Remotely Sensed Imagery Reveals Animal Feeding Operations Increase Surface Waterbody Concentrations of Dissolved Reactive Phosphorus*

Andrew Meyer [a]
Zach Raff [b]
Sarah Porter [c]

April 25, 2022

Abstract

Recently, the eutrophication of surface waterbodies in the US has led to increased and more intense harmful algal blooms. During the same time, animal agriculture in the US has changed to consist of larger and more concentrated operations. However, the contribution of more intense animal agriculture to surface waterbody eutrophication is unclear because data on animal feeding operations are sparse. Here, we use remotely sensed imagery to identify the location and size of animal feeding operations in the Maumee River Watershed, a key drainage area to Lake Erie's Western Basin, which has recently experienced severe harmful algal blooms. We then estimate the relationship between the intensity of animal feeding operations in the watershed and surface waterbody concentrations of dissolved reactive phosphorus, the pollutant most responsible for algal growth. We find that stream reaches with more upstream livestock exposure contain significantly higher concentrations of dissolved reactive phosphorus than stream reaches with less livestock exposure. Depending on livestock intensity measure, the average marginal upstream animal feeding operation in the watershed increases downstream dissolved reactive phosphorus concentrations by between 10 and 15%. Our work presents evidence that the increasing intensity of animal agriculture in the US contributes to water quality problems. Further, the geolocating of animal feeding operations is important for managing runoff and correctly identifying the causes of surface water quality degradation.

Keywords: aerial imagery; animal feeding operations; dissolved reactive phosphorus; Maumee River watershed; remote sensing; stream and river network

* *Acknowledgements*: We thank Ethan Bahe for help securing and processing the remotely sensed imagery data and we thank Adam Heisey for research assistance on the project. Raff and Meyer acknowledge funding from the Lake Erie Waterkeeper, which helped support this project. All remaining errors are ours.

[a] Marquette University, 1225 W. Wisconsin Ave., Milwaukee, WI 53233, (414) 288-5489, andrew.g.meyer@marquette.edu.

[b] University of Wisconsin-Stout, 712 Broadway Street S, Menomonie, WI 54751, (715) 232-2493, raffz@uwstout.edu.

[c] Environmental Working Group, Minneapolis, MN 55401, saporter1129@gmail.com.

1. Introduction

Anthropogenic eutrophication is recognized as a major cause of impairment in freshwater lakes and reservoirs worldwide, causing predictable increases in algae biomass and harmful algal blooms (Smith, 2002; Heisler et al., 2008). Eutrophication and its associated algal blooms reduce property values, degrade recreational usages, and harm drinking water supplies, imposing billions of dollars of annual economic damages on residents of the US (Dodds et al., 2009). It is well established that phosphorus is the nutrient that constrains eutrophication and the associated algal blooms in freshwater lakes (Carpenter, 2008; Schindler, 1974; Schindler et al., 2008). In many regions of the US, agricultural nonpoint source pollution is the primary source of excess phosphorus (Carpenter et al., 1998; Mooney, 2020). Thus, there is great interest in better understanding the relationship between nonpoint source contributors and downstream phosphorus concentrations and loads.

To date, commercial fertilizer has been the primary focus of research on agricultural nonpoint source phosphorus pollution. However, researchers are beginning to recognize animal feeding operations (AFOs) as another important nonpoint source contributor to phosphorus loads in many watersheds. One limitation of the existing literature estimating the source contribution to phosphorus loads is that there is a lack of reliable, publicly available information about where and how many AFOs exist in many watersheds and the amount of manure and phosphorus that they produce. In this paper, we fill this knowledge gap using remotely sensed imagery to identify the effects of increased animal agriculture on surface water quality in a largely agricultural watershed in the Midwestern US, the Maumee River Watershed (hereafter MRW). The MRW is important because the Maumee River is recognized as the largest source of phosphorus loading to both Lake Erie overall and its Western Basin (Maccoux et al., 2016). We leverage remotely sensed imagery to determine the size and scope of livestock operations in the MRW, obviating the need for permit

information. We then combine this information about livestock operations with publicly available water quality data at downstream locations. We focus on dissolved reactive phosphorus (DRP) because it is the most bioavailable form of phosphorus and previous research recognizes its importance in explaining eutrophication and harmful algal blooms (Baker et al., 2014; Bertani et al., 2016; De Pinto et al., 1986; Richards et al., 2010; Stumpf et al., 2016).

Lake Erie has suffered from eutrophication since at least the 1960s and provides an illustrative case study of the effects of excess phosphorus. Several international agreements and regulations in the 1970s primarily targeted point sources of phosphorus and resulted in water quality improvements that lasted until around 1990 (Baker and Richards, 2002; Paytan et al., 2017). However, harmful algal blooms in Lake Erie began appearing in the mid-1990s and have increased in severity over time (D'Anglada et al., 2018; Michalak et al., 2013; Scavia et al., 2014; Watson et al., 2016). Algal blooms impose millions of dollars of economic costs each year on residents through lost recreational opportunities and tourism revenue, decreased housing values, and from degraded drinking water supplies (Wolf and Klaiber, 2017; Wolf et al., 2022). The blooms are caused by excess phosphorus, primarily dissolved phosphorus, which upstream tributary watersheds deliver to the lake through streams and rivers. Like many other watersheds in the US and around the globe, nonpoint source agricultural release is recognized as the single largest source of excess phosphorus to the MRW (Dolan and Chopra, 2012; IJC, 2018). As such, many studies assume that manure application from AFOs has remained constant over time (IJC, 2018). But manure nutrients produced by AFOs in Western Ohio have increased considerably over the past several decades (Yang et al., 2016). Livestock operations above a certain size threshold (based on total animal units at the facility) are subject to regulation by government agencies. Many AFOs are below this threshold and therefore do not require Clean Water Act permits that would provide

more detailed information about the operation's location, size, and other characteristics. Additionally, many states, including those in the MRW, do not issue permits to all AFOs that are above the relevant size thresholds because of differences in state level Clean Water Act implementation and state agency interpretation of operations considered as potential dischargers (GAO, 2003; Raff and Meyer, 2022). As a result, researchers and governmental agency officials have had little detailed information about the scope of livestock production in the watershed; we address this gap in extant knowledge.

To develop our contributions, we empirically estimate the aggregate effect of upstream AFOs on downstream DRP concentrations, controlling for permanent differences in DRP concentrations across stream reaches and for unobservable factors common to the MRW that change over time. Intuitively, changes in DRP concentrations at stream reaches that experience little upstream expansion in AFO exposure serve as counterfactual outcomes to observed changes in DRP concentrations at stream reaches that experience more substantial increases in AFO exposure.

We find that upstream AFO exposure significantly increases downstream DRP levels in the MRW. In our sample, a marginal AFO that is within 20 km upstream of a monitoring stream reach increases the concentration of DRP at that downstream location by 13.9%. Moreover, we find that estimated DRP impacts vary with the size of livestock operations; larger operations have more animals and produce more manure. The addition of 1,000 animal units within 20 km upstream of a monitoring stream reach increases DRP levels by 10.2% and the addition of 1,000 tons of manure production from upstream AFOs within 20 km increases downstream DRP concentrations by 2.1%. Therefore, larger operations have a larger effect on downstream DRP concentrations and industry trends to more concentrated and larger operations will likely result in

more substantial nutrient loads in the future. These results are robust to alternative upstream distance measures (10 km and 30 km).

We make three primary contributions to the literature with this paper. First, we provide empirical evidence linking AFO expansion to downstream DRP concentrations. Previous studies assessing water quality and policy primarily rely on ex ante hydrological simulations using models such as the Soil and Water Assessment Tool (SWAT) to predict the impacts of land management changes on nutrient concentrations and loads. Previous research efforts use the SWAT to relate agricultural and landscape practices to water quality specifically in the MRW (e.g., Bosch et al., 2013; Cousino et al., 2015; Kast et al., 2021; Martin et al., 2021; Muenich et al., 2016; Scavia et al., 2017). These models require quantity and locational assumptions about the application of manure and other nutrients and conclusions about linkages in the watershed therefore depend on the modeling assumptions. In contrast, our entirely empirical approach is agnostic about these assumptions and relies instead on comparing changes in DRP concentrations downstream of areas with differential expansion in historical AFO intensity. Second, we demonstrate the usefulness of remotely sensed imagery data for producing a historical record of AFO geolocations and intensity. As mentioned, AFO counts from governmental permit data miss many operations, so remotely sensed AFO counts fill an important knowledge gap in the literature. We are not aware of previous studies linking remotely sensed imagery derived AFO intensity measures to downstream nutrient concentrations; we view this as an important contribution to the general agri-environmental literature beyond the MRW. Third, we leverage the National Hydrography Dataset Plus v2.1 (NHD) to determine the flow networks of streams and rivers in our sample and link each AFO to downstream stream reaches. This methodological approach is a major improvement over previous literature on nonpoint source pollution that relies on less precise spatial measures, such as averages

over a HUC8 hydrologic unit watershed (e.g., Grant and Langpap, 2019; Paudel and Crago, 2020; Raff and Meyer, 2022).

2. Study Area

In this section, we describe our study area. First, we discuss the MRW and the Western Basin of Lake Erie. Then, we discuss animal agriculture and manure management in the basin.

2.1. The MRW and the Western Basin of Lake Erie

The MRW spans over 17,000 square kilometers of Northwest Ohio, Northeast Indiana, and South Central Michigan, making it the largest drainage basin of any river system in the Great Lakes Watershed (Ohio DNR, 2018). For locational context, Figure 1 depicts the watershed on a map of the three states that it spans. The MRW lies primarily in Ohio (73 percent of land area), followed by Indiana (20 percent of land area) and Michigan (7 percent of land area). As shown in Figure 2, there are seven HUC8 subbasins within the MRW. Figure 3 further divides the MRW into HUC12 "subwatersheds" and Figure 4 displays the entire MRW stream network. Land use in the MRW is heavily agricultural, with many AFOs and over 70% coverage with row crop agriculture (Muenech et al., 2016; Ohio EPA, 2010). The watershed is characterized by flat topography and heavy, clayey soils with poor natural drainage. Therefore, most of the cropland is drained by subsurface tile (Ohio EPA, 2010).

The MRW drains into Maumee Bay and then the Western Basin of Lake Erie at Toledo, Ohio. Lake Erie is the shallowest and warmest of the Great Lakes, and of its three main basins, the Western Basin is the shallowest. These characteristics make Lake Erie particularly susceptible to environmental impacts from sediment and nutrient loads. Although trends point to decreasing commercial fertilizer application in the Lake Erie Watershed since the 1990s, DRP loads to the lake continue to rise, and harmful algal blooms continue to increase in frequency and severity.

Leading current hypotheses to explain these increasing DRP loads include legacy phosphorus in the soil, tile drainage, and tillage practices (Ohio EPA, 2010). Relatedly, Baker et al. (2014) find that, while total phosphorus loads to Lake Erie have continued to decline, DRP export from the MRW substantially increased beginning in the 1990s. This increase in bioavailable DRP corresponds to the period of Lake Erie's re-eutrophication, suggesting that researchers and policymakers should focus on addressing DRP concentrations and loads (Baker et al., 2014). Another theory is that changes in the Lake Erie ecosystem, potentially related to invasive species, have altered Lake Erie's internal P cycling (Matisoff and Ciborowski, 2005).

Maccoux et al. (2016) provide an accounting of total phosphorus (TP) and DRP loadings to Lake Erie by basin and source type and demonstrate that nonpoint sources now dominate point sources as contributors to phosphorus loadings. For Lake Erie as a whole, 76% of the TP load stems from tributaries. Most of the load is from nonpoint sources (71% of total Lake Erie TP load) versus point sources. Five percent of Lake Erie's TP comes from point sources upstream of tributary monitoring locations and 14% comes from industrial and municipal point sources discharged directly to the lake or downstream of monitoring locations (Maccoux et al., 2016). Furthermore, 60% of the TP loading in Lake Erie comes from the Western Basin (Maccoux et al., 2016). The International Joint Commission (2018) estimates that 80 percent of agricultural phosphorus generated in the Western Lake Erie Basin derives from commercial fertilizer, whereas approximately 20 percent derives from manure from AFOs (IJC, 2018). Kast et al. (2021) estimate that manure sources of phosphorus contribute 8% of the TP load in the watershed, given their modeling assumptions about manure spreading practices.

2.2. Animal Agriculture and Manure Management in the MRW

For our study, it is important to consider how manure from livestock operations likely reaches

surface waterbodies. AFOs use two stages to manage manure. First, AFOs may store manure onsite. Swine and cattle operations most often store manure within surface “lagoons”, whereas poultry operations typically store manure under tarps or in buildings. Second, after storing onsite, nearly all livestock operations spread manure onto farmland, often at agronomically inappropriate rates (Osterberg and Wallinga, 2004). Nonpoint source water pollution, or runoff, can happen in either storage stage. In the first stage, manure lagoons are frequently insecure and do not contain linings or retaining walls (Hribar, 2010). As such, especially during precipitation events, manure can run off from piles and lagoons into surface waterbodies or leach into groundwater (Waller et al., 2021). In the second stage, if manure is overspread on surrounding farmland, the soil cannot fully absorb nutrients from the manure (EPA, 2001; Kellogg et al., 2014). Rain or melting snow can then transport the manure to surface waterbodies. Moreover, artificial drainage tile systems, like those common in the MRW, can quickly transfer excess nutrients directly from farm fields to nearby streams.

Kast et al. (2019) study the manure management plans for all 48 operational CAFOs in the Ohio portion of the MRW. These CAFOs consist of 18 dairy, one beef, 23 swine, and six poultry operations. Among the fields controlled by the CAFOs, the average distance from a swine CAFO’s holding barn to the receiving fields where manure is spread is 1.43 miles. For cattle, this distance is 1.91 miles. Importantly, much of the manure is unaccounted for because of Distribution and Utilization, in which the manure is legally transferred from the CAFO to another entity. Over 75% of CAFO manure phosphorus is transferred in this way, including 100% of all poultry manure phosphorus. Nevertheless, at least for non-poultry manure, other studies find that the manure does not travel far from its source. Lory et al. (2001) and Long et al. (2018) find that most non-poultry manure from CAFOs is applied within five miles of the source.

In addition, regulations vary by state, making consolidation of data across state lines challenging. This study therefore improves upon previous work by using remote sensing to map all AFOs in the MRW between 2010 and 2018. Moreover, we estimate the number of animals housed at these facilities and the amount of manure that they produce.

3. Data

For our empirical analysis, we aggregate data from several sources. In this section, we describe these data sources, the construction of our analysis sample, and provide descriptive analyses and statistical summaries.

3.1. Water Quality

First, we gather water quality data from the publicly available Water Quality Portal (NWQC, 2022). The National Water Quality Council aggregates data from three sources to include in the portal: USGS National Water Information System (NWIS), EPA Storage and Retrieval (STORET), and USGS Biodata. In the MRW, DRP is the water pollutant of interest. We therefore use De Cicco et al. (2018) to scrape the Water Quality Portal and gather all nutrient readings for Ohio, Indiana, and Michigan, which we then filter for DRP readings for the MRW. Few DRP readings are available in the portal before 2010, so we begin our sample in 2010, when data are readily available. Next, we retain only water quality readings from streams and rivers, rather than, e.g., lakes, ditches, to better eliminate any impacts of legacy phosphorus or non-perennial waterbodies on the readings. Finally, like previous studies, we transform zero and non-detected measurements to ½ of the smallest positive value in the sample (Keiser and Shapiro 2019; Chen et al. 2019; Raff and Meyer, 2022). Then, we complement data from the Water Quality Portal with long term monitoring data from the Heidelberg Tributary Loading Program (HTLP). Heidelberg University's National Center for Water Quality Research (NCWQR) initiated the HTLP in 1974

as part of state and federal programs to improve water quality in Lake Erie. The NCWQR collects hundreds of water samples from each monitored location each year, analyzes the water samples to calculate concentrations of multiple pollutants, and publicly shares these data via the HTLP data portal (HTLP, 2022). There are nine HTLP monitoring locations in the MRW; of these, eight provide DRP readings for at least one year of our sample period. We use all available DRP measurements from the HTLP data portal. The HTLP reports DRP measurements for the Maumee Station dating back to 1975, for three other locations (Blanchard River, Tiffin River, and Lost Creek) beginning in 2007, and for four other upstream tributary locations in 2018. As the longest-running and most detailed program of its kind in the United States (NCWQR, 2022), these data help establish longer terms trends in DRP concentrations in the MRW.

There could be a concern that governmental agencies and citizen volunteers are more likely to sample water quality on days that they expect nutrient concentrations to be elevated, particularly following precipitation events. This potential bias could cause a threat to our empirical analysis if the probability of sampling following a precipitation event increases differentially at stream reaches with more upstream AFOs. Raff and Meyer (2022) provide evidence from Wisconsin that sampling timing is not endogenous to precipitation and that sampling does not differentially increase in areas with more CAFOs following precipitation events. Nevertheless, to avoid potentially oversampling certain locations that were monitored more frequently and to smooth daily noise in the data, we follow other studies and aggregate surface water quality readings to the monthly level (Kesier and Shapiro, 2019). Here, we average all DRP readings along each stream reach. Our measure of water quality is therefore an average reading at the stream reach-month level, from 2010-2018. (Regardless, our empirical results are robust to estimation using daily water quality readings.)

3.2. Animal Agriculture

Next, we obtain data on the location, type (i.e., animal), and size of livestock operations in the MRW from 2010 to 2018 using satellite imagery. Few livestock operations in the watershed have Clean Water Act permits (GAO, 2003; Kast et al. 2019), so it is difficult to identify their locations and the overall intensity of animal agriculture in the watershed. We overcome this lack of data by using National Agriculture Imagery Program (NAIP) aerial photography to visually locate AFOs in the MRW. Consistent imagery was available across the study area beginning in 2005, so we match on available DRP data and begin our study in 2010. We supplement NAIP data with planet satellite imagery to capture recent AFO construction (late 2018 to January 2019). For each AFO, we attribute several characteristics of the operation: the number of barns and their total square footage, animal type (poultry, swine, beef cattle or dairy cattle) and the year of expansion, if any. Figures 2-4 depict the location of the AFOs in the watershed as of 2018, in relation to their position within the stream and river network.

When available, we assign animal type using permit data for each operation. When unavailable, we assign animal type to each operation based on several attributes unique to each facility, including the size and shape of each barn, the presence and number of feed bins, the location of fans, and the presence of lagoons and of visually identifiable animals. We also use Google Street View to ground truth the assignments. In addition, we remove facilities from the analysis if they appear abandoned, as evidenced by dilapidated roofs or removal of infrastructure.

We obtain state permit data for AFOs in the MRW from various sources. We obtain CAFO permit data from the Ohio Department of Agriculture (Ohio AFOs), MIWaters website (Michigan AFOs), and the Indiana Department of Environmental Management (Indiana AFOs). Where possible, we match these permit data to mapped facilities.

Finally, we estimate animal counts and manure production for each facility by dividing the mapped square footage of each barn by a square footage per animal. We obtain recommended square footage per animal from a literature review of standards, which, along with their source, we provide in Supplementary Table S3. We estimate animal counts for each barn in the MRW and provide total counts in Supplementary Table S4. For permitted facilities, we use the permitted animal count rather than the estimates.

3.3. Surface Waterbody Network

We gather data from the Watershed Boundary Dataset v2.3 (WBD) and the NHD to determine watershed boundaries for each AFO and stream reach and to determine the flow networks of streams and rivers in our sample, respectively. The WBD and NHD are national geospatial surface water frameworks that EPA and USGS developed through a collaborative partnership. The datasets contain information on stream and river flow paths, watershed and catchment boundaries, and other important locations in the US hydrological network, such as USGS gage stations and catchment outlets. The WBD allows us to place each stream reach into its appropriate watershed, which we then use to develop variables that capture the characteristics of the area surrounding AFOs and stream reaches. Here, we use the HUC12 “subwatershed” level. More importantly, the NHD provides information about the US flow network at the stream reach level for all perennial streams and rivers in the contiguous US. Figure 4 depicts the NHD flow network for the MRW. A stream reach typically consists of a one to five km section of the waterbody, depending on if it is part of the mainstem or a tributary. The NHD provides information on, among other things, the type of waterbody, its flow direction, its outlet, the type of stream reach (e.g., mainstem v. tributary), and flow. From this stream and river flow network, we can navigate upstream and downstream from AFOs and DRP monitoring stream reaches.

3.4. Control Variables

In our empirical analysis, we control for time varying characteristics of surface waterbodies and watersheds that, if omitted and correlated with the treatment regressors, may bias our estimates. Inclusion of covariates can also reduce the variance of the error term and hence improve precision in our estimates. First, we control for precipitation in all specifications because it is well known that precipitation events substantially affect nutrient runoff and observed concentrations (Waller et al., 2021). We gather precipitation data from Schlenker and Roberts (2020), who use PRISM climate data and weather monitoring stations to create daily precipitation data for 2.5 by 2.5-mile grids throughout the conterminous US. We geocode each stream reach to the nearest PRISM grid centroid to get the average daily precipitation (cm) at each stream reach in the MRW for each month of our sample. Next, in some specifications we control for land cover and fertilizer usage. We gather land cover data from the National Land Cover Database (NLCD). The NLCD classifies land cover at 30 m resolutions and we overlay the NLCD maps onto HUC12 subwatersheds and calculate the percentage of each HUC12 that is developed and planted. Finally, the literature suggests that runoff from commercial fertilizer is a source of phosphorus in surface waterbodies in the MRW. We therefore control for commercial fertilizer application of phosphorus, both on- and off-farm. We gather these county level data from Falcone (2020) and convert them to HUC12 level measures using GIS area weights. For land cover and fertilizer usage, which USDA and the NLDC do not collect monthly, we interpolate values to the monthly level.

3.5. Construction of Analysis Sample

To construct our analysis sample, we must first match each stream reach with DRP concentration readings to its level of upstream AFO exposure over time. We derive an AFO intensity measure that is more spatially fine than an aggregate watershed approach (e.g., Raff and Meyer, 2022) by

using GIS to spatially match each water quality monitoring location and AFO to the nearest stream reach in the NHD. We then identify all monitoring stream reaches that are within 10, 20, and 30 km downstream (via the stream and river network) from the stream reach nearest each AFO. For each stream reach, we aggregate the number of upstream AFOs within these distances and their associated animal counts (in au) and their estimated manure production (in tons). We therefore have three treatment regressors at the stream reach-month level, from 2010-2018, within three distance measures: 1) number of upstream operating AFOs, 2) number of upstream animal units, and 3) amount of upstream manure production.

We complete our analysis sample by adding time varying control variables. We match monthly average precipitation data from the nearest PRISM grid to each stream reach. Then, we add HUC12-month level land use and commercial fertilizer application controls. Our final analysis sample is a 1,673 observation stream reach-month level panel dataset from 2010-2018.

3.6. Descriptive Analysis

In this sub-section, we discuss data trends, spatial relationships between our primary variables of interest, and statistical summaries. First, we explore how the intensity of animal agriculture has changed during our sample period. By the end of our sample period, we identify, with location information, 774 AFOs in the MRW; we map these operations in Figures 2-4. Of the operations in our sample, there are 298 swine, 183 cattle, 118 dairy, and 175 poultry facilities. Of the AFOs in the watershed in 2018, 27 percent expanded since their first year of construction, which means that the operations either added more buildings or increased the size of existing buildings. During our sample period, the mean barn size and the number of animals per facility increased for dairy, poultry, and swine operations, which suggests that barns at AFOs are growing and that the operations are housing more animals at each facility. Our results align with those from IJC (2018)

that show increased consolidation of livestock agriculture in the Western Lake Erie Basin.

Next, we examine the regulatory environment surrounding animal agriculture in the MRW by looking at permit data. Indiana, Ohio, and Michigan have their own requirements for livestock operations to obtain a permit, based mainly on the number of animals at each operation and the state's interpretation of potential dischargers. Overall, only 20 percent of AFOs that we identify in the MRW via remotely sensed imagery were permitted at the end of our sample period. Swine and dairy operations were permitted at the highest frequency, with 32 percent and 31 percent of facilities permitted, respectively. Only 12 percent of poultry and 2 percent of cattle facilities had permits in 2018. The overall low percentage of permitted facilities highlights the importance of identifying AFOs by means outside of official permits.

Third, we provide a series of maps of the MRW to visualize spatial variation in our main variables of interest. In each map (except Figure 9), we use the Jenks optimization method to cluster variable levels into classes. These natural breaks minimize the variation within each class while maximizing variation across the classes. Figure 5 shows the number of AFOs in each HUC12 region in 2018. Likewise, Figure 6 shows the number of animal units in each HUC12 region in 2018. These figures reveal considerable variation in the intensity of AFO presence across the MRW. The western HUC8 subbasins have the relative highest AFO intensity within the Maumee; this includes the St. Mary's, Auglaize, Upper Maumee, and St. Joseph's subbasins. Two other subbasins (Tiffin and Blanchard) also have substantial AFO presence, but the intensity is somewhat lower in comparison with the subbasins to the west. The Lower Maumee subbasin has the comparatively lowest intensity of AFOs. Figure 7 then shows how AFO presence interfaces with the stream and river network in the MRW. In the figure, we show the number of animal units that are within 20 km upstream of each stream reach; thicker blue lines represent stream reaches

downstream of higher intensities of AFOs. Stream reaches in the St. Mary's, Auglaize, and Upper Maumee subbasins are exposed to the highest number of upstream animal units. Some stream reaches within the St. Joseph's and Tiffin subbasins also display substantial exposure to upstream animal units.

Figures 8, 9, and 10 visualize average DRP concentrations for each HUC12 in the MRW. In Figure 8, we average all available DRP concentrations from 2017-2019 and use the Jenks optimization method to cluster DRP levels into classes. In general, we see the highest concentrations of DRP in the St. Mary's, Upper Maumee, and Auglaize subbasins. Then, in Figure 9, we compare average DRP levels to thresholds that have been proposed as target concentrations. For example, the Annex 4 Objectives and Targets Task Team (2015) recommends a flow weighted mean target DRP concentration of 0.05 mg/L for the Maumee River. The same report also details average flow weighted mean DRP concentrations of approximately 0.10 mg/L for the Maumee River from 2012-2015. In Figure 9, we see that virtually all HUC12 subwatersheds exceed average concentrations of 0.05 mg/L and most also exceed average concentrations of 0.10 mg/L. Lastly, in Figure 10, we overlay the AFO stream exposure from Figure 7 onto the DRP concentrations from Figure 8. Darker areas, with higher average DRP concentrations, tend to be in the same areas with stream reaches receiving the highest upstream AFO intensity. These maps represent a snapshot from one point in time, and hence do not establish a causal relationship. However, a strong spatial correlation between AFO intensity and average DRP concentrations is evident and motivates our empirical analysis of Section 4.

Finally, we discuss sample summary statistics. Table 1 provides statistical summaries for the observations of the final analysis sample; we focus on the 20 km distances throughout the main text. The mean stream reach-month observation has a DRP concentration of 0.155 mg/L. The mean

observation is exposed to nearly 3.5 AFOs, almost 4,700 total animal units, and 27,000 tons of manure production at AFOs. For our empirical analysis, these intensity measures represent our primary treatment regressors. The ranges of these measures suggest growing intensity and concentration of animal agriculture in the MRW over the past several years. For our sample, the maximum upstream AFO exposure is 50 operations within 20 km upstream of the monitoring stream reach.

To examine if there exists a descriptive relationship between our treatment regressors and dependent variable, we examine the heterogeneity in our outcome measure by upstream AFO exposure. For the final analysis sample, we examine the mean DRP concentration for those stream reaches with zero upstream AFO exposure and those with at least one AFO within 20 km upstream (the sample median). We tabulate the results for this analysis in Table 2. The mean DRP concentration for stream reaches without any upstream AFO exposure is 0.120 mg/L, while the mean level for stream reaches with at least one upstream AFO is 0.174 mg/L; this univariate difference of means is statistically significant (p=0.000). This difference in means is suggestive of the effect of upstream AFO exposure on DRP concentrations. However, time invariant differences across stream reaches and trends in the environment may create the difference. We address these potential confounding factors with our empirical approach in the following section.

4. Empirical Approach

To empirically analyze the effects of AFO exposure and intensity on surface water quality in the MRW, we use regression analysis to identify the relationship between the treatment regressors described above and stream reach-month level DRP concentrations. This approach allows us to model the aggregate effect of upstream AFOs on downstream DRP concentrations. Our approach facilitates the inclusion of treatment variables with varying intensities (number of operating AFOs,

animal counts, or manure production) which is important in the MRW where there are closely clustered AFOs throughout the watershed. We estimate the following regression specification:

$$ln(P_{jhmt}) = \beta_1 AFO_{jhmt} + \boldsymbol{P'_{jmt}}\beta_2 + \boldsymbol{X'_{hmt}}\beta_3 + \gamma_j + \psi_m + \lambda_t + \varepsilon_{jhmt}, \quad (1)$$

where P_{jhmt} is the mean DRP concentration (in mg/L) of stream reach *j* in HUC12 *h* in month *m* of year *t*. Here, we log-transform the outcome for two reasons. First, the outcome is leftward-skewed, so log-transforming normalizes its distribution. Second, log-transforming the outcome allows for interpretation of the primary treatment regressors in percentage terms (rather than in level terms), which is more appropriate for variables of interest with wide distributions. Next, we code AFO_{jhmt} as the count of upstream: 1) operating AFOs, 2) number of animal units at AFOs, and 3) amount of manure production (tons) at AFOs; that are present on the month of the mean DRP readings. In the main text, we focus on upstream intensity measures for operations within 20 km of the monitoring stream reach via the stream and river network. In the supplementary material, we provide results for AFO exposure within 10 km and 30 km of the monitoring stream reach. Importantly, equation (1) contains a series of fixed effects. γ_j are stream reach fixed effects that control for time invariant stream reach level characteristics that correlate with DRP concentrations. There are permanent differences in average DRP concentrations across stream reaches because of land slope, soil type, or other geographic/hydrologic characteristics near or at each stream reach. By including stream reach fixed effects in our empirical specification, we use a "within" estimator that identifies the effects of our treatment regressors using only within stream reach variation in the independent variables of interest. These fixed effects therefore control for the permanent differences in DRP concentrations at different stream reaches across time. ψ_m are month fixed effects that control for the seasonality of DRP concentrations and λ_t are year fixed effects that control for unobservable factors common to the entire watershed that change over time, e.g.,

regional or national policies to control water pollution. During our sample period, there were several policy changes that could possibly affect surface water quality in the MRW, including changes in the permitting requirements for AFOs. The year fixed effects absorb the common changes in nutrient concentrations due to these policies.

Next, $\boldsymbol{P_{jmt}}$ and $\boldsymbol{X_{hmt}}$ are vectors of time varying controls at the stream reach- and HUC12-month level, respectively. Inclusion of these controls can reduce the variance of the error term and improve the precision of our estimates. First, $\boldsymbol{P_{jmt}}$ controls for precipitation. Rain and snow events affect how much nonpoint source runoff occurs. Dilution occurs during the runoff process, so we control for both the mean of daily total precipitation and its square. Second, $\boldsymbol{X_{hmt}}$ is a vector that contains measures for land cover and land use. $\boldsymbol{X_{hmt}}$ includes NLCD land cover classifications in percentages, such as the percent of the HUC12 that is planted. These land cover measures control for the likelihood that other, non-AFO, nonpoint source pollution occurs in each HUC12. These measures also capture the likelihood of urban runoff, which can affect nutrient concentrations in surface waterbodies. Alternatively, more developed land may be associated with decreased nonpoint source nutrient runoff if there is no agriculture in the region. Certain land uses can also serve as a phosphorus "sink", such as wetlands, which would decrease the amount of runoff phosphorus that ultimately ends up in surface waterbodies. $\boldsymbol{X_{hmt}}$ contains measures for the percentage of land in each HUC12 that is planted, developed, or wetlands. Finally, commercial fertilizer contributes to phosphorus runoff in the MRW. We therefore include within $\boldsymbol{X_{hmt}}$ a control measure for the amount of phosphorus added to the land through commercial fertilizer application, both on- and off-farm, on a per acre basis (Falcone, 2020). ε_{jhmt} is the exogenous error term. We cluster standard errors at the stream reach level, which is our level of identifying variation, to allow for within stream reach correlation in the error term (spatial and serial

correlation).

Our approach represents a two-way fixed effects specification, where identification of our coefficient of interest, β_1, comes from changes in DRP concentrations within a stream reach coincident with variation in upstream AFO presence and intensity. Our identifying assumption is the standard parallel trends assumption for two-way fixed effects regression. Conditional on covariates, and in the absence of any changes to upstream AFO presence and intensity, average downstream phosphorus concentrations would have trended in parallel at monitoring locations with varying intensities of upstream AFO expansion.

5. Results

In this section, we provide the estimation results. Table 3 presents results for the estimation of equation (1), which examines the effects of upstream AFO exposure on downstream DRP levels in streams and rivers that monitor DRP levels in the MRW. In Table 3, we tabulate results for several regression specifications to assess their robustness to model specification and treatment definition. We also present these results for several definitions of treatment to examine heterogeneity in the results by the characteristics of each operation, such as its size. Panel A presents results for the treatment variables and Panel B presents results for the control variables. Columns 1, 3, and 5 present results for the parsimonious regression specification, where we include only our measures of treatment and precipitation on the righthand side, while columns 2, 4, and 6 add HUC12 level time varying controls to add precision to our estimates. Below, we focus our discussion on the marginal effects of the specification with the full set of controls but note that coefficient point estimates are qualitatively similar in specifications without the additional controls.

In our sample, the marginal AFO that is within 20 km of a monitoring stream reach

increases the concentration of DRP at that downstream location by 13.9%. (Because we log-transform the outcome, we interpret each coefficient as an $e^{\widehat{\beta}} - 1$ percent change in the outcome.) Here, the marginal effect represents the impact of any livestock operation, regardless of size, animal type, or other characteristics of the AFO, on downstream DRP concentrations. Because larger operations produce more manure in a concentrated location, it is likely that there exists heterogeneity in the effects of AFO exposure on downstream DRP levels. We therefore examine the effects on our outcome of interest for two other measures of treatment that better reflect the differences in the size and makeup of AFOs in the watershed. First, we examine the effect of animal unit exposure on downstream DRP levels. The results that we present in column 2 of Table 3 assume a uniform, average effect of the marginal AFO on downstream DRP levels. We now allow this effect to change based on the size of the operation, which we measure in animal units. The fourth column of Table 3 shows that the addition of 1,000 animal units within 20 km upstream of a stream reach that monitors DRP increases these levels by 10.2%. To put this value into greater perspective, in the last year of our sample period, the average AFO in the MRW contained 1,061 animal units. Our results therefore suggest that the addition of an average size AFO, which here, EPA considers a CAFO, increases downstream DRP concentrations by nearly 11%. Second, we examine as a treatment regressor the amount of manure that each AFO produces onsite. As mentioned, nearly all AFOs, both in the MRW and throughout the US, spread their manure onto nearby agricultural land, most often within five miles of the operation. Therefore, the livestock operations that produce more manure can have a larger effect on downstream DRP concentrations, through greater contributions to runoff, than AFOs that produce less manure. Using onsite manure production as a treatment regressor allows us to again examine the differences in effects by the size of the AFO, but also by the type of operation, because different livestock types produce

different amounts of manure each day. As a specific example, dairy cows produce 120 pounds of manure per day, while swine produce 10 pounds of manure per day. Therefore, AFOs with similar animal counts can produce different amounts of manure, which differentially impacts runoff pollution and therefore water quality. Column 6 of Table 3 shows that for every 1,000 tons of manure production from upstream AFOs, downstream DRP concentrations increase by roughly 2.1%. Again using an average AFO to put this effect into context, the average AFO in our sample produced 7,152 tons of manure in 2018. Therefore, the addition of one AFO, that produces the sample average amount of manure onsite, to the watershed results in an increase of downstream DRP levels of over 15%. (Tables S1 and S2, which contain estimation results using alternative distance measures, present results that are qualitatively and quantitatively like the primary estimation results presented in the main text.)

We next examine the results for the control variables, which we present in Panel B of Table 3. Because these measures serve only as controls, we do not interpret the estimates to represent causal effects. However, the signs of the control variable coefficients and their stability throughout all regression specifications lend support to the internal validity of our study. As expected, precipitation is positively correlated with DRP concentrations in the MRW while its square is negatively correlated; both signs reinforce the impact of precipitation on runoff and increasing pollutant concentrations in surface waterbodies, and that dilution occurs during the runoff process. The land use measures also have the expected signs, although they are imprecisely estimated. Again, as expected, more developed land is associated with higher DRP concentrations, while more area in the HUC12 containing wetlands is associated with lower DRP concentrations. The coefficients on planted land use are zero. For commercial fertilizer usage, estimation results confirm the hypothesized positive relationship between fertilizer application and DRP

concentrations. Also consistent with previous studies in the MRW, the estimated coefficients suggest that commercial fertilizer plays a larger role in DRP concentrations than animal agriculture.

Our collective estimation results provide evidence that the increasing concentration and intensity of animal agriculture in the MRW significantly contributes to surface water quality degradation in the watershed, particularly through increased concentrations of DRP in streams and rivers. In addition, our empirical results suggest that there is heterogeneity in these effects depending on the size of the operation. Although we find an average effect for the marginal operation in the watershed, we show that larger operations – both those with a higher number of animal units and those that produce more manure – have a larger effect on downstream DRP concentrations than smaller operations. As a result, the trend in the livestock industry to more concentrated and larger operations is likely to result in even more sizable water quality impacts in the future.

6. Conclusion

Lake Erie, which is the shallowest and warmest of the Great Lakes, has long struggled with eutrophication. More recently, Lake Erie, and its Western Basin in particular, has suffered from the increasing prevalence and severity of harmful algal blooms. Previous work attributes these blooms to nonpoint source agricultural pollution, primarily from commercial fertilizer usage near upstream tributaries that drain to the lake, most of which are in the MRW. But during this same time, animal agriculture in the US and the MRW has become more intensified, with more concentrated and larger operations dominating the landscape. These operations, which produce, store, and spread nearby vast quantities of manure, are not regulated like point source wastewater dischargers. The potential for runoff pollution to occur from these operations is therefore high.

However, because of the lack of available permit data on AFOs (only about 20% of AFOs in the MRW are permitted), it is difficult to identify the effects of animal agriculture on phosphorus loadings to Lake Erie and its Western Basin. Previous work primarily uses ex ante hydrological modeling to study the causes of nutrient loadings in the watershed, while holding constant over time the contribution of animal agriculture and manure.

We overcome the dearth of information on AFOs in the MRW and empirically study this question by using remotely sensed imagery to identify the location, size, and type of livestock operations in the watershed. Our approach allows us to geolocate livestock operations and examine the intensity of animal agriculture in the watershed over time. The primary purpose of this paper is to empirically identify the effects of livestock operations and their expansion on surface water quality in the MRW.

DRP, rather than TP, is the water pollutant most responsible for harmful algal blooms. Our outcome of interest is therefore the concentration of DRP in streams and rivers within the MRW. Unlike previous studies that use a watershed approach, we link AFOs and downstream DRP concentrations using the US hydrological network to examine the effects of interest. Importantly, we use stream reach and time fixed effects, land use measures, and commercial fertilizer application to control for variables that may bias our estimates if omitted and correlated with livestock intensity.

Our estimation results provide evidence that stream reaches with more upstream AFO exposure contain significantly higher concentrations of DRP than counterfactual stream reaches with less AFO exposure. Our results suggest that, depending on the definition of treatment (upstream operating AFOs, animal units at AFOs, or manure production at AFOs), the average AFO in the watershed increases downstream DRP concentrations by between 10 and 15%.

The results of our analysis suggest that animal agriculture in the MRW contributes to the water quality problems in the watershed and to the presence of harmful algal blooms in the Western Lake Erie Basin. Our results therefore present policy implications and motivate additional needed work. First, the lack of permits and local regulations in the watershed likely contribute to the runoff. In some states (e.g., Wisconsin), all AFOs of a certain size must be permitted, but all AFOs (regardless of size) are subject to local regulations, such as the requirement to complete a nutrient management plan. In addition, previous studies on surface water quality degradation in the MRW likely underestimate the contributions of AFOs because of a lack of data on their presence and size. Mooney et al. (2020) show that an analogous knowledge and policy gap related to nonpoint source pollution stems from regulatory and monitoring efforts focusing only on large tributaries. Future studies should expand their data collection outside of official permits to better account for animal agriculture's contributions to nonpoint source pollution and examine all streams and rivers in the watershed. Second, previous work has shown that the external costs to water quality from increasingly concentrated opportunities are the result of the land being unable to assimilate the excess manure phosphorus. Comparing present DRP concentrations to targets proposed by the Annex 4 Objectives and Targets Task Team (2015), it is difficult to see how DRP loading targets could be reached if more AFOs are added to the MRW, unless other, more substantial, nonpoint source reductions in DRP are made to compensate for the additional loads from new or expanded AFOs. Relatedly, we echo the conclusions of Kast et al. (2019) that more knowledge is needed about what happens to manure from livestock operations below CAFO thresholds and from CAFOs engaging in Distribution and Utilization practices. This additional knowledge about where AFOs apply manure could help in designing incentives and policies to decrease its environmental impact. Lastly, as we have shown, most livestock operations in the MRW are near the outside

edges of the watershed, not necessarily near the mouth of the river at Lake Erie. Therefore, future empirical studies should examine the downstream rate of DRP transfer (loads) in addition to concentrations near the operations.

References

Annex 4 Objectives and Targets Task Team. 2015, "Recommended Phosphorus Loading Targets for Lake Erie." Final Report to the Nutrients Annex Subcommittee. Available at: https://www.epa.gov/glwqa/report-recommended-phosphorus-loading-targets-lake-erie.

Baker, D.B., Confesor, R., Ewing, D.E., Johnson, L.T., Kramer, J.W. & Merryfield, B.J. 2014, "Phosphorus loading to Lake Erie from the Maumee, Sandusky and Cuyahoga rivers: The importance of bioavailability", *Journal of Great Lakes Research,* vol. 40, no. 3, pp. 502-517.

Baker, D.B. & Richards, R.P. 2002, "Phosphorus budgets and riverine phosphorus export in northwestern Ohio watersheds", *Journal of Environmental Quality,* vol. 31, no. 1, pp. 96-108.

Bertani, I., Obenour, D.R., Steger, C.E., Stow, C.A., Gronewold, A.D. & Scavia, D. 2016, "Probabilistically assessing the role of nutrient loading in harmful algal bloom formation in western Lake Erie", *Journal of Great Lakes Research,* vol. 42, no. 6, pp. 1184-1192.

Bosch, N.S., Allan, J.D., Selegean, J.P. & Scavia, D. 2013, "Scenario-testing of agricultural best management practices in Lake Erie watersheds", *Journal of Great Lakes Research,* vol. 39, no. 3, pp. 429-436.

Carpenter S., R. 2008, "Phosphorus control is critical to mitigating eutrophication", *Proceedings of the National Academy of Sciences,* vol. 105, no. 32, pp. 11039-11040.

Chen, C., G.E. Lade, J.M. Crespi, D.A. Keiser, 2019, "Size-based regulations and environmental quality: Evidence from the U.S. livestock industry", working paper, available at: https://ageconsearch.umn.edu/record/291262/files/Abstracts_19_05_15_21_01_17_38__129_186_248_69_0.pdf.

D'Anglada, L., Gould, C., Thur, S., Lape, J., Backer, L., Bricker, S., Clyde, T., Davis, T., Dortch, Q., Duriancik, L., Emery, E., Evans, M.A., Fogarty, L., Friona, T., Garrison, D., Graham, J., Handy, S., Johnson, M.-V., Lee, D., Lewitus, A., Litaker, W., Loeffler, C., Lorenzoni, L., Malloy, E.H., Makuch, J., Martinez, E., Meckley, T., Melnick, R., Myers, D., Ramsdell, J., Rohring, E., Rothlisberger, J., Ruberg, S., Ziegler, T. 2018. "Harmful Algal Blooms and Hypoxia in the United States: A Report on Interagency Progress and Implementation." National Oceanic and Atmospheric Administration.

De Cicco, L.A., Hirsch, R.M., Lorenz, D., Watkins, W.D., 2018, dataRetrieval: R packages for discovering and retrieving water data, available from Federal hydrologic web services, doi:10.5066/P9X4L3GE.

De Pinto, J.V., Young, T.C. & McIlroy, L.M. 1986, "Great lakes water quality improvement", *Environmental science & technology,* vol. 20, no. 8, pp. 752-759.

Dodds, W.K., Bouska, W.W., Eitzmann, J.L., Pilger, T.J., Pitts, K.L., Riley, A.J., Schloesser, J.T., and Thornbrugh, D.J. 2009, "Eutrophication of U.S. freshwaters: Analysis of potential economic damages", *Environmental Science and Technology,* vol. 43, no. 1, pp. 12-19.

Dolan, D.M. & Chapra, S.C. 2012, "Great Lakes total phosphorus revisited: 1. Loading analysis and update (1994–2008)", *Journal of Great Lakes Research,* vol. 38, no. 4, pp. 730-740.

EPA, 2001, "Environmental assessment of proposed revisions to the National Pollutant

Discharge Elimination System regulation and the effluent guidelines for Concentrated Animal Feeding Operations", Washington, DC: U.S. Environmental Protection Agency.

Falcone, J.A. 2020, "Estimates of County-Level Nitrogen and Phosphorus from Fertilizer and Manure from 1950 through 2017 in the Conterminous United States," US Geological Survey Open File Report 2020-1153.

GAO, 2003, "Livestock agriculture: Increased EPA oversight will improve environmental program for Concentrated Animal Feeding operations", Washington, DC: U.S. Government Accountability Office.

Grant, L. & Langpap, C. 2019, "Private provision of public goods by environmental groups", *Proceedings of the National Academy of Sciences,* vol. 116, no. 12, pp. 5334-5340.

Heisler, J., Glibert, P.M., Burkholder, J.M., Anderson, D.M., Cochlan, W., Dennison, W.C., Dortch, Q., Gobler, C.J., Heil, C.A., Humphries, E., Lewitus, A., Magnien, R., Marshall, H.G., Sellner, K., Stockwell, D.A., Stoecker, D.K. & Suddleson, M. 2008, "Eutrophication and harmful algal blooms: A scientific consensus", *Harmful Algae,* vol. 8, no. 1, pp. 3-13.

Hribar, C., 2010, "Understanding Concentrated Animal Feeding Operations and their impact on communities", Bowling Green, OH: National Association of Local Boards of Health.

HTLP, 2022, "HTLP Data Portal", https://ncwqr-data.org/.

International Joint Commission (2018). Fertilizer Application Patterns and Trends and Their Implications for Water Quality in the Western Lake Erie Basin.

Kast, J.B., Apostel, A.M., Kalcic, M.M., Muenich, R.L., Dagnew, A., Long, C.M., Evenson, G. & Martin, J.F. 2021, "Source contribution to phosphorus loads from the Maumee River watershed to Lake Erie", *Journal of Environmental Management,* vol. 279, pp. 111803.

Kast, J.B., Long, C.M., Muenich, R.L., Martin, J.F. & Kalcic, M.M. 2019, "Manure Management at Ohio Confined Animal Feeding Facilities in the Maumee River Watershed", *Journal of Great Lakes Research,* vol. 45, no. 6, pp. 1162-1170.

Keiser, D.A. & Shapiro, J.S. 2019, "Consequences of the Clean Water Act and the Demand for Water Quality", *The Quarterly Journal of Economics,* vol. 134, no. 1, pp. 349-396.

Kellogg, R.L., D.C. Moffitt, and N.G. Gollehon, 2014, "Estimates of recoverable and non-recoverable manure nutrients based on the Census of Agriculture", Washington, DC: U.S. Department of Agriculture, Natural Resources Conservation Service.

Long, C.M., Muenich, R.L., Kalcic, M.M. & Scavia, D. 2018, "Use of manure nutrients from concentrated animal feeding operations", *Journal of Great Lakes Research,* vol. 44, no. 2, pp. 245-252.

Lory, J.A., Massey, R., Zulovich, J., Millmier, A., Hoehne, J., And, P.E., Case, C., 2001. Chapter 4 on Farm Evaluation of Adopting Phosphorus Versus Nitrogen Limits for Manure Application on US Swine Operations, Economic Viability of US Swine Farms Implementing Water Quality Best Available Technologies.

Maccoux, M.J., Dove, A., Backus, S.M. & Dolan, D.M. 2016, "Total and soluble reactive phosphorus loadings to Lake Erie: A detailed accounting by year, basin, country, and tributary", *Journal of Great Lakes Research,* vol. 42, no. 6, pp. 1151-1165.

Martin, J.F., Kalcic, M.M., Aloysius, N., Apostel, A.M., Brooker, M.R., Evenson, G., Kast, J.B., Kujawa, H., Murumkar, A., Becker, R., Boles, C., Confesor, R., Dagnew, A., Guo, T., Long, C.M., Muenich, R.L., Scavia, D., Redder, T., Robertson, D.M. & Wang, Y. 2021, "Evaluating management options to reduce Lake Erie algal blooms using an ensemble of watershed models", *Journal of Environmental Management,* vol. 280, pp. 111710.

Matisoff, G. & Ciborowski, J.J.H. 2005, "Lake Erie Trophic Status Collaborative Study", *Journal*

of Great Lakes Research, vol. 31, pp. 1-10.

Michalak Anna, M., Anderson Eric, J., Dmitry, B., Steven, B., Bosch Nathan, S., Bridgeman Thomas, B., Chaffin Justin, D., Kyunghwa, C., Rem, C., Daloğlu Irem, DePinto Joseph, V., Evans, M.A., Fahnenstiel Gary, L., Lingli, H., Ho Jeff, C., Liza, J., Johengen Thomas, H., Kuo Kevin, C., Elizabeth, L., Xiaojian, L., McWilliams Michael, R., Moore Michael, R., Posselt Derek, J., Peter, R.R., Donald, S., Steiner Allison, L., Ed, V., Wright David, M. & Zagorski Melissa, A. 2013, "Record-setting algal bloom in Lake Erie caused by agricultural and meteorological trends consistent with expected future conditions", *Proceedings of the National Academy of Sciences*, vol. 110, no. 16, pp. 6448-6452.

Muenich, R.L., Kalcic, M. & Scavia, D. 2016, "Evaluating the Impact of Legacy P and Agricultural Conservation Practices on Nutrient Loads from the Maumee River Watershed", *Environmental Science & Technology,* vol. 50, no. 15, pp. 8146-8154.

NCWQR, 2022, “About the HTLP”, https://ncwqr.org/monitoring/about-the-htlp/.

NWQC, 2022, “Water Quality Portal”, https://www.waterqualitydata.us/.

Ohio EPA, 2010. Ohio Lake Erie Phosphorus Task Force (Final Report). Ohio Environmental Protection Agency, Division of Surface Water, Columbus, OH.

Osterberg, D. & Wallinga, D. 2004, "Addressing externalities from swine production to reduce public health and environmental impacts", *American Journal of Public Health,* vol. 94, no. 10, pp. 1703-1708.

Paudel, J. & Crago, C.L. 2021, "Environmental Externalities from Agriculture: Evidence from Water Quality in the United States", *American Journal of Agricultural Economics,* vol. 103, no. 1, pp. 185-210.

Paytan, A., Roberts, K., Watson, S., Peek, S., Chuang, P., Defforey, D. & Kendall, C. 2017, "Internal loading of phosphate in Lake Erie Central Basin", *Science of The Total Environment,* vol. 579, pp. 1356-1365.

Raff, Z. & Meyer, A. 2022, "CAFOs and Surface Water Quality: Evidence from Wisconsin", *American Journal of Agricultural Economics,* vol. 104, no. 1, pp. 161-189.

Richards, R.P., Baker, D.B., Crumrine, J.P. & Stearns, A.M. 2010, "Unusually large loads in 2007 from the Maumee and Sandusky Rivers, tributaries to Lake Erie", *Journal of Soil and Water Conservation,* vol. 65, no. 6, pp. 450.

Scavia, D., David Allan, J., Arend, K.K., Bartell, S., Beletsky, D., Bosch, N.S., Brandt, S.B., Briland, R.D., Daloğlu, I., DePinto, J.V., Dolan, D.M., Evans, M.A., Farmer, T.M., Goto, D., Han, H., Höök, T.O., Knight, R., Ludsin, S.A., Mason, D., Michalak, A.M., Peter Richards, R., Roberts, J.J., Rucinski, D.K., Rutherford, E., Schwab, D.J., Sesterhenn, T.M., Zhang, H. & Zhou, Y. 2014, "Assessing and addressing the re-eutrophication of Lake Erie: Central basin hypoxia", *Journal of Great Lakes Research,* vol. 40, no. 2, pp. 226-246.

Scavia, D., Kalcic, M., Muenich, R.L., Read, J., Aloysius, N., Bertani, I., Boles, C., Confesor, R., DePinto, J., Gildow, M., Martin, J., Redder, T., Robertson, D., Sowa, S., Wang, Y. & Yen, H. 2017, "Multiple models guide strategies for agricultural nutrient reductions", *Frontiers in Ecology and the Environment,* vol. 15, no. 3, pp. 126-132.

Schindler, D.W. 1974, "Eutrophication and Recovery in Experimental Lakes: Implications for Lake Management", *Science*, vol. 184, no. 4139, pp. 897-899.

Schindler D., W., Hecky, R.E., Findlay, D.L., Stainton, M.P., Parker, B.R., Paterson, M.J., Beaty, K.G., Lyng, M. & Kasian, S.E.M. 2008, "Eutrophication of lakes cannot be controlled by reducing nitrogen input: Results of a 37-year whole-ecosystem experiment", *Proceedings of the National Academy of Sciences*, vol. 105, no. 32, pp. 11254-11258.

Schlenker, W. and M. Roberts, 2020, “Fine-scaled weather data”, http://www.columbia.edu/~ws2162/links.html.

Smith, V.H. 2003, "Eutrophication of freshwater and coastal marine ecosystems a global problem", *Environmental Science and Pollution Research*, vol. 10, no. 2, pp. 126-139.

Stumpf, R.P., Johnson, L.T., Wynne, T.T. & Baker, D.B. 2016, "Forecasting annual cyanobacterial bloom biomass to inform management decisions in Lake Erie", *Journal of Great Lakes Research,* vol. 42, no. 6, pp. 1174-1183.

Waller, D.M., Meyer, A.G., Raff, Z., and Apfelbaum, S.I. 2021, “Shifts in precipitation and agricultural intensity increase phosphorus concentrations and loads in an agricultural watershed”, *Journal of Environmental Management*, vol. 284, 112019.

Watson, S.B., Miller, C., Arhonditsis, G., Boyer, G.L., Carmichael, W., Charlton, M.N., Confesor, R., Depew, D.C., Höök, T.O., Ludsin, S.A., Matisoff, G., McElmurry, S.P., Murray, M.W., Peter Richards, R., Rao, Y.R., Steffen, M.M. & Wilhelm, S.W. 2016, "The re-eutrophication of Lake Erie: Harmful algal blooms and hypoxia", *Harmful Algae,* vol. 56, pp. 44-66.

Wolf, D. and Klaiber, H.A. 2017, “Bloom and bust: Toxic algae’s impact on nearby property values”, *Ecological Economics,* vol. 135, pp. 209-221.

Wolf, D., Gopalakrishnan, S., and Klaiber, H.A. 2022, “Staying afloat: The effect of algae contamination on Lake Erie housing prices”, *American Journal of Agricultural Economics*.

Yang, Q., Tian, H., Li, X., Ren, W., Zhang, B., Zhang, X., and Wolf, J. 2016, “Spatiotemporal patterns of livestock manure nutrient production in the conterminous United States from 1930-2012”, *Science of the Total Environment*, vol. 541, pp. 1592-1602.

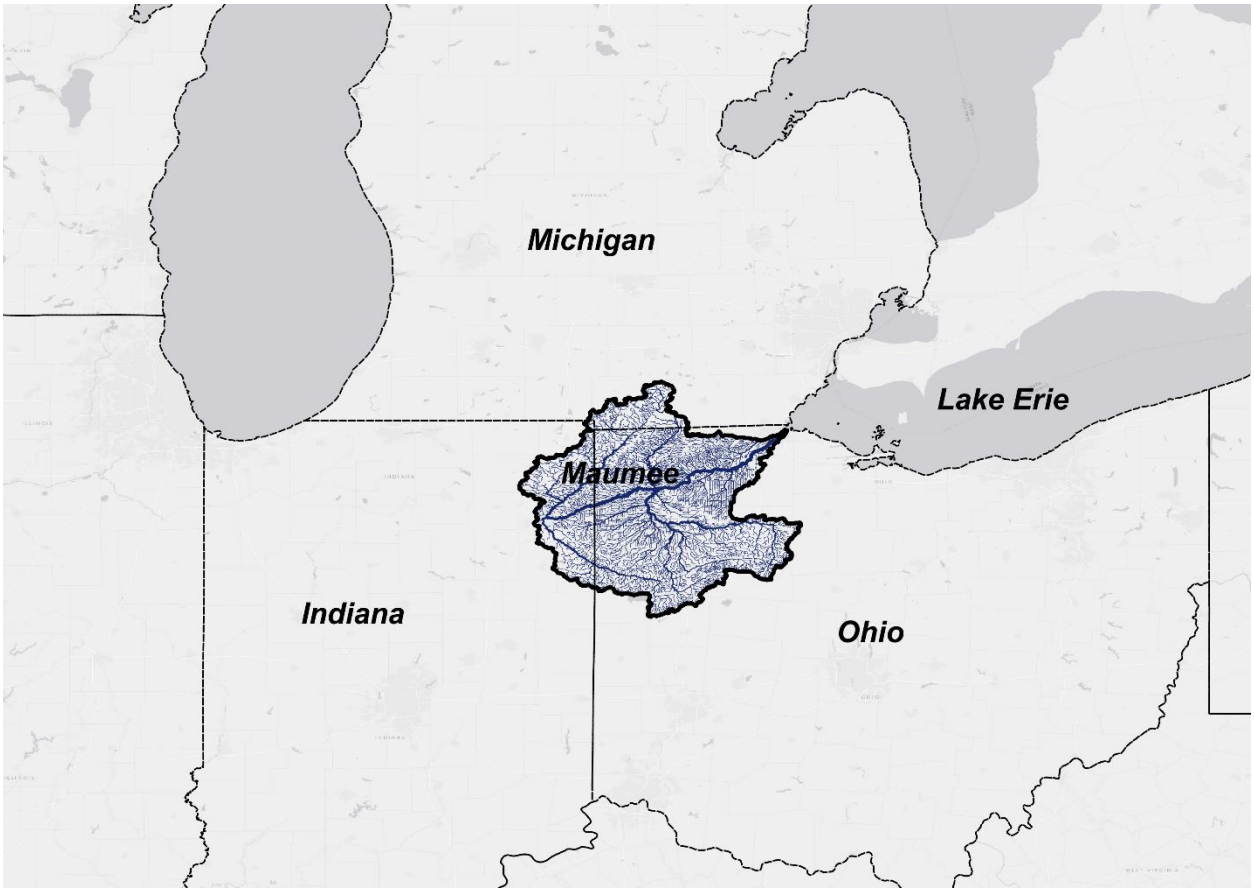

Figure 1. Location of the Maumee River watershed

Notes: This map shows the location of the Maumee River watershed, along with state borders for Indiana, Michigan, and Ohio.

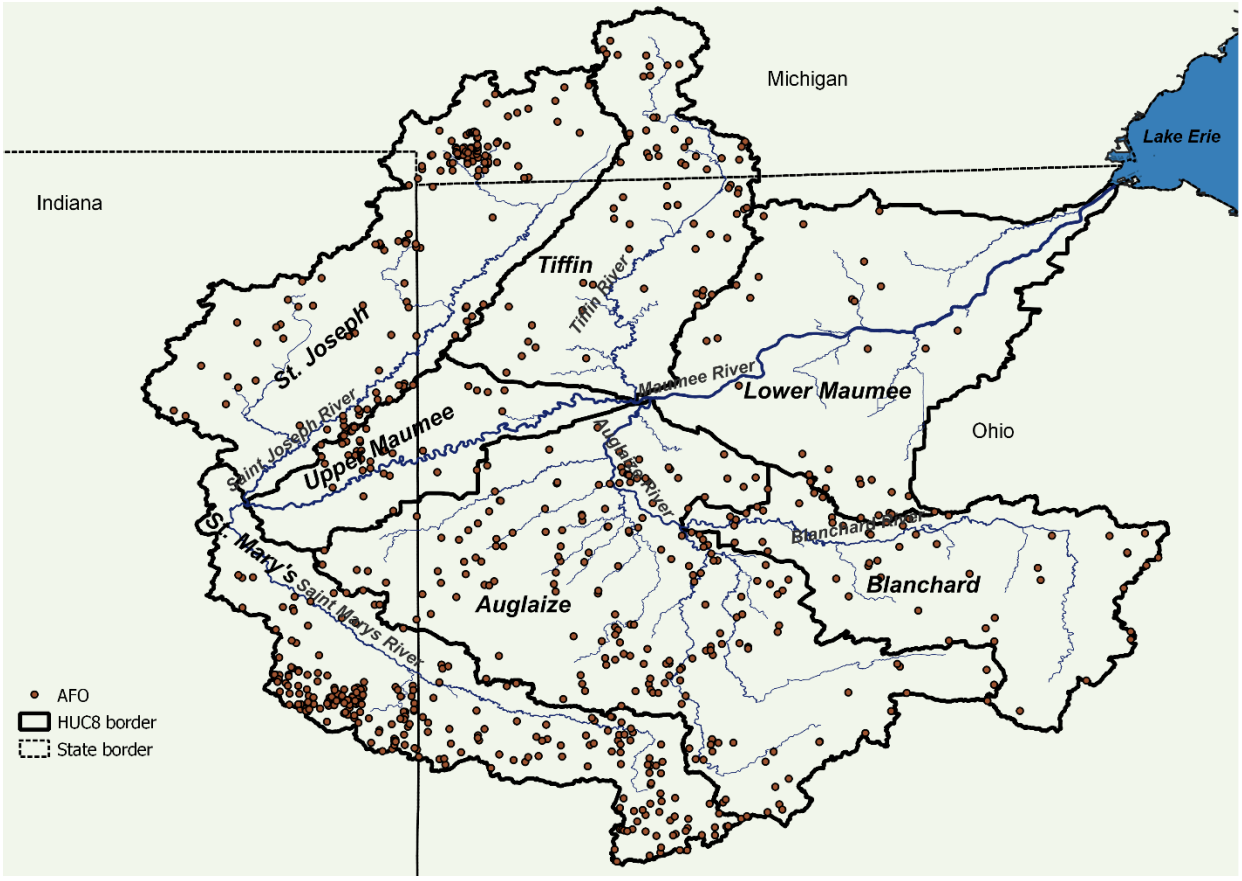

Figure 2. The Maumee River Watershed with HUC8 Borders and AFO Locations
Notes: This map shows the locations of AFOs within the Maumee River Watershed (2018), along with HUC8 borders.

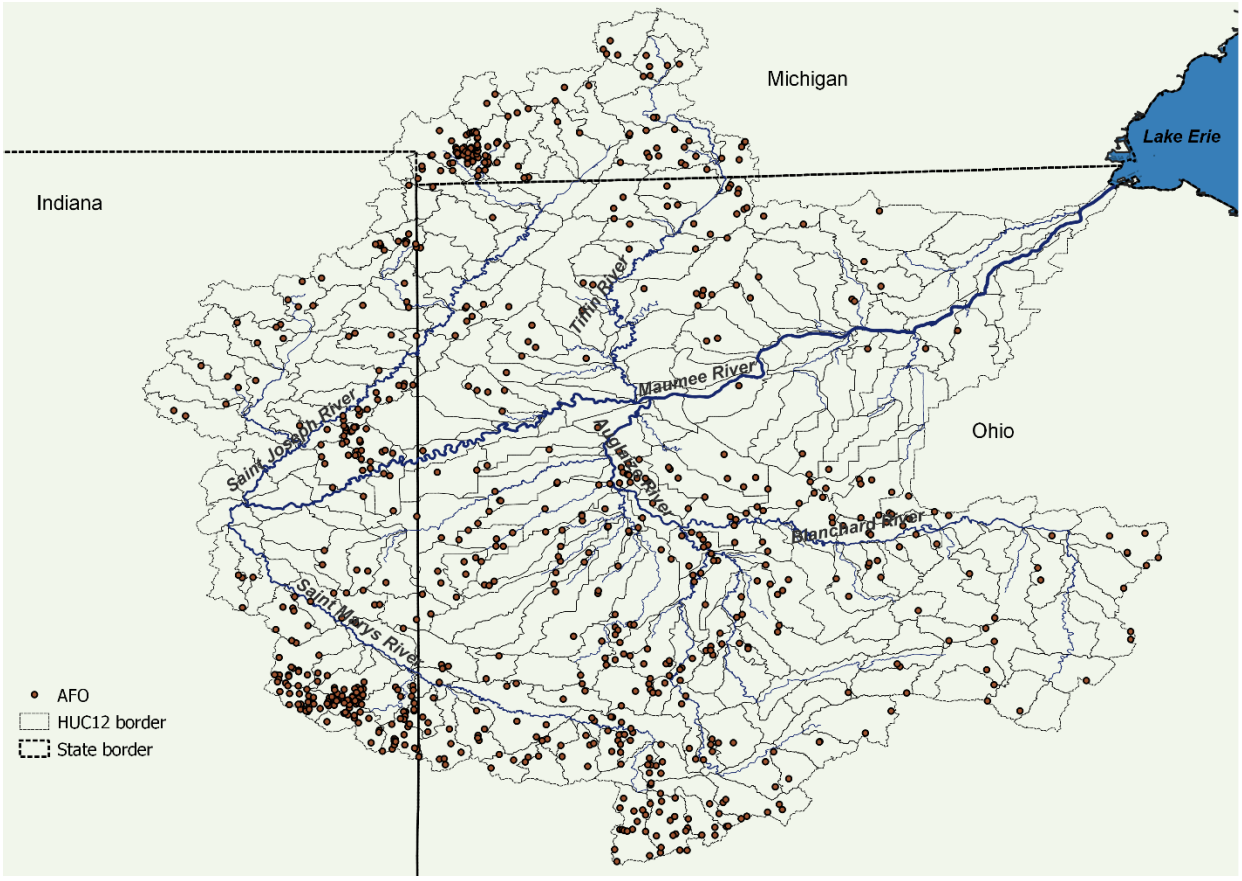

Figure 3. The Maumee River Watershed with HUC12 Borders and AFO Locations
Notes: This map shows the locations of AFOs within the Maumee River Watershed (2018), along with HUC12 borders.

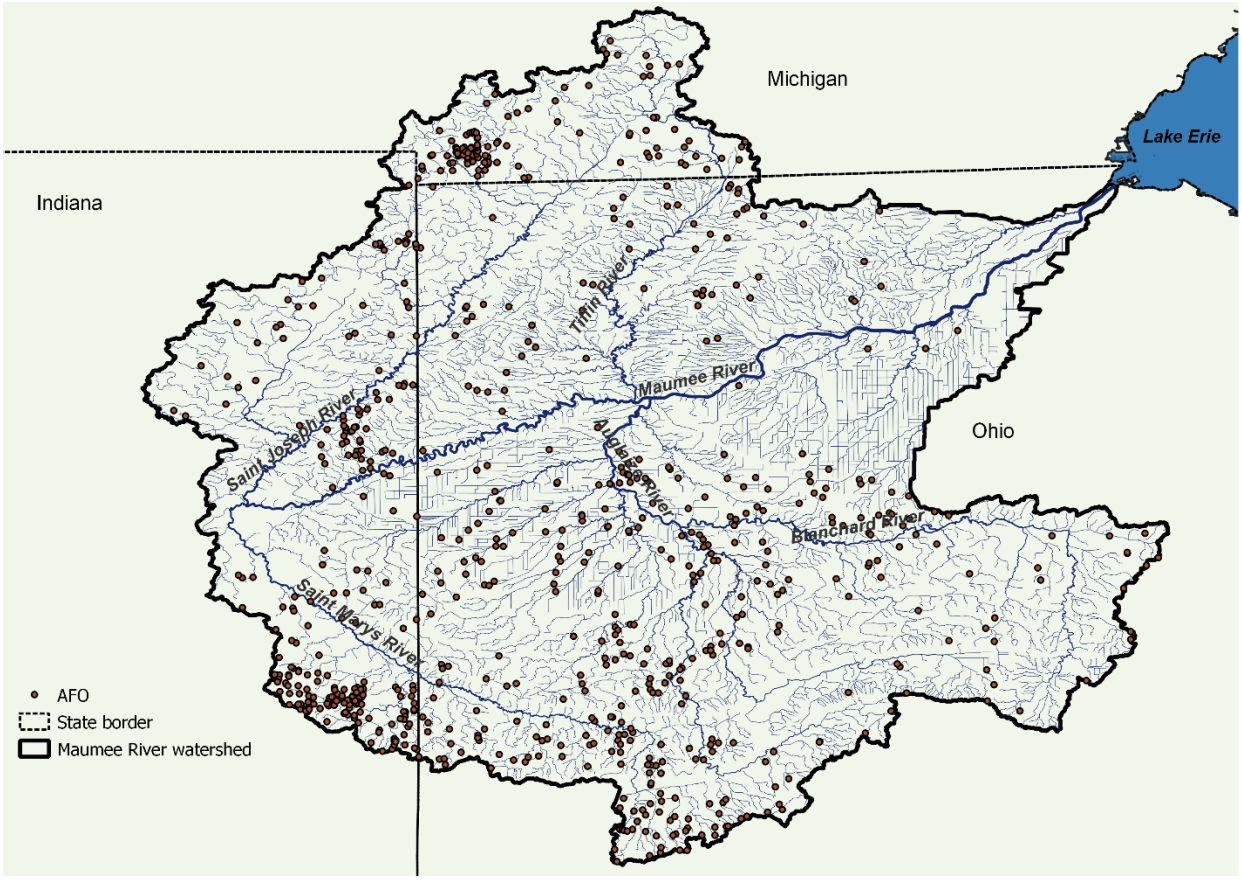

Figure 4. The Maumee River Watershed with the NHD Stream Network and AFO Locations
Notes: This map shows the locations of AFOs within the Maumee River Watershed (2018), along with stream reaches of the NHD stream and river network.

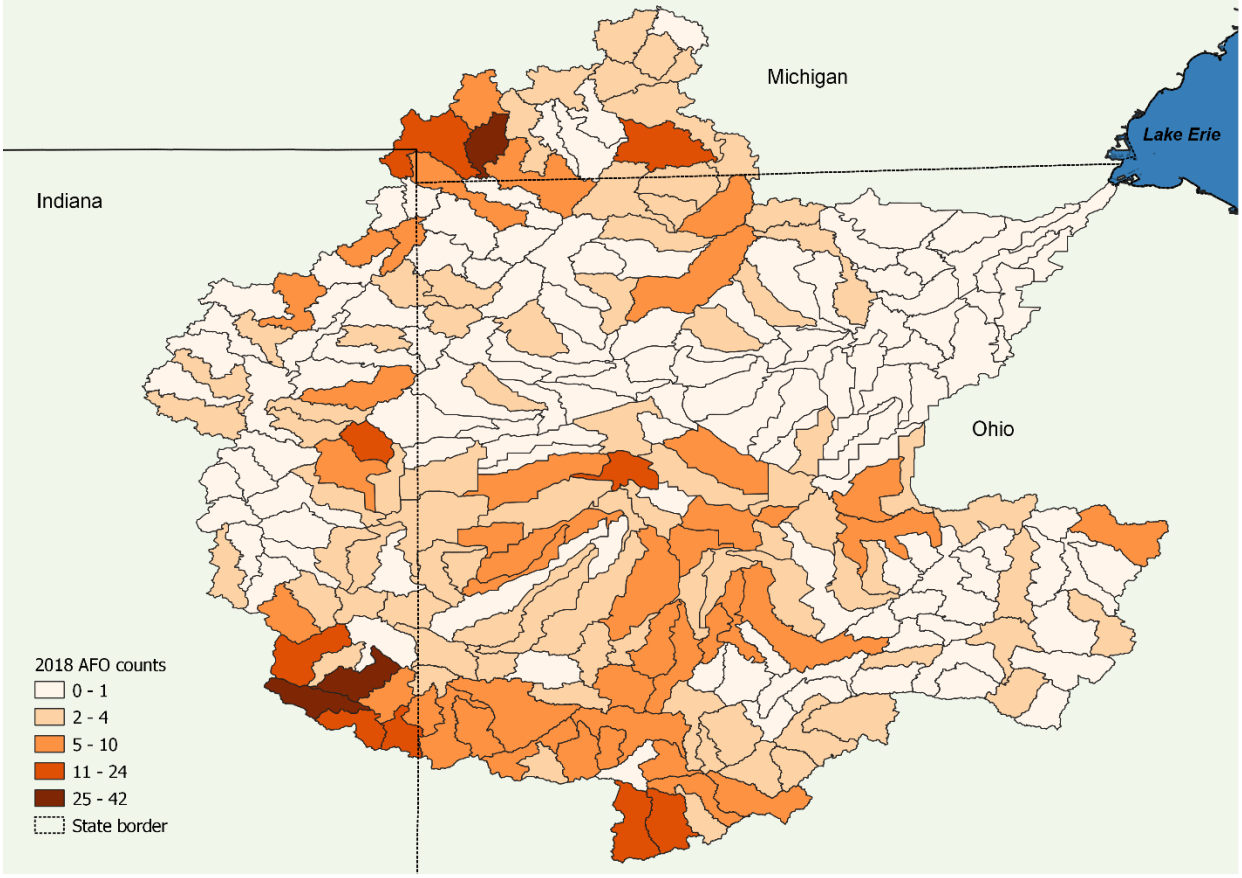

Figure 5. The Maumee River Watershed with 2018 AFO Counts

Notes: This map shows the number of AFOs operating in the Maumee River Watershed (2018) within each HUC12 subwatershed.

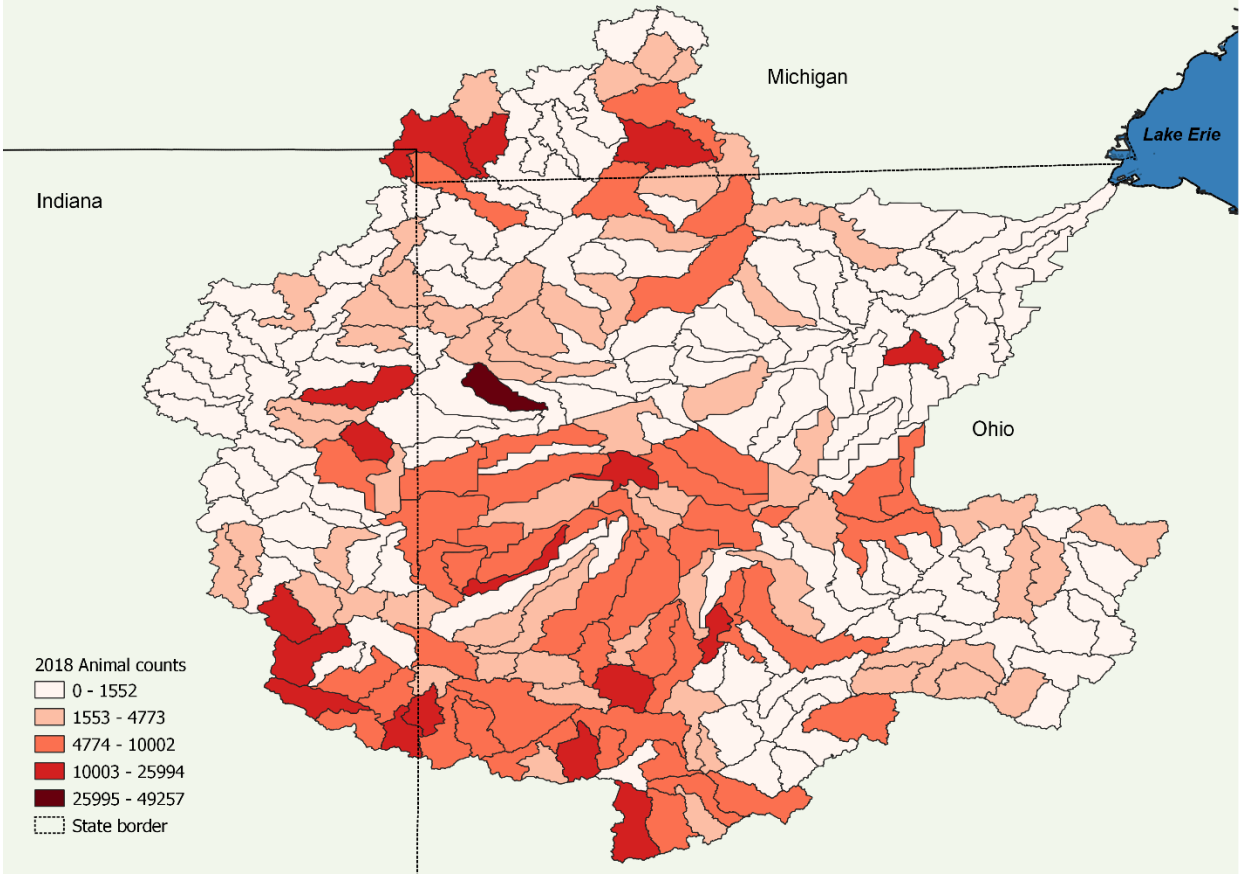

Figure 6. The Maumee River Watershed with 2018 animal counts

Notes: This map shows the number of animal units present in the Maumee River Watershed (2018) within each HUC12 subwatershed.

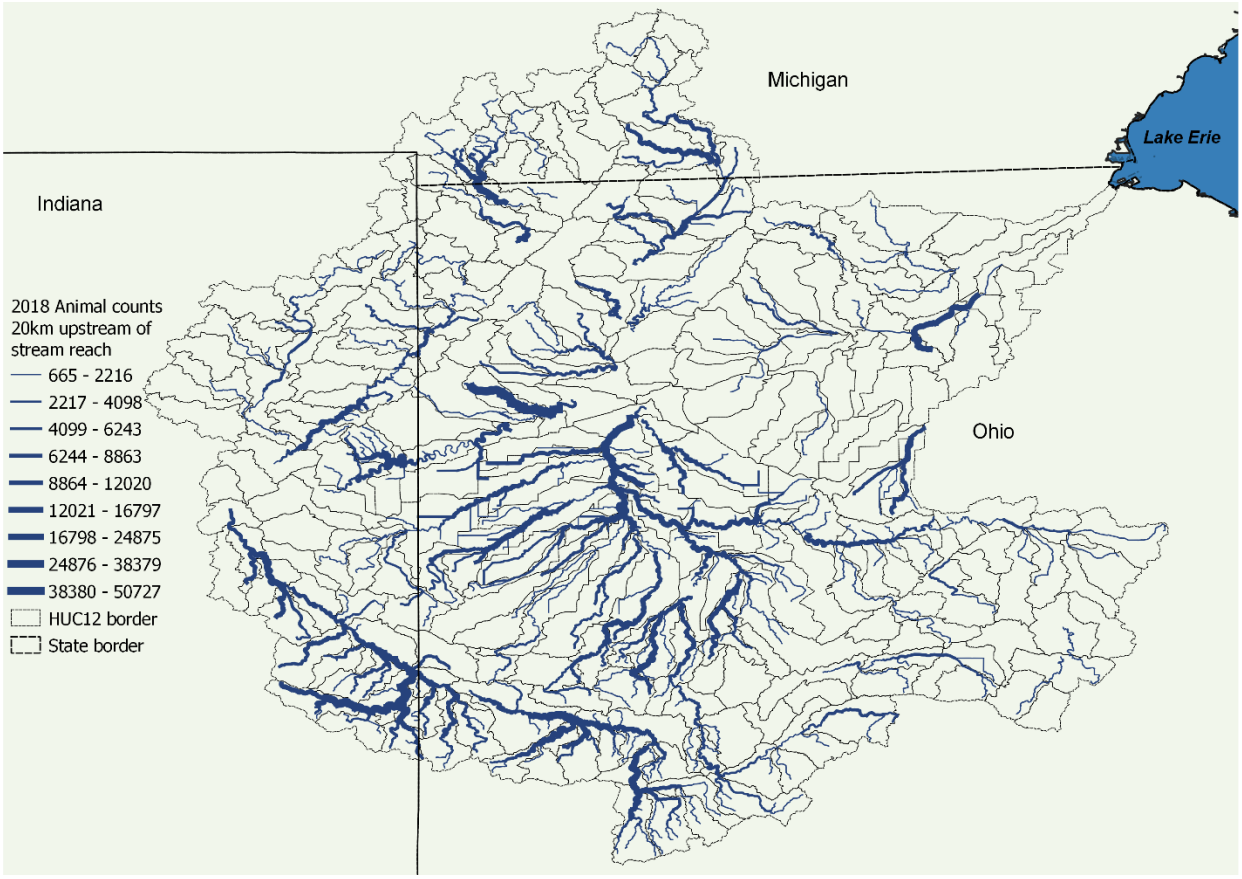

Figure 7. Stream Reach Exposure to Upstream AFOs in the Maumee River Watershed
Notes: This map shows the number of animal units present in the Maumee River Watershed (2018) within 20 km upstream of each stream reach. Wider lines indicate more upstream animal units. HUC12 borders are also shown.

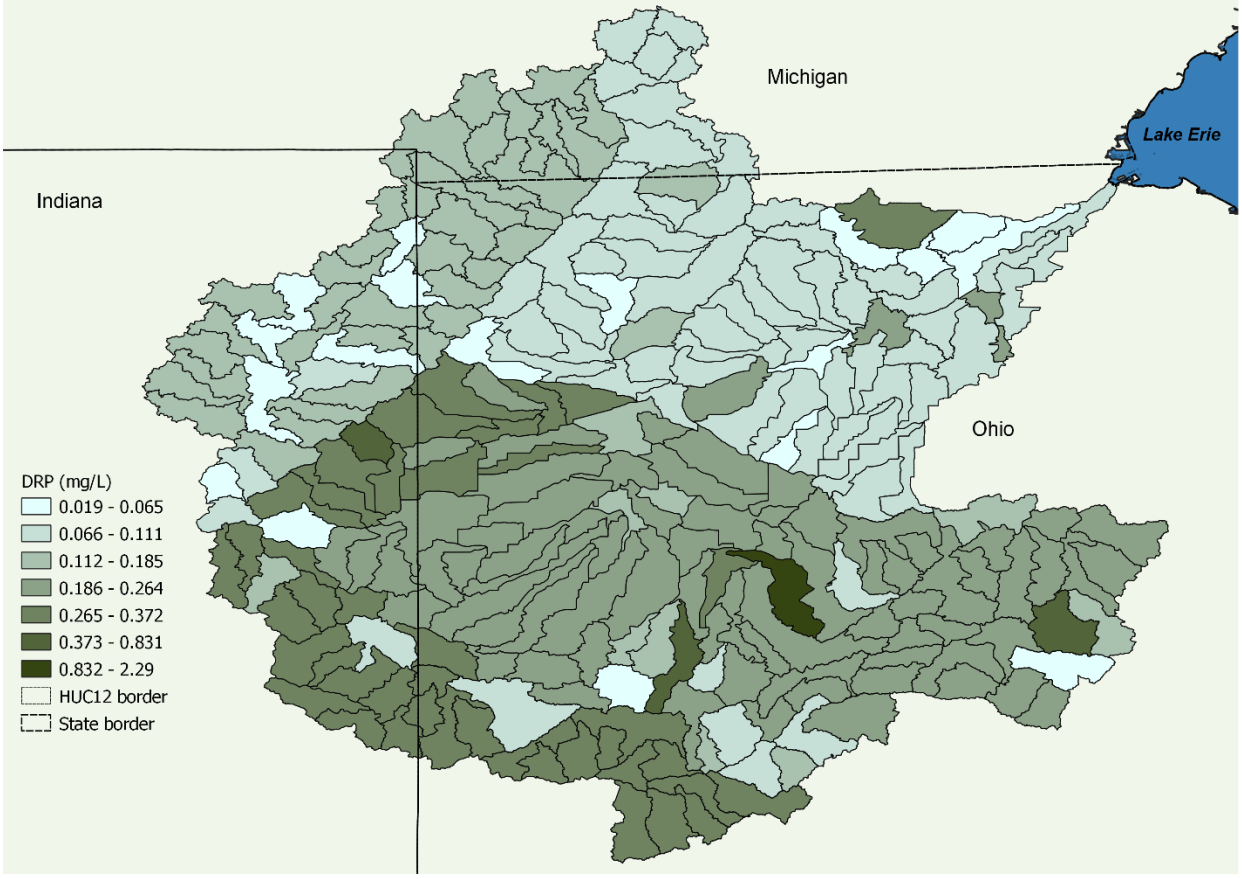

Figure 8. Variability in DRP Concentrations in the Maumee River Watershed
Notes: This map shows average DRP concentrations in the Maumee River Watershed from 2017-2019 within each HUC12 subwatershed. Classes are determined by Jenks natural breaks.

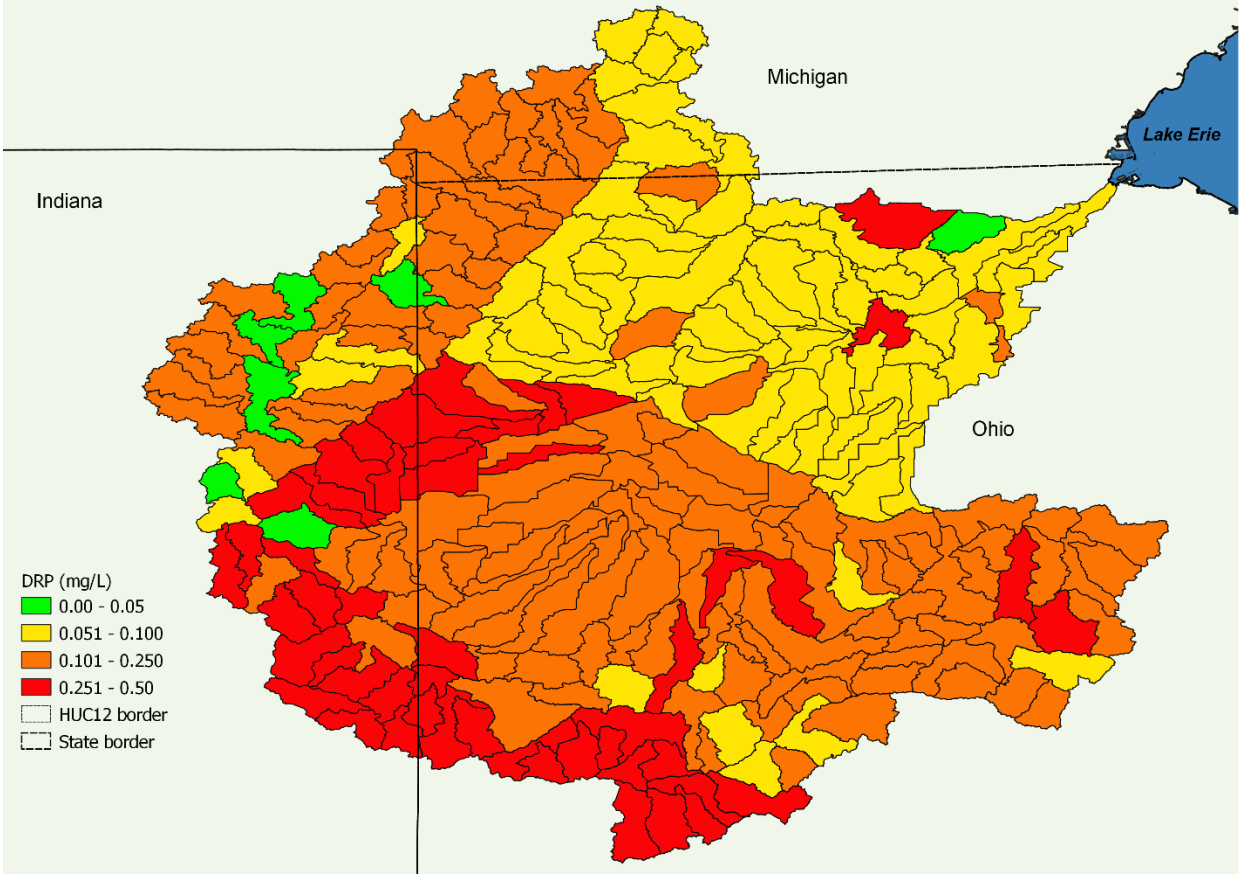

Figure 9. DRP Concentrations in the Maumee River Watershed vs. Potential Target Levels
Notes: This map shows average DRP concentrations in the Maumee River Watershed from 2017-2019 within each HUC12 subwatershed. Classes are manually set to reflect thresholds of interest from the Annex 4 Objectives and Targets Task Team (2015).

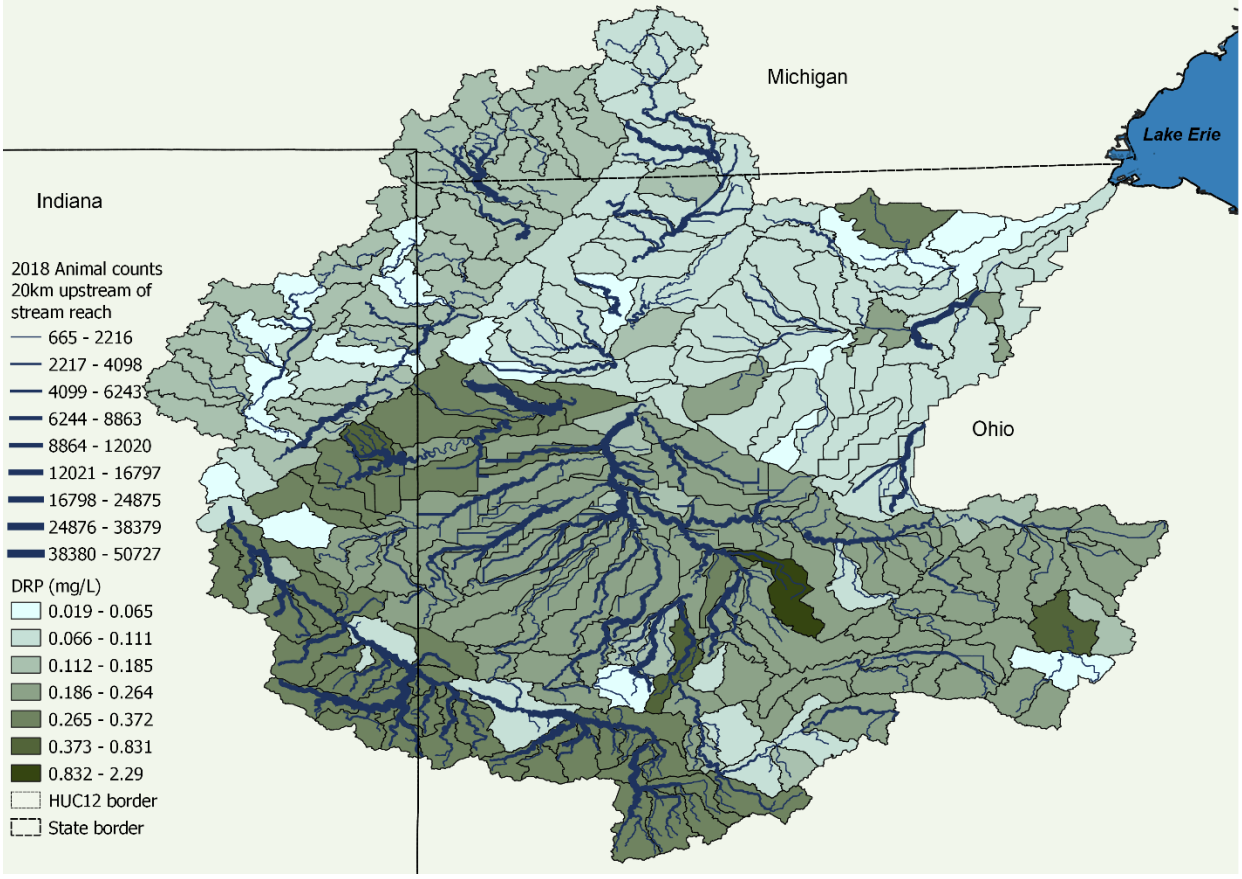

Figure 10. DRP Concentrations and Stream Reach Exposure to Upstream AFOs in the Maumee River Watershed

Notes: This map shows average DRP concentrations in the Maumee River Watershed from 2017-2019 within each HUC12 subwatershed and the number of animal units present in 2018 within 20 km upstream of each stream reach. Wider lines indicate more upstream animal units.

Table 1. Sample Summary Statistics

Variable	Mean	SD	Min	Max
Panel A. Dependent Variable				
DRP concentration (mg/L)	0.155	0.156	0.0005	2.38
Panel B. Treatment Regressors (within 20 km)				
Operating AFO	3.495	5.732	0	50
Animal count (000 au)	4.694	8.018	0	49.26
Manure production (000 tons)	26.98	42.30	0	234.9
Panel C. Control Variables				
Total precipitation (cm)	4.114	4.777	0.0336	45.91
Planted %	67.00	23.33	0.920	91.43
Developed %	19.58	19.84	4.780	83.55
Wetlands %	3.312	3.030	0.225	19.95
Commercial fertilizer/acre (000 kg)	3.976	0.876	2.268	6.336
Observations		1,673		

Notes: Summary statistics are at the stream reach-month (Panels A and B, total precipitation) and HUC12-month (planted %, developed %, wetlands %, commercial fertilizer) level and for the observations of the final analysis samples. Panel B presents summaries for AFO exposure within 20 km upstream of the stream reach, in km traveled via the stream and river network.

Table 2. Sample Summary Statistics of the Dependent Variables by AFO exposure

	(1) None	(2) >=1
Variable	Mean (SD)	Mean (SD)
DRP Concentration (mg/L)	0.120	0.174
	(0.104)	(0.175)
Difference of means p-value	0.000	
Observations	583	1,090

Notes: Summary statistics are at the stream reach-month level and for the observations of the final analysis sample. P-values represent univariate difference tests of mean DRP concentrations at stream reaches based on upstream AFO exposure. The first column shows mean DRP concentrations at stream reaches with zero upstream AFO exposure. The second column shows mean DRP concentrations at stream reaches with one or more upstream operating AFOs (the sample median).

Table 3. Effect of Upstream AFO Exposure on DRP Concentrations

Variable	(1)	(2)	(3)	(4)	(5)	(6)
Panel A. Treatment Measures						
Operating AFO	0.106**	0.130***				
	(0.0444)	(0.0390)				
Animal count (000 au)			0.0858**	0.0971***		
			(0.0381)	(0.0343)		
Manure produced (000 tons)					0.0182**	0.0206***
					(0.00835)	(0.00751)
Panel B. Control Variables						
Precipitation (cm)	0.0894***	0.0881***	0.0887***	0.0873***	0.0887***	0.0873***
	(0.0169)	(0.0167)	(0.0169)	(0.0167)	(0.0169)	(0.0167)
Precipitation^2 (cm)	-0.00198***	-0.00193***	-0.00196***	-0.00191***	-0.00196***	-0.00190***
	(0.000497)	(0.000491)	(0.000496)	(0.000490)	(0.000496)	(0.000490)
Planted %		0.00714		0.0188		0.0176
		(0.278)		(0.278)		(0.277)
Developed %		0.249		0.255		0.252
		(0.356)		(0.346)		(0.346)
Wetlands %		-3.10		-3.09		-3.09
		(2.16)		(2.14)		(2.14)
Commercial fertilizer/acre (000 kg)		0.561*		0.542*		0.540*
		(0.325)		(0.321)		(0.321)
Stream reach FE	X	X	X	X	X	X
Month FE	X	X	X	X	X	X
Year FE	X	X	X	X	X	X
Additional controls		X		X		X
Observations	1,673	1,673	1,673	1,673	1,673	1,673

Notes: Each column presents OLS regression results from a separate specification of equation (1). The dependent variable is the log-transformed concentration of DRP at the stream reach-month level. This table presents results for remotely sensed AFO exposure within 20 km upstream of the stream reach, in km traveled via the stream and river network. Robust standard errors in parentheses are clustered at the stream reach level. *** p<0.01, ** p<0.05, * p<0.1.

Supplementary Material (for online publication only)

Table S1. Effect of Upstream AFO Exposure on DRP Concentrations – Within 10 km

Variable	(1)	(2)	(3)	(4)	(5)	(6)
Panel A. Treatment Measures						
Operating AFO	0.0885**	0.126**				
	(0.0399)	(0.0564)				
Animal count (000 au)			0.0962	0.133		
			(0.0621)	(0.0868)		
Manure produced (000 tons)					0.0208	0.0281
					(0.0148)	(0.0204)
Panel B. Control Variables						
Precipitation (cm)	0.0892***	0.0879***	0.0891***	0.0877***	0.0890***	0.0877***
	(0.0169)	(0.0168)	(0.0169)	(0.0168)	(0.0169)	(0.0168)
Precipitation^2 (cm)	-0.00197***	-0.00192***	-0.00197***	-0.00192***	-0.00197***	-0.00191***
	(0.000498)	(0.000492)	(0.000497)	(0.000491)	(0.000497)	(0.000491)
Planted %		-0.0137		-0.0170		-0.0188
		(0.280)		(0.279)		(0.279)
Developed %		0.206		0.199		0.195
		(0.364)		(0.362)		(0.362)
Wetlands %		-3.25		-3.27		-3.27
		(2.21)		(2.21)		(2.22)
Commercial fertilizer/acre (000 kg)		0.535		0.526		0.523
		(0.328)		(0.326)		(0.326)
Stream reach FE	X	X	X	X	X	X
Month FE	X	X	X	X	X	X
Year FE	X	X	X	X	X	X
Additional controls		X		X		X
Observations	1,673	1,673	1,673	1,673	1,673	1,673

Notes: Each column presents OLS regression results from a separate specification of equation (1). The dependent variable is the log-transformed concentration of DRP at the stream reach-month level. This table presents results for remotely sensed AFO exposure within 10 km upstream of the stream reach, in km traveled via the stream and river network. Robust standard errors in parentheses are clustered at the stream reach level. *** p<0.01, ** p<0.05, * p<0.1.

Table S2. Effect of Upstream AFO Exposure on DRP Concentrations – Within 30 km

Variable	(1)	(2)	(3)	(4)	(5)	(6)
Panel A. Treatment Measures						
Operating AFO	0.127***	0.146***				
	(0.0398)	(0.0411)				
Animal count (000 au)			0.0931***	0.112***		
			(0.0275)	(0.0404)		
Manure produced (000 tons)					0.0162***	0.0195**
					(0.00607)	(0.00848)
Panel B. Control Variables						
Precipitation (cm)	0.0896***	0.0882***	0.0890***	0.0874***	0.0889***	0.0873***
	(0.0169)	(0.0167)	(0.0168)	(0.0166)	(0.0168)	(0.0167)
Precipitation^2 (cm)	-0.00200***	-0.00194***	-0.00197***	-0.00191***	-0.00197***	-0.00190***
	(0.000499)	(0.000494)	(0.000495)	(0.000490)	(0.000496)	(0.000490)
Planted %		0.161		0.269		0.228
		(0.287)		(0.314)		(0.314)
Developed %		0.277		0.270		0.283
		(0.345)		(0.331)		(0.335)
Wetlands %		-3.15		-3.21		-3.25
		(2.21)		(2.20)		(2.20)
Commercial fertilizer/acre (000 kg)		0.579*		0.554*		0.555*
		(0.317)		(0.314)		(0.315)
Stream reach FE	X	X	X	X	X	X
Month FE	X	X	X	X	X	X
Year FE	X	X	X	X	X	X
Additional controls		X		X		X
Observations	1,673	1,673	1,673	1,673	1,673	1,673

Notes: Each column presents OLS regression results from a separate specification of equation (1). The dependent variable is the log-transformed concentration of DRP at the stream reach-month level. This table presents results for remotely sensed AFO exposure within 30 km upstream of the stream reach, in km traveled via the stream and river network. Robust standard errors in parentheses are clustered at the stream reach level. *** p<0.01, ** p<0.05, * p<0.1.

Table S3. Square Footage per Animal Type as Derived from Industry, Academic or Government Guidelines

	Square Footage Allotted to Animal type	Source
Dairy	80 (based on 1100 - 1300 lb. heifer)	Penn State Extension
Cattle	35 (average of access to yard and no access to yard)	Midwest Planning Service
Swine	7.4 (average of optimal economic and productivity)	National Pork Board
Poultry	.465 (layers)	United Egg Producers

Table S4. Estimated Animal Counts in the MRW, as of 2018

Estimated Animal Counts in the MRW (2018)				
	Indiana	Michigan	Ohio	Total
Dairy	12,949	15,494	69,834	98,277
Cattle	21,527	29,288	18,652	69,467
Swine	239,595	18,560	789,904	1,048,059
Poultry	4,610,857	285,076	14,323,216	19,219,149
Total	4,884,928	348,418	15,201,606	20,434,952