respectively. Although the bottom water of A has a higher degree of anoxia than water B, it has

lower anoxic volume.

Anoxic intensity (AI). To overcome the above problem, a transformed feature of anoxia volume, the anoxic intensity (AI), is introduced. In the AI calculation, the volume of water is normalized (i.e., divided) by DO concentration for each cell i before the summation:

$$AI = m \sum V_i / DO_i \quad (2)$$

where, $m = 0.5$ DO-threshold (in mg/l).

The above calculation is based on the assumption, that, for example, if $DO=0.2$ mg/l is a problem for a parcel of water, then the water of a same volume but containing $DO=0.1$ mg/l is twice the degree of the problem. The multiplier, m serves for two purposes. One is to make the unit of AI to a unit of volume. The other is to provide comparability between values of AI and AV with the same DO-threshold: $AI = AV$ if the average DO in the anoxic water equals half the threshold; $AI > AV$ if the average DO is lower than half the threshold; $AI < AV$ if the average DO is higher than half the threshold. The unit of AI is “ km^3 equivalence for $DO = 0.5$ threshold”, briefly, km^3 -eqv.

In the above example, AI of $DO \leq 0.2$ mg/l are $0.005 km^3$ -eqv and $0.035 km^3$ -eqv for waters A and B, respectively. AI correctly represents the relative intensity of anoxia between waters A and B. Whereas, in the AV calculation, once DO concentrations in parcels of water meet the threshold, it considers the volumes only, without comparing the relative DO concentrations. Note that the anoxic volume of water B, having higher anoxic volume but lower anoxic intensity, can be generated from the water of Case A through partial mixing with the higher DO water in the upper layer.

C. Verification using nutrient based information

Summer anoxia is related to the decay of algae due to excessive nutrient inputs. High summer temperatures accelerate algal decay. Fresh water flow promotes Chesapeake stratification and is also related to the magnitude of nutrient loads delivered to the Bay [7], [5]. The spring nutrient loads (NL), spring flows (FLW), and summer water temperature (WTMP) are the three most important factors to summer anoxia. The Pearson product-moment correlation coefficients between anoxic intensity and NL, FLW and WTMP are about 0.85, 0.85, and 0.2, respectively [5]. Here, the NL uses the effective nutrient loads to the Bay’s anoxic center from the Chesapeake Bay watershed; WTMP uses average summer temperature in the greater Bay’s anoxic center; FLW uses the flow from the Susquehanna River, which contributes about 50% of the

freshwater to the Bay.

We use a simple formula to calculate an index for annual summer anoxia, called the nutrient-based anoxic index (NBAI):

$$NBAI = \{a(NL/average_NL) + b(FLW/average_FLW) + c(WTMP/average_WTMP)\} / (a+b+c) \quad (3)$$

where, NL, FLW and WTMP are annual values, and the average is based on twenty-one years (1985-2005) of data.

If NL, FLW, and WTMP from a year all equal their corresponding twenty-one years’ averages, the NBAI equals 1. The year with NBAI greater than 1 probably will have more anoxia than on average.

If the metric of AV or AI is representative of the anoxic conditions, the order of the annual AV or AI will be similar to the order of annual NBAI, especially for the years of greatest anoxia. We can use this to test which metric best represents anoxia.

III. RESULTS

A. Anoxic volume

Figure 2 plots anoxic volumes (AV) for five low DO-thresholds in a stacked column plot. Different thresholds yield different rankings of anoxic volumes among years. For example, the four years of the highest anoxic volumes defined by $DO \leq 0.3$ mg/l are, in the descending order, 2005, 1998, 1993, and 1996. Whereas, the ranking is 1993, 1998, 2005, and 1996 for anoxia defined by $DO \leq 0.2$ mg/l, and is 1993, 1996, 1998, and 1994 for anoxia defined by $DO < 0.2$ mg/l. Greater differences exist in the ranking of anoxic volumes for anoxia defined by higher DO-thresholds (not plotted). Therefore, it is difficult to select a DO-threshold upon which the calculated anoxic volumes can represent relative DO problems among years.

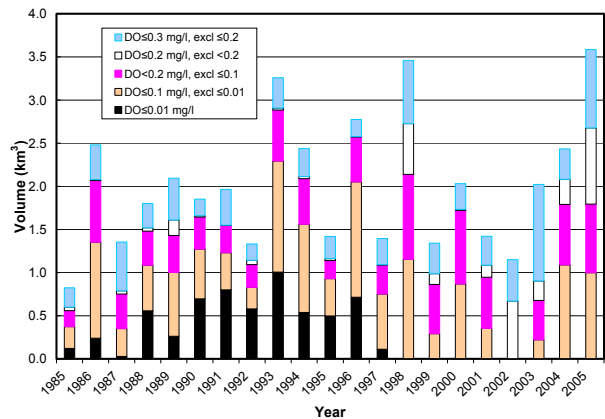


Fig. 2. Stacked columns for anoxic volume (AV) in summer of 1985-2005 for five low DO thresholds.

B. Anoxic Intensity

Figure 3 plots annual summer anoxic intensity (AI) for six thresholds. Note that in the AI calculation (Equation 2), for DO concentrations between 0-0.07 mg/l, we use a factor of 15 to multiply the volume, instead of dividing concentrations. The ranking orders of annual AI are almost the same among the six thresholds, especially for the 5 years of greatest anoxia, 1993, 1996, 1994, 1998, and 2005. These findings suggest that an AI, by normalization of AV using DO concentration, best represents the extent of DO problems in the Chesapeake.

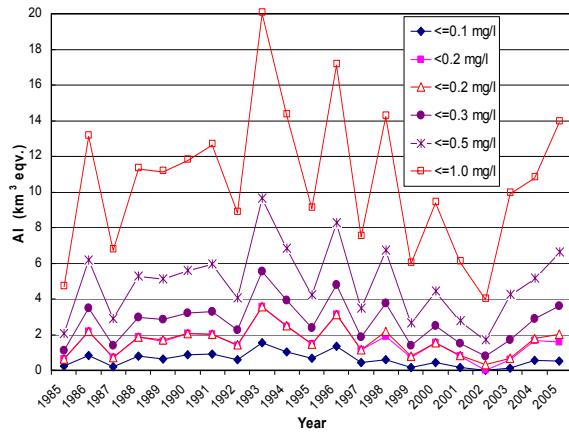


Fig. 3. Anoxic intensity (AI) of six thresholds for summer 1985-2005. Using value 15 as weighting factor for $\text{DO} \leq 0.07$ mg/l.

The AI defined by any of these DO-thresholds can be used to compare relative DO problems among years. However, using DO-thresholds higher than 1 mg/l to define AI may involve large volumes of water that are farther away from the center of anoxia (that is located in the deepest region of the Chesapeake) and may complicate the situation. To avoid this, it is better to use DO-thresholds lower than 1 mg/l for the AI calculation. On the other hand, using a low DO-threshold, such as ≤ 0.1 mg/l or lower, to define AI, omits the water containing slightly higher DO (e.g., 0.2 mg/l) that is unhealthy to most aquatic species. To avoid this, it is better to use DO-thresholds greater than 0.1 mg/l. Thus, the threshold of $\text{DO} \leq 0.2$ mg/l or ≤ 0.3 mg/l is appropriate for AI calculation. The metric of AI is used to indicate relative DO problems among years. Although the assessment of AI is not necessarily confined by DO requirements of living resources, we still like to refer the DO requirements by living resources in the selection of DO threshold for AI calculation. Considering the dissolved oxygen criteria

set by the USEPA Chesapeake Bay Program [9], we are inclined to use the threshold of $\text{DO} \leq 0.2$ mg/l for calculating AI in the determination of the best and worst years of anoxia in the Chesapeake Bay.

C. Nutrient-based anoxic index

Column 2 of Table 1 lists the estimated nutrient-based anoxic index (NBAI) for 1985-2005. Considering the key role of nutrients on anoxia, the weighting factors for LI, FLW, and WTMP in Equation 3 are: $a=10$, $b=5$, and $c=1$, respectively (Note: NBAI of 2005 is less accurate than other years, because it was calculated from a provisional database of nutrient loads and water temperature).

Years 1993, 1994, 1996, 1998, and 2005 have significantly higher NBAI than other years, consistent with the assessment of relative anoxic problems by AI of $\text{DO} \leq 0.2$ mg/l, confirming that AI of $\text{DO} \leq 0.2$ mg/l is a suitable index for anoxia.

We tested with a rather extreme case: increasing the weights on flow and temperature, and using a factor of 1 to weigh NL, FLW, and WTMP. The ranks of the top five NBAI (Columns 4 and 5, Table 1) are close to those calculated by weighting NL, FLW, and WTMP at 10, 5, and 1 (Columns 2 and 3, Table 1). This increases our confidence in the above analysis.

Table 1. Nutrient-based anoxic index (NBAI)

Year	NL, WTMP & FLW weigh 10, 5, 1		NL, WTMP & FLW weigh 1, 1, 1	
	NBAI	Order of Top 5	NBAI	Order of top 5
1985	0.81		0.85	
1986	1.11		1.06	
1987	0.94		0.95	
1988	0.88		0.91	
1989	1.00		0.99	
1990	0.95		0.97	
1991	0.98		1.00	
1992	0.71		0.80	
1993	1.73	1	1.50	1
1994	1.47	3	1.32	4
1995	0.63		0.76	
1996	1.50	2	1.35	2
1997	0.75		0.82	
1998	1.46	4	1.33	3
1999	0.67		0.79	
2000	0.85		0.91	
2001	0.64		0.76	
2002	0.70		0.81	
2003	1.11		1.06	
2004	1.11		1.08	
2005	1.39	5	1.26	5

D. Threshold of anoxic volume to indicate anoxic problem

As addressed earlier, AV is a useful measure of volume containing DO less than a certain threshold critical to living resources. Water quality standards can be used to setup a DO-threshold to define anoxia for the calculation of anoxic volume. Although anoxic volume is not the best indicator of relative anoxia among years, many researchers may wish to use the conventional anoxic volume as the indicator of relative anoxia. For this reason, the following looks for an optimal DO-threshold for anoxic volume (AV) that may be used to compare relative anoxia among years.

The order of annual AI (of $DO \leq 0.2$ mg/l), especially for the top 4 years (Fig. 3), is similar to the order of annual AV of $DO < 0.2$ mg/l (Fig. 2), but is much less similar to the rankings of annual AV defined by other DO-thresholds, including the AV of $DO \leq 0.2$ mg/l, (Fig. 2). Therefore, we suggest using $DO < 0.2$ mg/l as the threshold for anoxic volume if we want to use anoxic volume as an alternative measure to compare relative anoxic problems among years in the Chesapeake Bay.

IV. DISCUSSION

1) The anoxic intensity metric weighs volume by the inverse of the DO concentration. This feature serves as an index for DO problems and can be used to compare relative DO problems among years for the Chesapeake Bay. The AI's of different DO-thresholds (0.2-1 mg/l, even higher) result in almost the same order in annual anoxic intensity, therefore, the aforementioned various DO-thresholds may be used for AI calculation. However, we recommend use the AI with a threshold of $DO \leq 0.2$ mg/l, though $DO \leq 0.3$ mg/l could be considered.

The years 1993, 1996, 1994, and 1998 are ranked, in descending order, as the top four most anoxic years for thresholds higher than $DO < 0.2$ mg/l (Fig. 3). We also calculated AI for a very low threshold ($DO \leq 0.01$ mg/l), The years 1998-2005 have zero values and are ranked after the 13th year. The ranks of AI in 1998-2005 for DO-thresholds higher than 0.1 mg/l are not always after 13th, and can even be ranked 4th or 3rd. This indicates that the weighting method does not entirely lean to the lowest DO concentration, but considers contributions from the water with higher DO concentrations. These suggest that such an index normalized by DO concentration could reasonably represent the extent of DO problems.

2) Anoxic volume may not fully represent the extent of DO problems. A water body having a larger anoxic volume could have less anoxia intensity than another water body having a smaller anoxic volume. This phenomenon is more evident if a larger volume of water body is involved, e.g., using a higher DO

concentration as the threshold. Therefore, the anoxic volumes defined by a low DO concentration may represent DO problems better. For example, the orders of anoxic volume among the five high years, 1993, 1994, 1996, 1998, and 2005, are significantly different between AV of $DO < 0.2$ mg/l and AV of $DO \leq 0.2$ mg/l (Fig. 1). This is due to years 1998 and 2005 having large volumes of $DO = 0.2$ mg/l, but not the other three years. Using a lower concentration (i.e., $DO < 0.2$ mg/l) as the threshold may avoid the influence of higher DO water on the anoxic volume calculation. Therefore, $DO < 0.2$ mg/l is preferred to $DO \leq 0.2$ mg/l as a threshold for anoxic volume in the comparison of anoxic problems. Such a choice is also proved correct when comparing the said AV with AI of $DO \leq 0.2$ mg/l. This also indicates that the normalization of DO concentration in the AI calculation is an important step in the assessment of anoxic problem.

Using a very low DO concentration to define anoxic volume is not recommended either, because many years will have zero anoxic volume. Also, very low DO concentrations may have much greater relative errors in DO measurement.

The AV and AI are two different metrics representing anoxia. The latter is a transformed feature from the former. The criterion of selecting a DO threshold for anoxic volume is based on the DO levels critical for living resources management, and the threshold of $DO \leq 0.2$ mg/l is commonly used [9]. The criterion for selecting a threshold for anoxic intensity is whether it can represent relative anoxic problems among years. Therefore, the optimal DO threshold for AV and AI can be different. However, considering the aforementioned DO value for living resources, here we also selected the threshold of $DO \leq 0.2$ mg/l for AI.

3) The anoxic intensity (AI) in Figure 3 is normalized by the DO concentration of each cell. The AI would be extremely large for the cells containing DO near zero. We need to set a quantitation limit (QL) for DO, below which a constant factor is used in the normalization. In fact, in the AI calculation for Figure 3 we used a factor of 15 for $DO \leq 0.07$ mg/l, i.e., treat (i.e., lump) all concentrations below this limit (0.07 mg/l) to be the limit value. There, $QL = 0.07$ mg/l. This quantitation limit is different from the "practical quantitation limit" (PQL) in the evaluation of laboratory analysis by USEPA guidelines [8], since the data we are using was after the evaluation. Although the DO sensors have a resolution of 0.01 mg/l for DO ranges of 0-20 mg/l, many data from the DO profiles in deep water in the summer were rounded at 0.1 mg/l interval (i.e., 0.3, 0.2, 0.1 and 0.0 mg/l). Our data used to calculate anoxic volume or anoxic intensity were after interpolation. In order not to have false flat out in the interpolated values, one more decimal point, i.e., two decimal points, of DO values were output and used to calculate AV or AI.

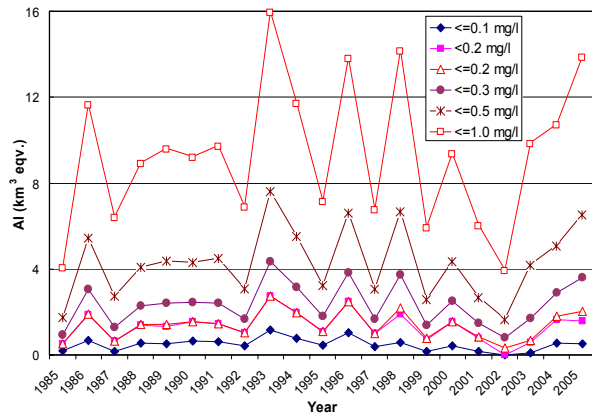


Fig. 4. Anoxic intensity (AI) of six thresholds for summer 1985-2005. Using value 10 as weighting factor for $\text{DO} \leq 0.1 \text{ mg/l}$.

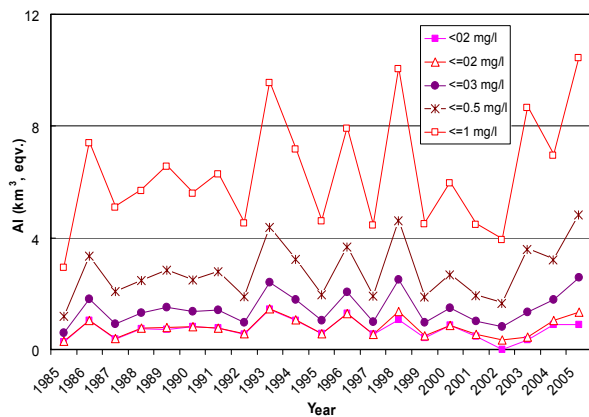


Fig. 5. Anoxic intensity (AI) of five thresholds for summer 1985-2005. Using value 5 as weighting factor for $\text{DO} \leq 0.2 \text{ mg/l}$.

Nevertheless, we still need to define a quantitation limit for AI calculation. In order to examine whether 0.05 or 0.07 or 0.1 mg/l or other concentrations are suitable as the quantitation limit for DO normalization using the interpolator outputs, we calculated AI for various DO-thresholds using a weighting factor of 20 or 15 or 10 or 5 to normalize anoxic volume for DO lower than 0.05 or 0.07 or 0.1 or 0.2 mg/l, respectively.

Using the quantitation limit of 0.07 mg/l (Fig. 3), the order of anoxic intensity among years is consistent with other analysis such as nutrient loading index and fishery responses. The order of annual AI by using 0.1 mg/l as the quantitation limit (Fig. 4) is close to that using 0.07 mg/l as the quantitation limit (Fig. 3). The order of annual AI by using 0.05 mg/l as the quantitation limit (not shown) differs somehow from that using 0.07 or 0.1 mg/l as the quantitation limit. The order of annual AI using 0.2 mg/l as the quantitation limit (Fig. 5), as well as using 0.3 mg/l (not shown), has large differences

from those in Fig. 3 and are inconsistent with other analysis of anoxia.

The above suggests that it is better to use 0.07 or 0.1 mg/l as the quantitation limit (QL), rather than 0.05 or 0.2 mg/l or higher concentrations. It is difficult to choose which one, 0.07 or 0.1 mg/l, as the QL. If we consider that the concentration intervals in the lab report are mostly at 0.1 mg/l, we may use 0.1 mg/l as the QL. On the other hand, if we consider that most DO values in the data interpolation and other intermediate calculations consists of at least 2 decimal points, and 2 decimal points of DO measurements were also reported from the lab, $\text{QL}=0.07 \text{ mg/l}$ may also be acceptable. Ironically, $\text{DO}=0.07 \text{ mg/l}$ corresponds approximately to 1% air saturation at 30°C in 10 ppt saline water and is used to define anoxia by some marine biologists [1]. Further data mining is required for the choice, which is beyond the scope of this work.

Note: The above method of lumping DO concentrations under a quantitation limit does not affect anoxic volume calculation, because anoxic volume is a lumped volume without considering DO differences in individual cells as long as the DO is below the threshold.

4) The mass of DO deficit in water under a specific depth (depth-volume), in reference to DO mass in full saturation, may be used to compare relative DO problems among years. However, we found that high DO deficit years and low DO deficit years are quite different for different depth-volumes and disagree with other analysis of anoxic problems. DO mass in the Bay can increase due to inputs from ocean water and freshwater runoff. The input conditions vary in different sampling cruises. The anoxic water in lower layers is usually affected less by the above processes. If we have a fine time-scale of sampling, the annual average DO mass could represent relative DO problems. However, we only have two sampling cruises in a month. The cruise met with higher runoff may overestimate DO mass. In addition to the complicated transport system in the Chesapeake estuary, different spatial distributions of anoxic water and pycnocline at different times cause difficulty in finding a referent depth suitable to calculate the DO mass deficit. Therefore, DO mass is insufficient to compare relative DO problems among years in the Chesapeake Bay.

5) The data used in the annual summer anoxia assessment was based on observations. It is important to select comparable number of sampling cruises during the same period in each year. Anoxic volume or anoxic intensity can change greatly between two cruises. Using the average summer anoxia from selected sampling cruises is a proper way in the above assessment of DO problems, though data mining on individual cruises is prerequisite.

In this work, data mining occurred throughout the

processes, including selection of sampling cruises, analyzing lab reports and the quality of DO measurement, data interpolation, feature transformation, data grouping and pattern comparison, regression, and correlation analysis. An omitted or inaccurate calculation in one step could introduce significant errors into the next.

V. CONCLUSIONS

This work demonstrated that data mining is an important step before data utilization, especially for spatial and temporal data in a complex system. Through data mining we learned the following lessons on the assessment of dissolved oxygen problems in the Chesapeake Bay.

The orders of anoxic volumes among years are different by different DO-thresholds of anoxia. Anoxic volume is not an optimal metric of relative anoxia among years. Anoxic volume can be affected by partial mixing or other mechanisms, which cause a year to have higher anoxic volume but less anoxic intensity than other years.

Anoxic intensity weighs volume by the inverse of DO concentration. This feature serves as an index for DO problems and can be used to compare relative DO problems among years for the Chesapeake Bay. Though AI of various DO-thresholds (from 0.2 to 1 mg/l) could be utilizable because they have almost the same order in anoxic intensity among years, the threshold of $DO \leq 0.2$ mg/l or ≤ 0.3 mg/l is recommended.

Anoxic volume measures the volume of water under a certain threshold. It reflects the volume of water that is unavailable to the aquatic species that requires DO above the threshold, and is a useful indicator in aquatic living resources management. Therefore, the threshold for anoxic volume should be set water quality standards for example, $DO \leq 0.2$ mg/l as established by the Chesapeake Bay Program [9], [10].

In the case when anoxic volume is used as an alternative to indicate relative anoxic problems among years, the anoxic volume with DO-threshold < 0.2 mg/l may be used.

Accuracy in DO measurement is important for accurate data interpretations, especially for low DO concentrations. The examination of the quality of DO monitoring data through data mining is important in the anoxia assessment.

The above conclusions are based on data mining from the Chesapeake Bay water quality database for the years 1985-2005. Different protocols in field sampling and instrument analysis in different areas may yield different conclusions on the thresholds for dissolved oxygen in the anoxia comparison. Nevertheless, data mining is important to ensure accurate assessments of problems in complicated systems.

APPENDIX – Acronyms

AI -- anoxic intensity; AV -- anoxic volume;
DO -- dissolved oxygen;
NBAI -- nutrient based anoxic index;
NLI -- nutrient loading index;
PQL -- practical quantitation limit;
QL -- quantitation limit; V -- volume;
WTMP -- water temperature.

REFERENCES

- [1] S. M. Baker and R. Mann, "Effect of hypoxia and anoxia on larval settlement, juvenile growth, and juvenile survival of the oyster *Crassostrea virginica*", *Biol. Bull.* 182: 265-269, 1992.
- [2] D. Jasinski, G. Shenk, E. Parry, W. Dennison, B., Longstaff, and M. Williams, *Summer 2005 ecological forecast technical documentation*. Chesapeake Bay Program, Annapolis, MD, USA, July. 10pp, 2005.
- [3] R. V. Thomann, J. R. Collier, A. Butt, E. Casman, and L. C. Linker, *Response of the Chesapeake Bay water quality model to loading scenarios*, Chesapeake Bay Program Office, Annapolis, MD, USA. CBP/TRS 101/94, 1994.
- [4] P. Wang, L. C. Linker, R. A. Richard, and C. F. Cerco, "Surface analysis of Chesapeake Bay water quality response to different nutrient and sediment loads", *J. Envir. Engr.*, 132: 377-383, 2006.
- [5] P. Wang, L. Linker, D. Jasinski, W. Dennison, G. Shenk, and R. Batiuk, "Forecast of summer anoxia for the Chesapeake Bay", *Estuarine and Coastal Modeling: Proc. Ninth International Conference*, Charleston, SC, USA, Oct 31-Nov 2, 2005, M.L. Spaulding (ed.). 20pp, in press, 2006.
- [6] L. Bahner, *The Chesapeake Bay and tidal tributary volumetric interpolator, software VOL3D version 4.0*. NOAA Chesapeake Bay Office, Annapolis, MD, 2001.
- [7] J. D. Hagy, Boynton, W. R., Keefe, C. W., and K. V. Wood, "Hypoxia in Chesapeake Bay, 1950-2001: long term change in relation to nutrient loading and river flow", *Estuaries* 27(4) 634-658, 2004.
- [8] USEPA, *Revised assessment of detection and quantitation approaches*, EPA-821-B-04-005, 2004.
- [9] CBPO, *Ambient water quality criteria for dissolved oxygen, water clarity and chlorophyll a for the Chesapeake Bay and its tidal tributaries*, USEPA Chesapeake Bay Program Office, Annapolis, MD, USA. EPA 903-R-03-002, April 2003.
- [10] S. C. Jordan, M. Stenger, M. Olson, R. Batiuk, and K. Mountford, *Chesapeake Bay dissolved oxygen goal for restoration of living resource habitats*, Chesapeake Bay Program Reevaluation Report 7c, Annapolis, MD, CBP/TRS 88/93, 81pp, 1992.