ACS ES&T Water

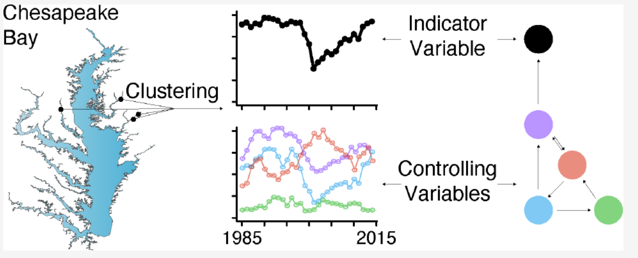
pubs.acs.org/estwater Article

Inferring Controls on Dissolved Oxygen Criterion Attainment in the Chesapeake Bay

Ryan E. Langendorf,* Vyacheslav Lyubchich, Jeremy M. Testa, and Qian Zhang*

Cite This: ACS EST Water 2021, 1, 1665−1675 Read Online

ABSTRACT: Environmental monitoring programs generate multivariate time series for the assessment of ecosystem health. Recent developments in causal inference offer ways to translate these observational data into networks able to explain gains and losses in the trajectories of indicator variables. Here, we present a case study of this approach using surface water dissolved oxygen (DO) criteria attainment across the Chesapeake Bay. We sought to understand the long-term fluctuations of DO in response to external watershed inputs (freshwater, nutrients, sediments) and internal water properties (water temperature, chlorophyll-a, nutrient concentrations). We found contrasting controls across different regions of the Chesapeake Bay. Summer freshwater and sediment inputs reduced surface water DO criteria attainment in landward regions but elevated attainment in open, mainstem waters. Algal biomass was often positively associated with AD in surface waters, in contrast to deep waters where algae are well-understood to lower oxygen levels. Using estimates of net effects, we also discovered that sediment flowing into many of the mainstem segments during the winter-spring season has had a net negative effect on oxygen levels. These segments have yet to deteriorate, which has masked this risk. This work demonstrates the utility of long-term monitoring programs to better understand and manage complex ecosystems.



KEYWORDS: *Chesapeake Bay, water quality, dissolved oxygen, causal inference, network, management*

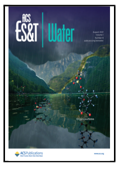
1. INTRODUCTION

Managing ecosystems involves tracking systemic properties designed to identify the risk of degradation or failure to achieve stated restoration goals. Indicator variables are useful representations of systemic risk, providing early warning signs like slowed responses to shocks,[1] changes in soil biological and chemical properties,[2] or proxies for habitat suitability.[3] Successful management designs often include quantitative ways to identify ecosystems at risk[4−6] but lack clear recommendations for how to manage these systems toward prescribed goals under uncertainty in climate, intervention effectiveness, and system understanding.

Advances in the analysis of time series data may help management programs by identifying causes of change in indicator variables that interventions can be designed to prevent or ameliorate. This is especially helpful because, with a renewed focus on remote and observational data collection (e.g., the creation of NEON;[7] see also refs 8−10), time series on relevant covariates are increasingly available. The central idea in this paradigm is to convert time series of variables into a network describing interactions between the variables. This network can then be interrogated to explain the causes of indicator variable dynamics and predict interventions most likely to help.

Here, we explore the value of this time series approach to indicator variable-based management of ecosystems, using the Chesapeake Bay as a case study. The Chesapeake Bay's watershed of more than 100 rivers spans six states along the Atlantic coast of the United States, supporting more than 3000 species of plants and animals. The Bay has suffered eutrophication for decades resulting in excessive algal growth, decreased submerged aquatic vegetation acreage, and low concentrations of dissolved oxygen (DO).[11−16] In response, the Chesapeake Bay Program partnership developed a guidance framework to establish water quality criteria for DO, water clarity, and chlorophyll-a for 92 segments across the Bay's tidal waters[17,18] (Figure 1). To reduce nutrient and sediment loads delivered to the Bay, coordinated management efforts among the Bay state jurisdictions began in the 1980s, and in 2010, the Chesapeake Bay total maximum daily load (TMDL) was established to enforce nutrient reductions to achieve habitat health that can fully support living resource survival, growth, and reproduction.[19] The TMDL is designed to ensure that all pollution control measures needed to fully restore the Bay and its tidal tributaries are in place by 2025.

Received: December 21, 2020
Revised: July 1, 2021
Accepted: July 2, 2021
Published: July 23, 2021



https://doi.org/10.1021/acsestwater.0c00307
ACS EST Water 2021, 1, 1665−1675

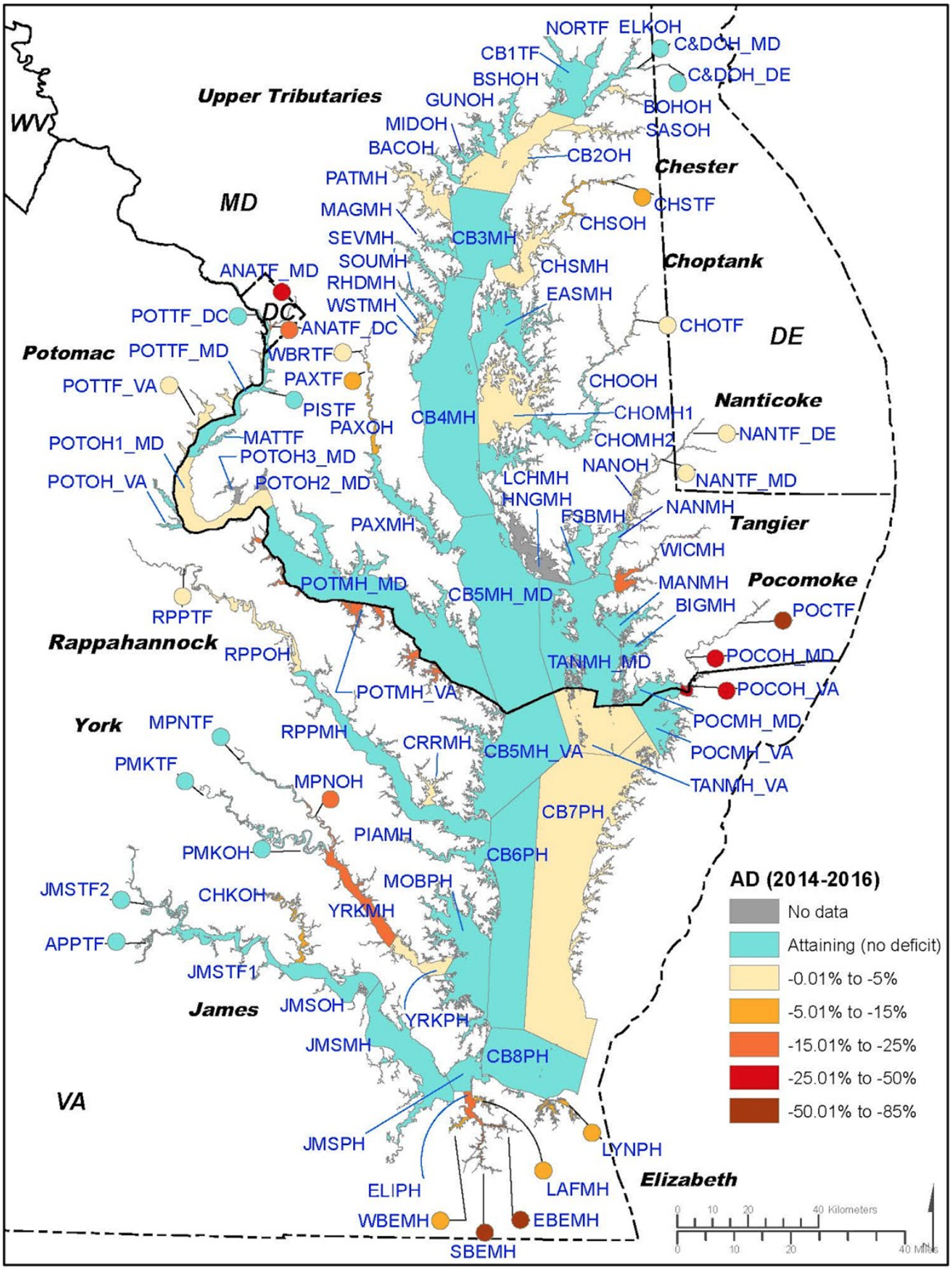


Figure 1. Chesapeake Bay open-water segments with attainment deficit (AD) in 2014−2016. Full names of the segments are available from Table S1; AD data are available from Table S2. The authors would like to thank Zhaoying Wei (Chesapeake Bay Program) for her assistance on this figure.

The indicator variable "attainment deficit" (AD) was developed as part of these Bay-wide monitoring and assessment efforts.[20] AD measures dissolved oxygen status as an indicator of water quality and ecological health as they pertain to habitat for metazoan life. Zhang et al.[20] documented the estimated AD for each tidal segment for each running three-year periods during 1985−2016. The authors analyzed AD across different depth-based habitats and detected significant,

https://doi.org/10.1021/acsestwater.0c00307
ACS EST Water 2021, 1, 1665−1675

improving trends during 2000−2016 in 15 open-water segments. However, clear linkages between AD status in the 92 tidal segments and underlying external and internal drivers remain elusive and have yet to be explored. Furthermore, while the majority of prior research in coastal ecosystems has explored controls on oxygen in deep waters where depletion is most severe, fewer, if any, studies have sought to understand controls on oxygen availability as it pertains to habitat in surface water environments.

In the above context, the main objectives of this work are two-fold: (1) explore the spatial and temporal patterns of AD in 92 open-water segments of the Chesapeake Bay that represent the full diversity of environments in the estuary, and (2) use tools from the burgeoning fields of nonlinear dynamics, network science, and causal inference to investigate whether and how AD has been affected by potential external and internal drivers, including management actions (e.g., reduction of nutrient loads), hydrological and climatic variability, and hydrodynamic and biogeochemical characteristics.

2. MATERIALS AND METHODS

2.1. Attainment Deficit (AD). AD quantifies the amount of space-time exceedance of the dissolved oxygen criterion. Because dissolved oxygen varies substantially over multiple spatial and temporal scales, this integrated measure of oxygen provides a useful metric of the duration, extent, and intensity of habitat condition. For a specific tidal segment and assessment period, AD can be computed by the following steps: (1) collect station-level DO monitoring data, (2) interpolate DO concentrations among stations, (3) compare interpolated values with designated-use specific criterion values to quantify criteria exceedance, (4) construct the cumulative frequency curve of criteria exceedance for all sampling events (also called "attainment curve"), (5) develop a reference curve of allowable criteria exceedance, (6) intersect the attainment curve and the reference curve to determine the nonallowable criteria exceedance, and (7) scale the nonallowable criteria exceedance to obtain AD.[3,16,17,21]

AD is always in the range from 0% to −100%. An attainment deficit of 0%, which is the best possible condition, indicates that the minimum water quality requirements for protecting aquatic life have been met. An attainment deficit of −100%, which is the worst possible condition, implies complete noncompliance, which has substantial negative effects on living resources' survival, growth, and reproduction. Any other values also indicate noncompliance, with values closer to −100% implying more harmful conditions. So while AD is a less intuitive negative percent, relationships with variables influencing AD are intuitive in that positive associations describe beneficial relationships and negative associations describe harmful relationships.

AD is an important indicator of a habitat inclusive of many types of organisms and also provides a summary of ecological conditions, given that dissolved oxygen levels integrate the effects of a variety of physical and biological processes. An advantage of using the Chesapeake Bay as a case study is the availability of long-term monitoring data. We focused on the so-called open water (OW) segments of the Bay because of their wide spatial coverage and ecological relevance, as the OW designated use covers the entire tidal surface of the Chesapeake Bay (i.e., all 92 tidal segments). These surface waters are a potential habitat for a wide variety of organisms in the Chesapeake Bay, including sport fish (striped bass, bluefish, mackerel), forage fish (menhaden, silversides), and demersal organisms in the most shallow, vertically mixed segments. Additionally, long-term changes in oxygen within OW segments of the Chesapeake Bay have been less studied than hypoxia-vulnerable bottom waters, even though oxygen conditions can reach harmful levels in OW environments.

For this study, AD values were obtained from Zhang et al.[20] for each of the 92 OW segments of the Chesapeake Bay (Figure 1) for three-year assessment periods between 1985−1987 and 2014−2016. For OW DO, June−September was evaluated for the 30 day mean criterion in the attainment assessment procedures, in which the criterion thresholds were 5.5 mg/L for very low-salinity regions and 5 mg/L otherwise. Full names of the segments are available from Table S1, and AD data are available from Table S2.

2.2. Water Quality Parameters. Water quality measurements were obtained from the U.S. Environmental Protection Agency Chesapeake Bay Program Water Quality database[22] and included Secchi depth (Secchi), light attenuation coefficient (KD), surface-layer (i.e., at depth ≤1 m) salinity (SAL), water temperature (WTemp), and concentrations of chlorophyll-a (CHLA), DO, nutrient (dissolved inorganic nitrogen, DIN; total nitrogen, TN; total phosphorus, TP), and total suspended solids (TSS). These variables were measured every 2−4 weeks at 133 tidal monitoring stations in the Chesapeake Bay and its tributaries. Riverine discharge (Flow) and loadings (SSload, TNload, and TPload) were recorded at nine river input monitoring stations located at the main tributaries. We aggregated the flow and loadings across the 133 monitoring stations assuming that the downstream stations and stations further south in the bay (i.e., closer to the Atlantic Ocean) experience cumulatively aggregated loadings.[23]

All water quality variables at the station level were processed to create a consistent set of monthly average values by following the approach from ref 23 for filling-in missing data using information from neighboring stations. Time series from the stations with documented changes in data collection laboratories or methodology were adjusted to remove potential level shifts at the times of the changes.[23] The monthly station data were then interpolated onto a dense 250 m grid and averaged within each segment of the bay. The interpolated monthly values within each segment were then aggregated within the following periods in each year: January−May (variables with the suffix '_JanMay') and June−September (suffix '_JunSep') for the external loadings and June−September (suffix '_JunSep') for all estuarine variables. Finally, a three-year moving average was applied so the temporal resolution of water quality parameters matched that of the AD data.

2.3. Statistical Methods. The local-scale heterogeneity of ecological data is important to consider in ecosystem studies, since the aggregation of data might mask existing relationships or even lead to the phenomenon of "ecological fallacy" where conclusions differ across levels of data aggregation.[24,25] One of the ways to minimize ecological fallacy and aggregation bias is by controlling how units of analysis are aggregated, with a goal of obtaining homogeneous groups (clusters).[26] Then, the models can be estimated for each cluster separately.

In many applied studies, including this one, the best way to assign units of data to clusters is unknown. Even the total number of clusters is typically unknown. Hence, methods of unsupervised learning are used to identify clusters on the basis of patterns in the data. Here, we clustered the bay segments to

https://doi.org/10.1021/acsestwater.0c00307
ACS EST Water 2021, 1, 1665−1675

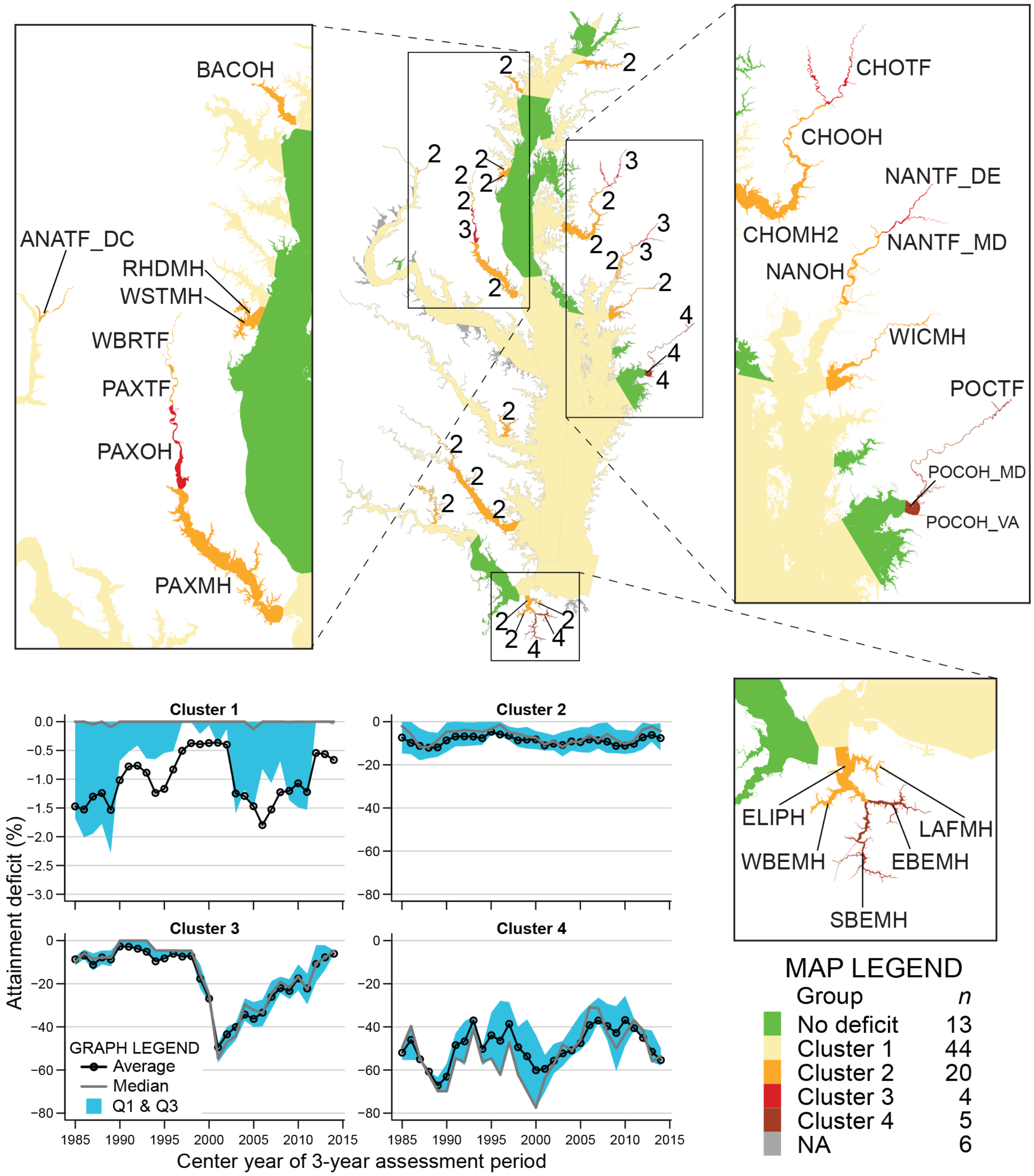


Figure 2. Map with results of hierarchical clustering of attainment deficit in 1985−2014 in open-water segments of the Chesapeake Bay (n denotes the number of segments in each group). Graphs show the aggregated dynamics of attainment deficit in each cluster (note a different scale for cluster 1). The 92 open-water segments and their assigned clusters are listed in Table S1.

identify clusters with homogeneous dynamics of AD. Due to specifics of the watershed, homogeneous dynamics may be observed at locations spatially distant from one another (e.g., see clustering results for water temperature[27] and TSS[28]), so we used agglomerative clustering without specifically considering spatial proximity of the segments. Euclidean distances between AD time series and Ward's method for finding minimum-variance compact clusters were used to cluster the segments.[29,30] The number of clusters k was identified as the minimal number of clusters that achieve at least 75% reduction of the total within-cluster sum of squared errors. The analysis was done using R package `stats`.[31]

We attempted to infer the direct effects of variables on one another with two state-of-the-art causal discovery methods: convergent cross mapping (CCM[32]) and the Momentary Conditional Independence extension of the PC algorithm (PCMCI).[33] CCM quantifies the probabilities of bidirectional causation between pairs of variables in nonlinear deterministic systems with variables that are weakly to moderately coupled. This method uses the insight of Takens theorem[34] that the

https://doi.org/10.1021/acsestwater.0c00307
ACS EST Water 2021, 1, 1665−1675

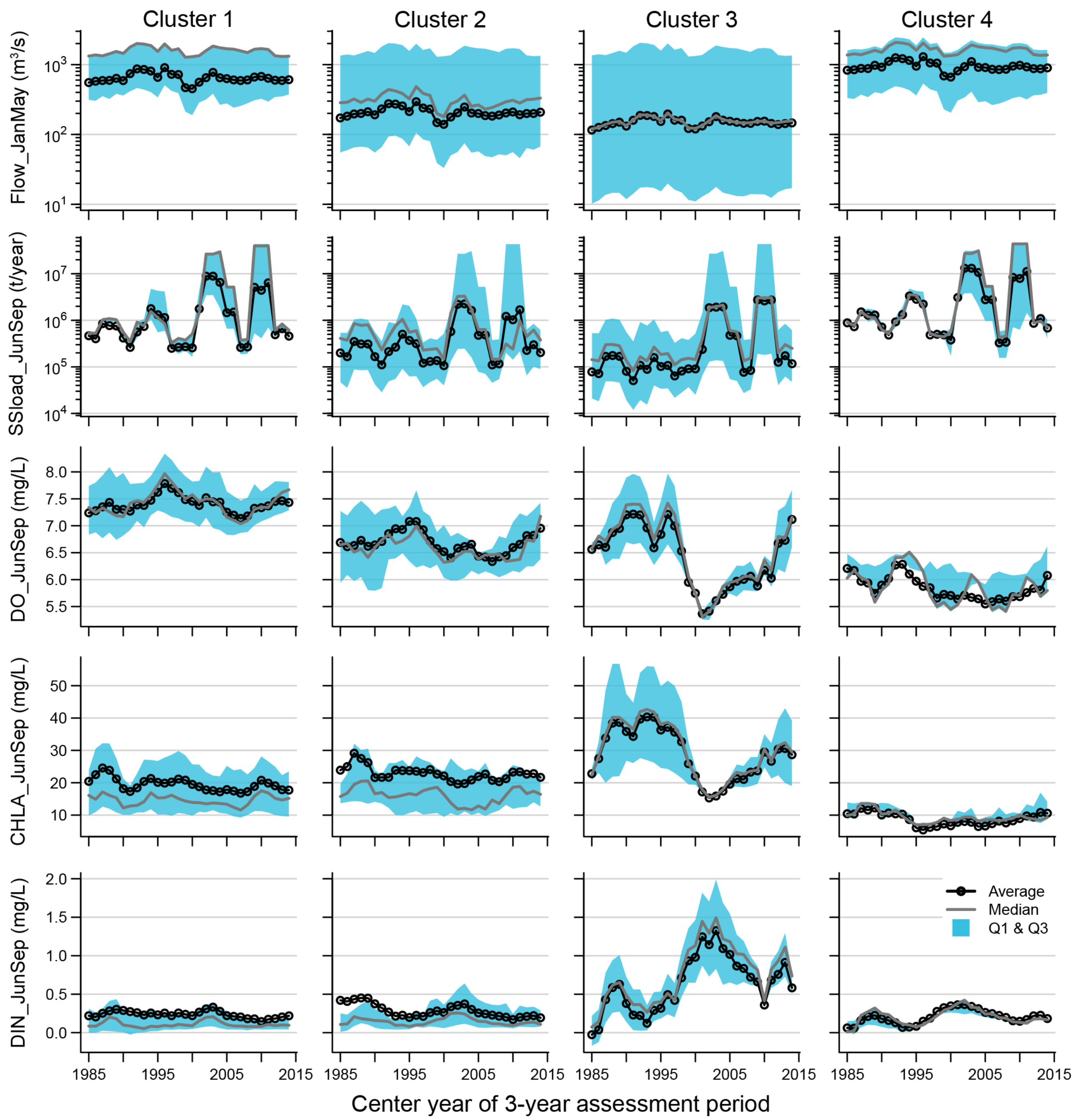


Figure 3. Dynamics of selected water quality parameters per cluster (cf. attainment dynamics in Figure 2). Note that Flow_JanMay and SSload_JunSep are plotted on $\log_{10}$ axes.

dynamics of a system of coupled variables can be reconstructed using time lags of only one of the variables. When the resulting "shadow manifold" of a variable does a better job predicting another variable's shadow manifold (called cross mapping) with increasing amounts of data (called convergent cross mapping), the second variable is thought to have caused the first. PCMCI, the other kind of causal discovery we applied, works in two steps: (1) use the PC algorithm,[35] named after its two authors, to find the possible causes of each variable, and (2) refine the causal graph produced in the first step by using its interactions as conditions in tests of momentary conditional independence (MCI) between putative cause and effect variables. PCMCI has the advantage of not assuming a kind of dynamical system and so not requiring time series with particular characteristics.

We also integrated our own expert opinions into a consensus set of interactions, which we then estimated by fitting structural equation models (SEM;[36] R package `lavaan 0.6−6`[37]) for each cluster. SEM is a well-established tool for causal inference in systems with complex interactions (e.g., see studies in refs 38−40) that relies on assumptions of the specific causal interactions on the basis of the knowledge of researchers.[41] To tailor our consensus set of interactions (56 total) to each cluster in a data-driven way, we iteratively removed nonsignificant (if any; $\alpha = 0.05$) interactions until only statistically significant interactions remained. The removal was done one at a time within each cluster, starting with the largest p-value.

The direct effects commonly shown in SEM pathways (e.g., Figure 4) are difficult to interpret with as many interacting

https://doi.org/10.1021/acsestwater.0c00307
ACS EST Water 2021, 1, 1665−1675

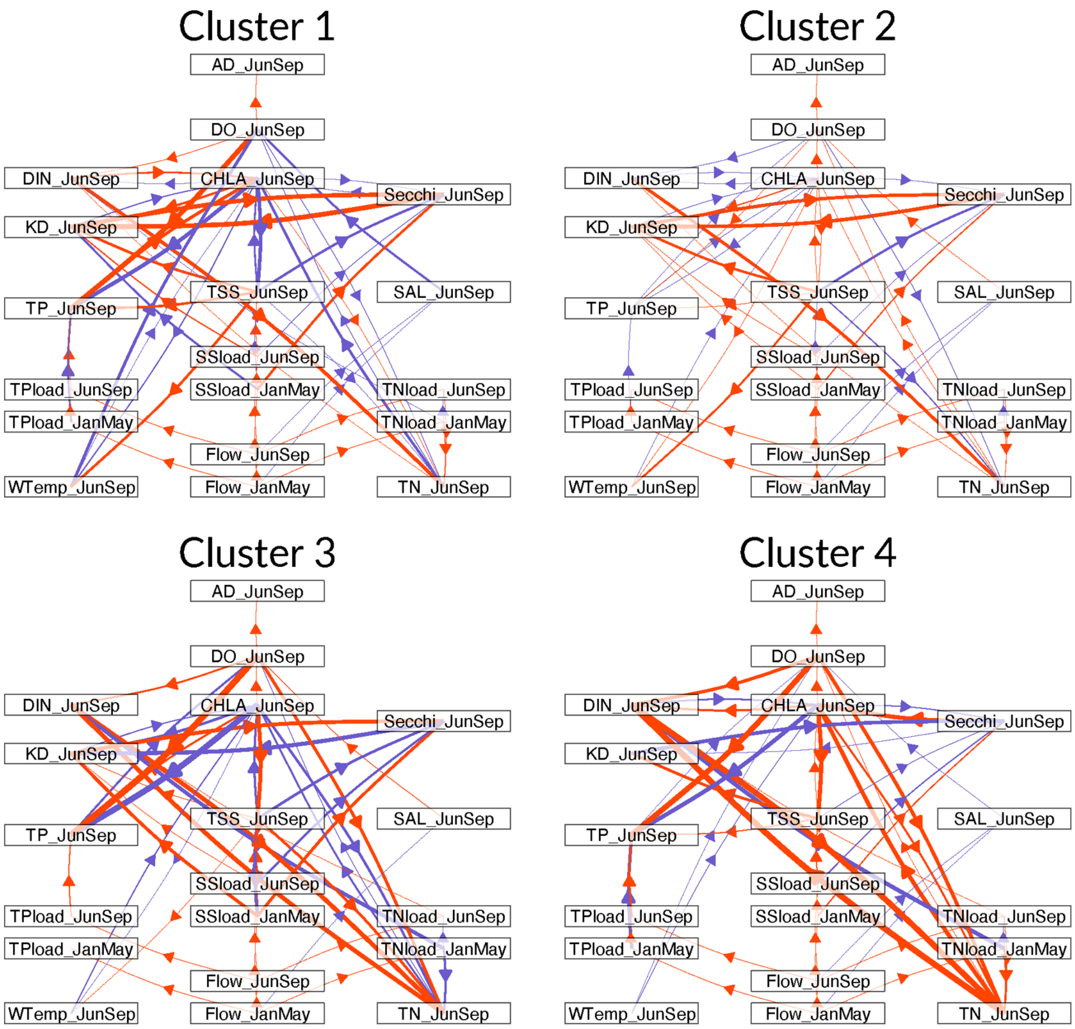


Figure 4. Cluster-specific structural equation models. Arrows show direct causal effects (pointing from the cause to the effect variable), with the width being proportional to the strength, and color (blue/red) representing the sign (negative/positive) of the effect, such that when the cause variable increases, the effect variable decreases/increases. Each cluster's interactions started with the same expert opinion consensus graph that was iteratively winnowed down to retain only statistically significant (α = 0.05) effects. The strength and statistical significance of each link are summarized in Tables S4−S7.

variables as govern dissolved oxygen in the Chesapeake Bay. While these pathways show useful mechanistic relationships, the combined effects of multiple interactions are hard to intuit from these networks. Since we are primarily concerned with drivers of AD, we have calculated the total (net) effects of each variable on AD within each cluster-specific SEM. These net effects were calculated as the sum of the effects of every acyclic path connecting that variable to AD in each cluster's SEM.[42] To find all such paths, we used a breadth-first search using R package `igraph 1.2.5`.[43] The effect of each path was calculated as the product of the standardized SEM coefficients at each interaction along the path from the causal variable to AD.

3. RESULTS AND DISCUSSION

3.1. Patterns of AD by Cluster. We analyzed a 32-year record (1985−2016) of open-water attainment of DO conditions in 92 segments of the Chesapeake Bay and its tributaries (Figures 1 and 2). Thirteen segments were removed that attained dissolved oxygen criteria at every time during the period of record (AD = 0% always, hence, not correlated with time series of explanatory variables; green segments on the map in Figure 2). We also removed six other segments (shown as "NA" in Figure 2), including five segments that had more than half of AD records missing and one segment with a small area (0.22 km^2).

Across the 73 segments of the Chesapeake Bay, we found four clusters with unique temporal AD dynamics (Figure 2). Marked differences (up to 50 percentage points) among these four clusters illustrate substantial spatial heterogeneity and suggest that each cluster may have been responding to different controlling factors. Oxygen attainment was high in both clusters 1 and 2 (AD > −2.5% and > −20%, respectively) and mostly constant through time, with cluster 1 having the highest attainment (AD closest to 0) and lowest absolute variability among the four clusters. The segments in cluster 1 (where AD ~ 0) and where no deficit existed share a large surface area, suggesting that exposure to wind stress may allow for surface water oxygen concentrations to stay nearer to equilibrium with the atmosphere (and thus in attainment) than the smaller, more protected segments. Interestingly, summer chlorophyll-a and DIN both appeared to decrease modestly over time in clusters 1 and 2 (Figure 3), consistent with recent reports of modest long-term water-quality improvements in some regions

https://doi.org/10.1021/acsestwater.0c00307
ACS EST Water 2021, 1, 1665−1675

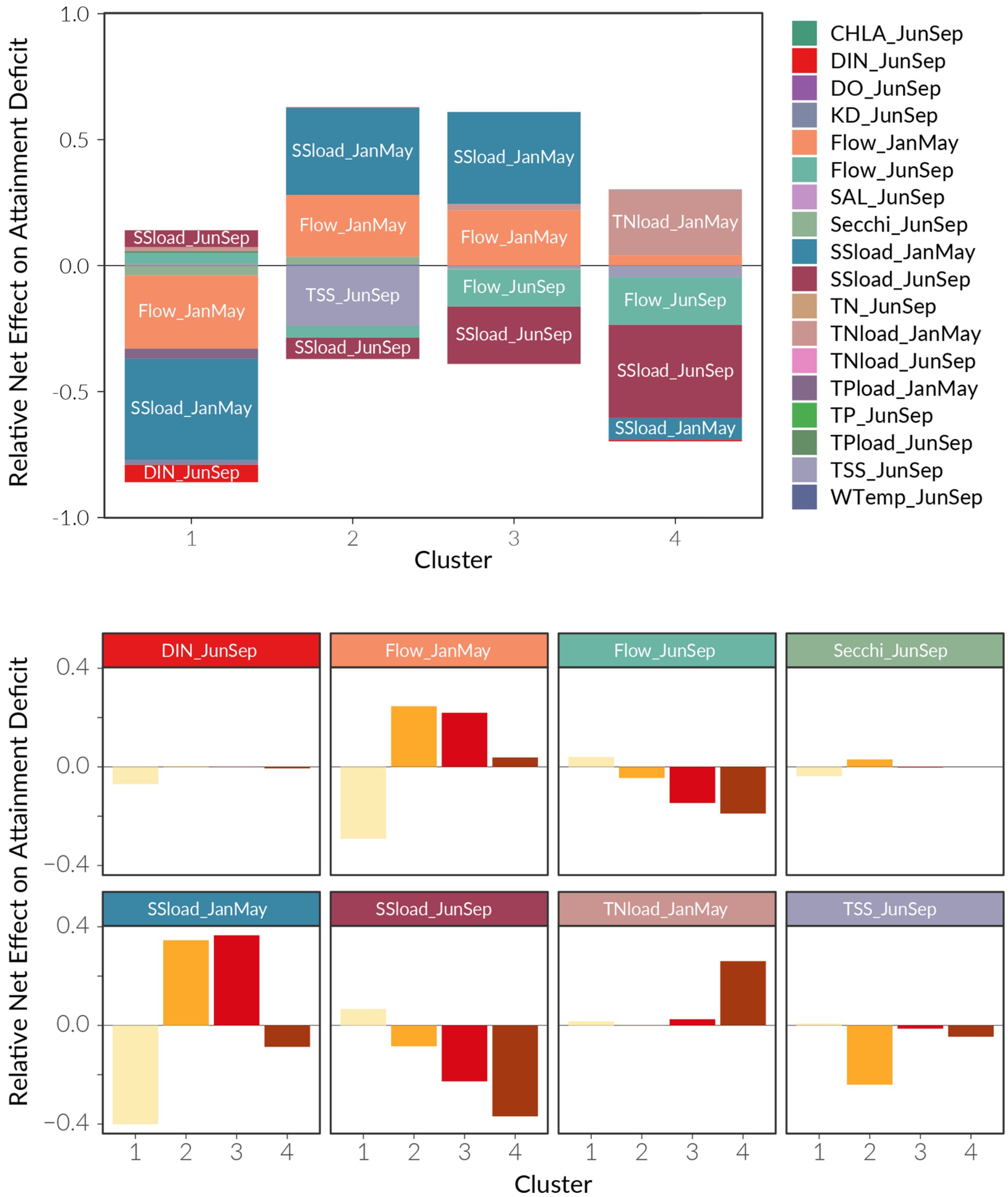


Figure 5. Relative total (net) effects on AD. These are the sum of every acyclic path connecting each variable to AD in each cluster's SEM (Figure 4), scaled to sum to one within each cluster. The strength of each path is the product of the standardized coefficients along it. Important drivers are plotted individually on the right showing cluster-specific causes of AD dynamics. Winter−spring flow (Flow_JanMay) and corresponding suspended solids (SSload_JanMay) are most responsible for the net reduction of attainment in clusters 1 and 4.

of this system.[15,40] In contrast, attainment was relatively low in clusters 3 and 4, with minimum AD values of about −60% and −80%, respectively (Figure 2). Cluster 3 included the most substantial variability in oxygen, chlorophyll-a, and DIN (Figure 3), consistent with its strong temporal variability in AD (Figure 2). The difference in AD between clusters was likely driven by differences in segment size, proximity and connectivity to land and associated water quality dynamics. For example, the mainstem segments are mostly in cluster 1, which is the biggest cluster in segment count and total surface area, and cluster 2 contains many of the larger, more downstream tributary segments connected to the mainstem. Both clusters 3 and 4 comprise a small number of segments (4 and 5 segments, respectively) and are primarily located in very upstream reaches of tributaries surrounded by land (Figure 2); cluster 3 segments are mostly in low-salinity regions, whereas cluster 4 segments are in high-salinity waters (Figure 2 and Figure S1).

3.2. Causes of AD Dynamics. We relied on prior knowledge to create a consensus set of interactions, which were estimated as a structural equation model. In tailoring the SEM to each cluster, we found 54, 41, 32, and 35 significant interactions for clusters 1−4, respectively (output in Tables S12−S23). The resulting causal graphs relating AD to biogeochemical and physical processes in the bay inferred from their respective time series (Figure 3) found similar driving variables, primarily including nutrient and sediment inputs, but the relative importance, season, and direction of

https://doi.org/10.1021/acsestwater.0c00307
ACS EST Water 2021, 1, 1665−1675

their effects on AD were cluster-specific (Figure 4). Freshwater inputs were most responsible for AD dynamics across all four clusters of the Chesapeake Bay, which is consistent with prior work showing that sediment and nutrient runoff from the Chesapeake Bay's watershed impact stratification and residence times, affecting light availability and associated phytoplankton growth, which can result in eutrophication.[11,44,45]

This strong impact of freshwater inputs on AD, however, was seasonally dependent and did not consistently improve (positive association) or worsen (negative association) AD (Figure 5). In clusters 2 and 3, winter–spring riverine inputs (Flow_JanMay) improved AD. The corresponding strong covariation of chlorophyll-a and DO in cluster 3 (Pearson correlation = 0.89, but also compare their time series in Figure 2) is consistent with several prior investigations in the Chesapeake Bay that have related winter–spring river flow to spring and summer phytoplankton biomass and primary production (e.g., refs 46 and 47), where river inflows and associated nutrient loads create nutrient-replete conditions that support elevated primary production. The fact that both winter–spring flow and suspended sediment load improved AD in cluster 3 illustrates the known linkage between flow and suspended sediment input[48] but also suggests that the beneficial effects of flow on AD outweighed the co-occurring but harmful effects of sediments on light available for primary production in this cluster. It is also possible that the elevated flushing or turbidity effects from the winter–spring flow, which would otherwise reduce phytoplankton growth and AD, had expired by summer but that the nutrient enrichment effects persisted to support phytoplankton growth. In contrast, summer river flow worsened AD, particularly in clusters 3 and 4 where AD is low when the summer river flow is high. In these clusters, the proximal effect of summer river flow appeared to be that either (a) elevated flushing reduced biomass and production (e.g., ref 49) or (b) increased turbidity associated with sediment inflows[23] limited photosynthesis and associated oxygen production during this season. In fact, cluster 4 includes the highly enriched small and deep tributaries of the Elizabeth River where harmful algal blooms often develop in response to high residence time (e.g., refs 50 and 51), consistent with our findings here. In general, the net causes of AD (Figure 5) highlight the contrast between downstream estuarine regions and more upstream regions. In the former, elevated river flows and nutrient inputs generate high phytoplankton production and elevated DO, and thus, AD improves. In contrast, in more upstream regions, elevated river flows result in declines in chlorophyll-a associated with flushing and light limitation that worsen AD.[52]

Compared to external freshwater inputs to the Chesapeake Bay, the impact of water temperatures on AD was muted over 1985–2016. Although increasing water temperature worsened AD in clusters 1 and 4 and was linked to a lower AD in 49 of 73 segments, there was little covariance between temperature and AD in clusters 2 and 3 (Pearson correlation −0.02 and −0.27, respectively), and the net effect of temperature on AD was less than 1% of all drivers in all four clusters (0.001%, 0.4%, 0.007%, and 0.0009%, respectively; Figure 5). This result is somewhat surprising given the multiple mechanisms by which elevated temperature can reduce DO concentrations, including lower solubility and elevated respiration (e.g., refs 53 and 54) and recent evidence for water temperature increases in the Chesapeake Bay over the past three decades.[54,55] The relative impact of observed summer water temperature changes (e.g., ref 15) on AD may simply be muted relative to other variables in our analyses. For example, the three-year running means we used to analyze AD may preclude temperature impacts from emerging in the causal inference, given that temperature may not impart a "memory effect" on water quality that persists for many years. Even so, given the high interannual variability in discharge to the Chesapeake Bay and most estuarine ecosystems, our results indicate that future changes in discharge (e.g., ref 56) will be a primary driver of the Chesapeake Bay's water quality.

In all but cluster 2, total phosphorus and nitrogen improved AD (Figure 5), which is not consistent with the well-described eutrophication process where elevated nutrient concentrations enhance algal biomass and increase oxygen consumption in bottom waters (e.g., refs 11 and 57). However, because the AD we analyzed represents surface water conditions; the positive link between nutrients and AD is consistent with the nutrient stimulation of phytoplankton biomass and oxygen-generating primary production. Indeed, nitrogen and phosphorus are incorporated into growing phytoplankton biomass, which is consistent with our finding that chlorophyll-a (a proxy for phytoplankton biomass) directly improved AD in all but cluster 1 (Figure 4, Tables S4–S7) and indirectly improved AD in all four clusters (Tables S16–S19). A consideration from these results is that nutrient reduction goals, if met, may have their intended effect of reducing primary production and associated surface water oxygen production, thereby having the unintended consequence of worsening surface water AD. Thus, management actions aimed at reaching dissolved oxygen criteria may need to be adapted to consider discrepancies between the impacts of eutrophication in surface versus bottom waters and the trade-offs associated with spatially decoupled surface oxygen production and bottom oxygen consumption processes. We should also emphasize that the oxygen concentrations used in this analysis are measured during the day, thus reflecting conditions when phytoplankton are actively growing and generating oxygen and that this analysis is potentially biased toward the positive impacts of phytoplankton biomass on oxygen.

This tension between surface and bottom water management warrants further attention, particularly because it appears that winter–spring flows were responsible for the improvement in AD in cluster 3 beginning in 2001. There was a prolonged drought between 1999 and 2002, especially in the Susquehanna River basin where annual discharges were at contemporary minima for four years.[15,58] During this period, chlorophyll-a concentrations declined and DIN concentrations increased in cluster 3 (Figure 3), presumably because of the lessened nutrient demand from phytoplankton. This drought period was, however, associated with a resurgence of living resources (i.e., submerged macrophytes) whose restoration is a target of water quality management,[58] a reduction in chlorophyll-a in seaward, open regions of the Chesapeake Bay,[15] and a period of low hypoxic volume in the estuary.[12] Thus, the fact that the natural nutrient reduction experiment of a drought created conditions consistent with those sought by ecosystem restoration efforts, that the drought led to worsened AD in surface waters, and the fact that postdrought increases winter–spring flows led to improved AD in cluster 3, highlight that nutrient inputs may benefit surface water AD in some regions.

https://doi.org/10.1021/acsestwater.0c00307
ACS EST Water 2021, 1, 1665–1675

3.3. Summary and Management Implications. Quantifying direct and indirect effects that putative drivers have on an indicator variable like AD can identify management actions most likely to meet restoration targets. Our analysis reinforces that freshwater inputs—and manageable nutrient and sediment levels they carry—are where most interventions should be aimed. Flows from January to May led to improved AD in all regions except the large open central segments of the Bay, which was also the only part of the Bay where dissolved inorganic nitrogen meaningfully impacted AD. These differences add to a growing recognition of the need for spatially explicit management that considers regional differences in the ways that upstream, low-salinity regions differ from more expansive seaward waters.[23,59−61] Similarly, this study is itself context-dependent in its focus on open water segments of the Bay, which we focused on because they are characterized by a relatively large diversity of species and habitats. In contrast, freshwater inputs have been clearly linked to oxygen depletion in deep water segments due to the combined effects of stratification and nutrient-induced organic matter production and sinking.[44]

The central segments of the Bay in cluster 1 also illustrate how our approach can identify negative water quality consequences before they manifest. There were more negative harmful drivers of AD than positive drivers in cluster 1, more so than in any other cluster. This is shown as the proportion of negative net effects in Figure 5 and suggests that, despite a record of low AD (close to zero) over the time frame we analyzed, these segments of Chesapeake Bay are at risk for future deficits given large changes in those drivers. The dominance of winter−spring river and sediment inputs in cluster 1 thus makes AD in this system vulnerable to predicted future increases in winter−spring spring precipitation and discharge in the Chesapeake Bay watershed.[56]

An important result from this study is that the two methods we attempted to use to discover causal drivers of AD without prior knowledge, using nothing but coincident time series data, were unsuccessful. Methods like CCM and PCMCI have recently garnered attention as ways to use observational data to understand complex systems.[32,62] However, both CCM and PCMCI produced qualitatively refutable interactions (Tables S12−S23), notably failing to consistently identify DO as a cause of AD despite AD being calculated directly from DO. These methods may be more suited to carefully testing individual interactions than reconstructing entire systems, especially highly complex multivariate systems like the Chesapeake Bay. It may also be the case that they infer mathematical relationships, such as those involving derived indicator variables, differently than physical interactions. We were able to rely on expert knowledge of which variables could interact, but the Chesapeake Bay is an unusually well-studied ecosystem. We suggest causal inference research focus on ways of incorporating prior knowledge, as in the construction of Bayesian networks,[63] particularly for cyclic systems where variables can indirectly have an effect on themselves.

Indirect effects are also as important as direct effects in illustrating controls on water quality variables. Consider that river flow and associated suspended sediment loads were the strongest drivers of AD across the entire Bay but exerted comparatively weak effects on the variables they directly caused (AD has only DO as a direct cause, which it is calculated from). Net effects can always be calculated once standardized direct effects are known, and we find that they more usefully explain how complex systems function. Large, highly connected systems may require converging approximations of net effects as the number of paths connecting each pair of variables becomes too large to store in memory.

A final consideration is to emphasize the value of spatially replicated long-term monitoring data for understanding long-term change in ecosystems. The Chesapeake Bay, like most ecosystems, is dynamic and spatially heterogeneous. High-resolution monitoring programs can track these behaviors and benefit the development of mathematical and statistical methods to explain and predict them and, as with AD, can produce useful synthetic indicators of system health. With these data, management decisions benefit from the kinds of exploratory inference we have conducted here to understand the water quality dynamics in the Chesapeake Bay.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acsestwater.0c00307.

Figure of water quality variables and attainment deficit across all four clusters of open-water segments of the Chesapeake Bay (PDF)

Tables of the Chesapeake Bay segments and their cluster association, attainment deficit in three-year assessment periods from 1985−1987 to 2014−2016, matrix of expert consensus interactions between water quality variables in the Chesapeake Bay, cluster-specific matrixes of estimated direct effects, cluster-specific matrixes of *p*-values for estimated direct effects, cluster-specific tables of nonsignificant interactions that were removed, cluster-specific matrixes of estimated net effects, cluster-specific matrixes of multispatial CCM *p*-values, cluster-specific matrixes of PCMCI coefficients, and cluster-specific matrixes of PCMCI *p*-values (ZIP)

■ AUTHOR INFORMATION

Corresponding Authors

Ryan E. Langendorf − *CIRES, University of Colorado, Boulder, Colorado 80309, United States;* orcid.org/0000-0003-1956-0764; Email: ryan.langendorf@colorado.edu

Qian Zhang − *University of Maryland Center for Environmental Science/U.S. Environmental Protection Agency Chesapeake Bay Program, Annapolis, Maryland 21403, United States*; Email: qzhang@umces.edu

Authors

Vyacheslav Lyubchich − *Chesapeake Biological Laboratory, University of Maryland Center for Environmental Science, Solomons, Maryland 20688, United States*

Jeremy M. Testa − *Chesapeake Biological Laboratory, University of Maryland Center for Environmental Science, Solomons, Maryland 20688, United States*

Complete contact information is available at: https://pubs.acs.org/10.1021/acsestwater.0c00307

Author Contributions

All authors contributed equally, and have given approval to the final version of the manuscript.

Funding

R.E.L. and V.L. acknowledge support of the Maryland Sea Grant (NOAA) at the initial stage of the project (grants

#SA075281850-H and SA075281850-U NA18OAR4170070). V.L. has been partially supported by NSF CNS 1739823 and NOAA NA20OAR4310481. J.M.T. acknowledges support from the National Science Foundation (CBET-13606395, CBET-1706416). Q.Z. acknowledges support of the Environmental Protection Agency (CBP Technical Support Grant No. 07-5-230480). The authors would like to thank the US EPA Chesapeake Bay Program, the Maryland Department of Natural Resources, the Virginia Department of Environmental Quality, and the US Geological Survey for collecting the datasets used in this analysis. This is UMCES Contribution Number 6029 and CBL Ref No. [UMCES] CBL 2022-001.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

The authors would like to thank the USEPA Chesapeake Bay Program, the Maryland Department of Natural Resources, the Virginia Department of Environmental Quality, and the US Geological Survey for collecting the datasets used in this analysis. This is UMCES Contribution Number 6029 and CBL Ref No. [UMCES] CBL 2022-001. The authors would also like to thank Matthew G. Burgess for helpful discussions about causal inference.

■ REFERENCES

(1) Carpenter, S. R.; Cole, J. J.; Pace, M. L.; Batt, R.; Brock, W.; Cline, T.; Coloso, J.; Hodgson, J. R.; Kitchell, J. F.; Seekell, D. A.; et al. Early warnings of regime shifts: a whole-ecosystem experiment. *Science* **2011**, *332*, 1079−1082.

(2) Muscolo, A.; Settineri, G.; Attinà, E. Early warning indicators of changes in soil ecosystem functioning. *Ecol. Indic.* **2015**, *48*, 542−549.

(3) Batiuk, R. A.; Breitburg, D. L.; Diaz, R. J.; Cronin, T. M.; Secor, D. H.; Thursby, G. Derivation of habitat-specific dissolved oxygen criteria for Chesapeake Bay and its tidal tributaries. *J. Exp. Mar. Biol. Ecol.* **2009**, *381*, S204−S215.

(4) Majumder, S.; Tamma, K.; Ramaswamy, S.; Guttal, V. Inferring critical thresholds of ecosystem transitions from spatial data. *Ecology* **2019**, *100*, e02722.

(5) Clements, C. F.; McCarthy, M. A.; Blanchard, J. L. Early warning signals of recovery in complex systems. *Nat. Commun.* **2019**, *10*, 1−9.

(6) Chen, S.; O'Dea, E. B.; Drake, J. M.; Epureanu, B. I. Eigenvalues of the covariance matrix as early warning signals for critical transitions in ecological systems. *Sci. Rep.* **2019**, *9*, 1−14.

(7) Kampe, T. U.; Johnson, B. R.; Kuester, M. A.; Keller, M. NEON: the first continental-scale ecological observatory with airborne remote sensing of vegetation canopy biochemistry and structure. *Journal of Applied Remote Sensing* **2010**, *4*, 043510.

(8) Bastin, J.-F.; Finegold, Y.; Garcia, C.; Mollicone, D.; Rezende, M.; Routh, D.; Zohner, C. M.; Crowther, T. W. The global tree restoration potential. *Science* **2019**, *365*, 76−79.

(9) Marvin, D. C.; Koh, L. P.; Lynam, A. J.; Wich, S.; Davies, A. B.; Krishnamurthy, R.; Stokes, E.; Starkey, R.; Asner, G. P. Integrating technologies for scalable ecology and conservation. *Global Ecology and Conservation* **2016**, *7*, 262−275.

(10) Steenweg, R.; Hebblewhite, M.; Kays, R.; Ahumada, J.; Fisher, J. T.; Burton, C.; Townsend, S. E.; Carbone, C.; Rowcliffe, J. M.; Whittington, J.; et al. Scaling-up camera traps: Monitoring the planet's biodiversity with networks of remote sensors. *Frontiers in Ecology and the Environment* **2017**, *15*, 26−34.

(11) Kemp, W. M.; Boynton, W. R.; Adolf, J. E.; Boesch, D. F.; Boicourt, W. C.; Brush, G.; Cornwell, J. C.; Fisher, T. R.; Glibert, P. M.; Hagy, J. D.; et al. Eutrophication of Chesapeake Bay: historical trends and ecological interactions. *Mar. Ecol.: Prog. Ser.* **2005**, *303*, 1−29.

(12) Murphy, R. R.; Kemp, W. M.; Ball, W. P. Long-term trends in Chesapeake Bay seasonal hypoxia, stratification, and nutrient loading. *Estuaries Coasts* **2011**, *34*, 1293−1309.

(13) Testa, J. M.; Clark, J. B.; Dennison, W. C.; Donovan, E. C.; Fisher, A. W.; Ni, W.; Parker, M.; Scavia, D.; Spitzer, S. E.; Waldrop, A. M.; et al. Ecological forecasting and the science of hypoxia in Chesapeake Bay. *BioScience* **2017**, *67*, 614−626.

(14) Lefcheck, J. S.; Wilcox, D. J.; Murphy, R. R.; Marion, S. R.; Orth, R. J. Multiple stressors threaten the imperiled coastal foundation species eelgrass (Zostera marina) in Chesapeake Bay, USA. *Global Change Biology* **2017**, *23*, 3474−3483.

(15) Testa, J. M.; Kemp, W. M.; Boynton, W. R. Season-specific trends and linkages of nitrogen and oxygen cycles in Chesapeake Bay. *Limnol. Oceanogr.* **2018**, *63*, 2045−2064.

(16) Zhang, Q.; Murphy, R. R.; Tian, R.; Forsyth, M. K.; Trentacoste, E. M.; Keisman, J.; Tango, P. J. Chesapeake Bay's water quality condition has been recovering: Insights from a multimetric indicator assessment of thirty years of tidal monitoring data. *Sci. Total Environ.* **2018**, *637*, 1617−1625.

(17) Ambient water quality criteria for dissolved oxygen, water clarity and chlorophyll a for the Chesapeake Bay and its tidal tributaries. Region III, Chesapeake Bay Program, Annapolis MD; Region III, Water Protection Division, Philadelphia, PA; US EPA, 2003.

(18) Cerco, C. F.; Noel, M. R. The 2017 Chesapeake Bay Water Quality and Sediment Transport Model A Report to the US Environmental Protection Agency Chesapeake Bay Program. *US Army Corps of Engineers*; US Army Engineer Research and Development Center: Vicksburg MS, 2017; pp 1−108.

(19) Chesapeake Bay total maximum daily load for nitrogen, phosphorus and sediment; US Environmental Protection Agency Region, 2010; Vol. 3, pp 1−366.

(20) Zhang, Q.; Tango, P. J.; Murphy, R. R.; Forsyth, M. K.; Tian, R.; Keisman, J.; Trentacoste, E. M. Chesapeake Bay Dissolved Oxygen Criterion Attainment Deficit: Three Decades of Temporal and Spatial Patterns. *Front. Mar. Sci.* **2018**, *5*, 422.

(21) Tango, P. J.; Batiuk, R. A. Deriving Chesapeake Bay water quality standards. *J. Am. Water Resour. Assoc.* **2013**, *49*, 1007−1024.

(22) Chesapeake Bay Program. Data. https://www.chesapeakebay.net/what/data (accessed 2017-04-30).

(23) Testa, J. M.; Lyubchich, V.; Zhang, Q. Patterns and trends in Secchi disk depth over three decades in the Chesapeake Bay estuarine complex. *Estuaries Coasts* **2019**, *42*, 927−943.

(24) Liang, Z.; Dong, F.; Qian, S. S.; Liu, Y.; Chen, H.; Lu, W. Ecoregional or site-specific lake nutrient criteria? Evidence from ecological fallacy. *Ecol. Indic.* **2020**, *111*, 105989.

(25) Pappas, C.; Fatichi, S.; Rimkus, S.; Burlando, P.; Huber, M. O. The role of local-scale heterogeneities in terrestrial ecosystem modeling. *J. Geophys. Res.: Biogeosci.* **2015**, *120*, 341−360.

(26) Duque, J. C.; Artís, M.; Ramos, R. The ecological fallacy in a time series context: evidence from Spanish regional unemployment rates. *Journal of Geographical Systems* **2006**, *8*, 391−410.

(27) Schaeffer, E. D.; Testa, J. M.; Gel, Y. R.; Lyubchich, V. *On information criteria for dynamic spatio-temporal clustering. Proc. of the 6th Int. Workshop on Climate Informatics* **2016**, 5−8.

(28) Huang, X.; Iliev, I. R.; Lyubchich, V.; Gel, Y. R. Riding down the Bay: Space-time clustering of ecological trends. *Environmetrics* **2018**, *29*, e2455.

(29) Murtagh, F.; Legendre, P. Ward's Hierarchical Agglomerative Clustering Method: Which Algorithms Implement Ward's Criterion? *Journal of Classification* **2014**, *31*, 274−295.

(30) Reddy, C. K.; Vinzamuri, B. *Data Clustering: Algorithms and Applications*; Aggarwal, C. C., Reddy, C. K., Eds.; Chapman and Hall/CRC: New York, NY, 2013; Chapter 4, pp 87−110.

(31) R Core Team. *R: A Language and Environment for Statistical Computing*, version 4.0.2; R Foundation for Statistical Computing: Vienna, Austria, 2020.

(32) Sugihara, G.; May, R.; Ye, H.; Hsieh, C.-h.; Deyle, E.; Fogarty, M.; Munch, S. Detecting causality in complex ecosystems. *Science* **2012**, *338*, 496−500.

(33) Runge, J.; Nowack, P.; Kretschmer, M.; Flaxman, S.; Sejdinovic, D. Detecting and quantifying causal associations in large nonlinear time series datasets. *Science Advances* **2019**, *5*, eaau4996.

(34) Takens, F. *Dynamical systems and turbulence, Warwick 1980*; Springer, 1981; pp 366−381.

(35) Spirtes, P.; Glymour, C. N.; Scheines, R.; Heckerman, D. *Causation, prediction, and search*; MIT Press, 2000.

(36) Schumacker, R. E.; Lomax, R. G. *A Beginner's Guide to Structural Equation Modeling*; 4th ed.; Routledge: New York, 2015.

(37) Rosseel, Y. Lavaan: An R package for structural equation modeling and more. Version 0.5−12 (BETA). *J. Stat. Soft.* **2012**, *48*, 1−36.

(38) Ryberg, K. R.; Blomquist, J. D.; Sprague, L. A.; Sekellick, A. J.; Keisman, J. Modeling drivers of phosphorus loads in Chesapeake Bay tributaries and inferences about long-term change. *Sci. Total Environ.* **2018**, *616*, 1423−1430.

(39) Ryberg, K. R.; Stone, W. W.; Baker, N. T. Causal factors for pesticide trends in streams of the United States: Atrazine and deethylatrazine. *J. Environ. Qual.* **2020**, *49*, 152−162.

(40) Lefcheck, J. S.; Orth, R. J.; Dennison, W. C.; Wilcox, D. J.; Murphy, R. R.; Keisman, J.; Gurbisz, C.; Hannam, M.; Landry, J. B.; Moore, K. A.; Patrick, C. J.; Testa, J. M.; Weller, D. E.; Batiuk, R. A. Long-term nutrient reductions lead to the unprecedented recovery of a temperate coastal region. *Proc. Natl. Acad. Sci. U. S. A.* **2018**, *115*, 3658−3662.

(41) Bollen, K. A.; Pearl, J. *Handbook of Causal Analysis for Social Research*; Springer, 2013; Chapter 15, pp 301−328.

(42) Thébault, E.; Fontaine, C. Stability of ecological communities and the architecture of mutualistic and trophic networks. *Science* **2010**, *329*, 853−856.

(43) Csardi, G.; Nepusz, T.; et al. The igraph software package for complex network research. *InterJournal, complex systems* **2006**, *1695*, 1−9.

(44) Hagy, J. D.; Boynton, W. R.; Keefe, C. W.; Wood, K. V. Hypoxia in Chesapeake Bay, 1950−2001: long-term change in relation to nutrient loading and river flow. *Estuaries* **2004**, *27*, 634−658.

(45) Zhang, Q.; Brady, D. C.; Boynton, W. R.; Ball, W. P. Long-term trends of nutrients and sediment from the nontidal Chesapeake watershed: An assessment of progress by river and season. *J. Am. Water Resour. Assoc.* **2015**, *51*, 1534−1555.

(46) Miller, W. D.; Harding, L. W., Jr. Climate forcing of the spring bloom in Chesapeake Bay. *Mar. Ecol.: Prog. Ser.* **2007**, *331*, 11−22.

(47) Harding, L. W., Jr.; Mallonee, M. E.; Perry, E. S.; Miller, W. D.; Adolf, J. E.; Gallegos, C. L.; Paerl, H. W. Variable climatic conditions dominate recent phytoplankton dynamics in Chesapeake Bay. *Sci. Rep.* **2016**, *6*, 1−16.

(48) Palinkas, C. M.; Testa, J. M.; Cornwell, J. C.; Li, M.; Sanford, L. P. Influences of a river dam on delivery and fate of sediments and particulate nutrients to the adjacent estuary: case study of Conowingo dam and Chesapeake Bay. *Estuaries Coasts* **2019**, *42*, 2072−2095.

(49) Dijkstra, Y. M.; Chant, R. J.; Reinfelder, J. R. Factors controlling seasonal phytoplankton dynamics in the Delaware River Estuary: an idealized model study. *Estuaries Coasts* **2019**, *42*, 1839−1857.

(50) Shen, J.; Qin, Q.; Wang, Y.; Sisson, M. A data-driven modeling approach for simulating algal blooms in the tidal freshwater of James River in response to riverine nutrient loading. *Ecol. Modell.* **2019**, *398*, 44−54.

(51) Mulholland, M. R.; Morse, R. E.; Boneillo, G. E.; Bernhardt, P. W.; Filippino, K. C.; Procise, L. A.; Blanco-Garcia, J. L.; Marshall, H. G.; Egerton, T. A.; Hunley, W. S.; et al. Understanding causes and impacts of the dinoflagellate, Cochlodinium polykrikoides, blooms in the Chesapeake Bay. *Estuaries Coasts* **2009**, *32*, 734−747.

(52) Boynton, W. R.; Hagy, J. D.; Cornwell, J. C.; Kemp, W. M.; Greene, S. M.; Owens, M. S.; Baker, J. E.; Larsen, R. K. Nutrient budgets and management actions in the Patuxent River Estuary, Maryland. *Estuaries Coasts* **2008**, *31*, 623−651.

(53) Testa, J.; Kemp, W. M. *Treatise on Estuarine and Coastal Science*; Wolanski, E., McLusky, D. S., Eds.; Academic Press: Waltham, 2011; pp 163−199.

(54) Ni, W.; Li, M.; Testa, J. M. Discerning effects of warming, sea level rise and nutrient management on long-term hypoxia trends in Chesapeake Bay. *Sci. Total Environ.* **2020**, *737*, 139717.

(55) Ding, H.; Elmore, A. J. Spatio-temporal patterns in water surface temperature from Landsat time series data in the Chesapeake Bay, USA. *Remote Sensing of Environment* **2015**, *168*, 335−348.

(56) Ni, W.; Li, M.; Ross, A. C.; Najjar, R. G. Large projected decline in dissolved oxygen in a eutrophic estuary due to climate change. *J. Geophys. Res.: Oceans* **2019**, *124*, 8271−8289.

(57) Nixon, S. W. Coastal marine eutrophication: a definition, social causes, and future concerns. *Ophelia* **1995**, *41*, 199−219.

(58) Gurbisz, C.; Kemp, W. M. Unexpected resurgence of a large submersed plant bed in Chesapeake Bay: Analysis of time series data. *Limnol. Oceanogr.* **2014**, *59*, 482−494.

(59) Wang, P.; Linker, L. C.; Shenk, G. W. Using geographically isolated loading scenarios to analyze nitrogen and phosphorus exchanges and explore tailored nutrient control strategies for efficient management. *Environmental Modeling & Assessment* **2016**, *21*, 437−454.

(60) Wainger, L.; Yu, H.; Gazenski, K.; Boynton, W. The relative influence of local and regional environmental drivers of algal biomass (chlorophyll-a) varies by estuarine location. *Estuarine, Coastal Shelf Sci.* **2016**, *178*, 65−76.

(61) Zhang, Q.; Fisher, T. R.; Trentacoste, E. M.; Buchanan, C.; Gustafson, A. B.; Karrh, R.; Murphy, R. R.; Keisman, J.; Wu, C.; Tian, R.; et al. Nutrient limitation of phytoplankton in Chesapeake Bay: Development of an empirical approach for water-quality management. *Water Res.* **2021**, *188*, 116407.

(62) Runge, J.; Bathiany, S.; Bollt, E.; Camps-Valls, G.; Coumou, D.; Deyle, E.; Glymour, C.; Kretschmer, M.; Mahecha, M. D.; Muñoz-Marí, J.; et al. Inferring causation from time series in Earth system sciences. *Nat. Commun.* **2019**, *10*, 1−13.

(63) Tulupyev, A. L.; Nikolenko, S. I. Directed cycles in Bayesian belief networks: probabilistic semantics and consistency checking complexity. *Artificial Intelligence.* **2005**, *3789*, 214−223.

https://doi.org/10.1021/acsestwater.0c00307
ACS EST Water 2021, 1, 1665−1675