Factors Driving the Spatiotemporal Variation of Dissolved Organic Carbon Flux from Croplands in the Midwestern USA

Yong Qian Tian (✉ yong.tian@cmich.edu)
Central Michigan University

Qian Yu
University of Massachusetts Amherst

Hunter Carrick
Central Michigan University

Brian Becker
Central Michigan University

Remegio Confesor
Norwegian Institute of Bioeconomy Research: Norsk Institutt for Biookonomi

Mark Francek
Central Michigan University

Olivia Anderson
Central Michigan University

Research Article

Keywords: dissolved organic carbon, inland water, anthropogenic activity, crop residue management, baseflow, agricultural landscape, hydrological modelling

Posted Date: January 14th, 2022

DOI: https://doi.org/10.21203/rs.3.rs-1223810/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License. Read Full License
Factors Driving the Spatiotemporal Variation of Dissolved Organic Carbon Flux from Croplands in the Midwestern USA

Yong Q. Tian\textsuperscript{a*}, Qian Yu\textsuperscript{b}, Hunter J. Carrick\textsuperscript{c}, Brian L. Becker\textsuperscript{a}, Remegio Confesor\textsuperscript{d}, Mark Francek\textsuperscript{a}, and Olivia C. Anderson\textsuperscript{c}

\textsuperscript{a*}: Department of Geography and Environmental Studies & Institute for Great Lakes Research, Central Michigan University, MI 48858

Email: yong.tian@cmich.edu

\textsuperscript{b}: Department of Geosciences, University of Massachusetts-Amherst, Amherst, MA 01003,

\textsuperscript{c}: Department of Biology & Institute for Great Lakes Research, Central Michigan University, MI 48858

\textsuperscript{d}: Environment and Natural Resources, Norwegian Institute of Bioeconomy Research, Norwegian

* Corresponding author
Abstract:

Improving understanding of dissolved organic carbon (DOC) cycling from farmlands to rivers is a challenge due to the complex influence of farming practices, the hydrology of predominantly flat lowlands, and seasonal snowpack effects. Monthly field DOC concentrations were measured throughout the year at sub-basin scale across the Chippewa River Watershed, which falls within the Corn Belt of the Midwestern United States. The observations from croplands were benchmarked against the data sampled from hilly forested areas in the Connecticut River Watershed. The Soil Water Assessment Tool (SWAT) was used to simulate daily soil water properties. This method tests for a framework for using the combination of new field data, hydrological modelling, and knowledge-based reclassification of Land Use/Land Cover (LULC) to analyze the predictors of both the spatial and temporal changes of DOC over farmlands. Our results show: 1) DOC concentrations from cropland baseflow were substantially high throughout the year, especially for spring runoff/snowmelt scenarios, 2) gradient analysis with spatial factors only was able to explain ~82% of observed annual mean DOC concentrations, and 3) with both spatial and temporal factors: [Evapotranspiration, Soil Water, and Ground Water], the analysis explained ~81% of seasonal and ~54% of daily variations in observed DOC concentrations.

Keywords: dissolved organic carbon, inland water, anthropogenic activity, crop residue management, baseflow, agricultural landscape, hydrological modelling
INTRODUCTION

Terrestrial organic carbon in the dissolved form (DOC) is easily transported to inland or coastal waters via hydrological processes (Dusek et al. 2017; Heppell et al. 2017). Excessive riverine carbon has a pronounced impact on aquatic ecosystems via processes such as amplifying microbial activity, diminishing the quantity and quality of light penetrating the water column, and controlling toxic metal availability (Butman and Raymond 2011; Spencer et al. 2013; Stedmon et al. 2006). As a major use of land (Ellis et al. 2010), agricultural landscapes are a significant source of DOC in freshwater ecosystems (Tranvik 2014). Recent studies raised a public health concern that rivers receiving runoff from croplands have elevated DOC concentrations at regional and worldwide scales (Bhattacharya and Osburn 2020; Qiao et al. 2017). DOC originating from croplands are enriched in aromatic structures and exhibit high labile properties, thereby significantly affecting freshwater nutrient pathways (Bhattacharya and Osburn 2020; Holgerson and Raymond 2016; Kellerman et al. 2020; Tranvik 2014).

In recent decades, there has been an increased interest in studying terrestrial carbon-source dynamics and their impact on freshwater interactive processes (Heppell et al. 2017). As such, studies have made significant progress towards gaining a better understanding of the impacts of climate change and extreme storm events on variations of riverine DOC concentrations originating from terrestrial environments (Kellerman et al. 2020). However, these studies are often limited in scope because they were set at entire river basin scales without considering the heterogeneity and do not include sub-basin land-surface characteristics and hydrological processes. In addition, these studies have largely focused on semi-natural habitats involving little or no agriculture such as estuarine wetland, scrub-shrub, and forested areas (Raymond and Saiers 2010; Rudolph et al. 2020; Tian et al. 2012). Studies focused on forested landscapes are typically dominated by sloped landscapes that generate DOC fluxes predominantly through surface runoff (Raymond and Saiers 2010; Singh et al. 2015). A representative
DOC model for studying hilly forested regions is the Integrated Catchments model for Carbon (INCA-C) (Futter et al. 2011). By using the INCA-C model, a study revealed that variations in DOC concentrations was primarily driven by soil temperature and surface runoff over forested, mountainous watersheds in Sweden (Clark et al. 2007; Fu et al. 2019; Pers et al. 2016). However, several other studies further clarified that event-based fluxes do not necessarily occur without baseflow during rainfall events; rather, event-based flows usually account for both surface runoff and baseflow during, and a few days after, a rainfall event (Ågren et al. 2008; Olsson et al. 2009). Therefore, DOC models that are based on measurements sampled from semi-natural habitats are not readily adaptable to agricultural landscapes in which carbon sources are typically more heavily influenced by complex anthropogenic factors, have a higher degree of heterogeneity of land cover types, and exhibit the major role of baseflow in lowland/flatland hydrological processes.

The DOC dynamics in streamflow generated from agricultural landscapes are significantly different than those reported for semi-natural habitats. This is especially true for the Corn Belt region of the Midwestern USA, where flatland hydrology dictates temporal variations in DOC concentrations via baseflow rather than overland surface runoff (Qiao et al. 2017). This is in sharp contrast to hilly forested areas where the majority of annual DOC fluctuations are dictated predominantly by surface runoff during rainfall events. Even over forested lands, it has been reported that low-flow conditions are ideal in order to best understand the transport processes of nutrients (e.g., N and P) from the soil profile and land surface (Lindström et al. 2010; Pers et al. 2016).

In general, the factors driving spatiotemporal variation of DOC concentrations in streams of croplands between rainfall events are not well understood. These temporal DOC variations are assumably controlled by soil properties such as organic matter content, level of saturation, and temperature. The spatial variations are likely associated with heterogeneity of the landscape, such as the
typical patchwork of small water bodies, wetlands, forest, and a broad range of crop density. More variation or heterogeneity is introduced when considering the varied crop residue management policies implemented by land managers. It remains largely unproven if both the spatial and temporal variations of DOC fluxes from croplands to receiving waters can be reliably quantified, and what the appropriate sub-basin scales are for quantifying these spatiotemporal variations.

The goal of this study is to analyze how DOC concentrations originating from crop-dominant lowlands vary by season and spatial distribution. We aim to identify a range of potential drivers that are useful to modelling the spatial disparities and trends of DOC variations instead of using soil C:N ratio as a surrogate (Aitkenhead-Peterson et al. 2003). The research niche includes the incorporation of detailed hydrologic variables, the adaptation of high spatial-resolution land surface features, and the collection and utilization of primary field observations. This study is focused on identifying a workable spatiotemporal scale of both field observations and quantitative analysis of DOC source dynamics, fate, and transport processes. This research focus is both novel and urgent because DOC export from agricultural landscapes occurs at the expense of the water quality of adjacent freshwater systems and to a lesser extent the land sustainability. Detrimental agricultural land management policies and practices can inherently lead to degraded water quality through both physical and biochemical processes (Jones et al. 2004; Jones and Knowlton 2005; Monaghan et al. 2007; Valentin et al. 2008). Clearly, DOC is part of a delicately balanced synergistic system and thus, modelling the concentration of DOC in lakes and rivers can provide a useful index of land-water ecology (Gómez-Gener et al. 2021).

MATERIALS AND METHODS

Study site:
Two contrasting study sites (cropland versus sloped forestland) were referenced in this study to allow direct comparison of DOC dynamics from croplands to a more homogeneous, forested landscape. The primary study site consisted of the Chippewa River watershed, which is one of the three tributaries draining the western portion of the larger Saginaw River watershed. The Saginaw River ultimately empties into Saginaw Bay, Lake Huron, which is one of the most bio-productive coastal regions in the entire Great Lakes Ecosystem (Millie et al. 2006). The Chippewa watershed cuts across the Lower Peninsula of Michigan and extends nearly 92 miles to Midland, Michigan, flowing predominantly eastward and ultimately merging with the Tittabawassee River. The croplands in our primary study site were dominated by row crops of corn, soy, wheat, and sugar beets, that were situated adjacent to mixed woodlots and wetlands, with a small proportion of lands dedicated to hay production and areas of development. The topography (mean slope < 0.3°), and climate in the primary study site were typical to the flat lowlands of the midwestern United States. In-situ DOC samples (153) were collected monthly at key locations across the Chippewa River watershed from 2012 to 2014.

The secondary study site for this study included five largely forested sub-basins within the greater Connecticut River Watershed, which is the longest river in the northeastern United States. DOC samples extracted from these forested sub-basins provided direct comparisons to those collected in the primary study site. The Connecticut River Watershed is one of the larger watersheds in the Northeast, draining approximately 11,000 square miles. In-situ DOC samples (125 samples) were collected monthly from 2011 to 2016 across these five forested sub-basins within the Connecticut River Watershed (Li et al. 2018). Two of the five sub-basins had landcover types dominated by northeastern deciduous forests. Northeastern conifer forests were the dominant vegetation in two of the sub-basins. The fifth sub-basin’s land cover was classified as a northeastern conifer/deciduous mixed forest. These 125 forested, in-situ
field observations collected across multiple years and seasons were paramount to better quantify the significant DOC contributions from croplands in comparison to forested landscapes.

Field Sample Protocols

There were 278 water samples collected in total, from 26 sampling locations via 43 field visits spread between the two study sites. Within the primary study site, a range of locations were selected and organized into two sampling groups (crop dominated and mixed croplands) as displayed in Figure 1. The 13 sampling locations (IDs 1-13) associated with the first group focused on crop-dominated drainage sub-basins. The 8 sampling locations (IDs 14-21) associated with the second group was added to include drainage sub-basins dominated by the croplands mixed with a large proportion of forests and wetlands. Monthly field visits were arranged between rainfall events to best reflect seasonal variations and to reduce the influence of any single event. The entire cropland field campaign ran for 14 months beginning in October 2012 and ending in January 2014. Field sampling activities were primarily conducted during the spring and fall when soil biochemical processes and nutrient fluxes are more prevalent (Sela et al. 2019; Sharratt et al. 1998). DOC concentrations are generally elevated in the springtime months due to snowpack dependent biochemical processes, soil moisture conditions, and runoff from snowmelt (Qiao et al. 2017). DOC concentrations are also elevated in the autumn months when crop biomass is at its peak and when such biomass is actively manipulated due to harvest, residue management practices, and fall tillage (Kelly et al. 2015; Vetsch and Randall 2002). A reduced number of field samples were collected for the purpose of analyzing the transition of DOC concentrations between seasons. In total, 153 water samples were collected to satisfactorily quantify seasonal patterns in DOC variation within receiving waters.
Laboratory Protocols to Estimate DOC concentration

Field samples were obtained from all sampling locations using clean 500 mL bottles (acid washed) to collect stream water for DOC analysis. Bottles were fully immersed while capping to ensure that no air remained. The samples were immediately stored on ice until arrival at the lab for immediate filtering. Water samples were filtered through pre-combusted glass-fiber filters (nominal 0.7 µm pore size) to remove any non-dissolved organic matter, and the filtrate was then stored (acidified and refrigerated) until analysis for DOC content. All Laboratory processes were completed within 12 hours of sample collection. The DOC concentration of each water sample was measured using a Shimadzu TOC-V analyzer with high temperature combustion (Vlahos et al. 2002). For each, 50 µl injections of sample water was combusted at 800 °C, from which the DOC concentration was calculated from the resultant CO$_2$ yield and measured with a non-dispersive infrared detector.

Land Cover Composition

National Land Cover Data (NLCD) data from 2011 was used because this date was closest to our sampling dates (2012-2014). The original baseline NLCD data referenced 13 land cover categories. We reclassified these baseline land cover data into three more broadly defined classes according to their general DOC transforming rates based on preliminary modelling runs (Figure 2). The percent-areal composition of the three new classes extracted from the NLCD data were calculated for each drainage sub-basin in the primary study site and are shown in Table 1. These data were used to analyze DOC dynamics at the sub-basin level. It is important to note that land cover data was needed for the entire area draining to an individual sampling location and not just the sub-basin in which the sampling
location resided. For instance, sampling point 15 receives runoff not only from its home (immediate) sub-basin, but also from upstream sub-basins that contained points 16, 17 and 18.

**Hydrological Modeling**

The Soil Water Assessment Tool (SWAT) was used to generate daily hydrologic characteristics for all sub-basins of the primary study site. We examined the simulated outputs to identify which variables, if any, were significant estimators of seasonal variation in DOC concentrations from croplands. The SWAT model is recognized worldwide as one of the most effective tools for estimating hydrological processes from agricultural lands (Arnold and Fohrer 2005; Olaoye et al. 2021; Tian et al. 2012). For this study, we implemented SWAT using 38 years of daily hydrologic input data spanning from 1981 to 2018. In addition, spatially explicit, high resolution (4 km\(^2\) grid) daily weather forcing data (precipitation, maximum temperature, and minimum temperature) were downloaded from the PRISM website ([https://prism.oregonstate.edu/recent/](https://prism.oregonstate.edu/recent/)) and also used as SWAT inputs. Daily discharge flow of the Chippewa River was obtained from the USGS 04154000-gauge station and used to calibrate the SWAT model. LULC types were extracted from NCLD 2011 as described above and used as inputs. Soil input parameters were extracted from SWAT’s built-in STATSGO data. Elevation data were extracted from a 30m Digital Elevation Model (DEM) of Isabella County, Michigan. As a result, the SWAT model delineated our primary study site into 126 sub-basins.

Next, Chippewa River daily flow data were categorized into high (top 10 percentile), medium (10th-50th percentile), and low (lowest 50\(^{th}\) – 100\(^{th}\) percentile) flows. The SWAT model was then calibrated through a multi-objective auto-calibration for these three flow regimes for the period from 2011 to 2018 using the method developed by (Confesor Jr and Whittaker 2007). The overall daily Nash–
Sutcliffe model efficiency coefficient (NSE) was 0.67 and the corresponding $R^2$ was 0.886. The simulated hydrologic properties for each of the 126 SWAT modeled sub-basins were used as independent variables for quantitative analysis of daily DOC observations.

**Exploratory Statistical Analysis**

Multiple linear regression analysis was used to explore which independent variables were most relevant in explaining variation in DOC concentrations at various spatial and temporal scales. In essence, these linear regressions aimed to fit observed/sampled DOC data to a linear equation with $m + n$ independent variables, where $m$ is the number of variables associated with the characteristics of each drainage sub-basin and $n$ is the number of temporal/seasonal variables. The specification of the linear model was as follows:

$$y(i, t) = c + \sum_{j=1}^{m} \alpha_j x_j(i) + \sum_{k=1}^{n} \beta_k S_k(i, t)$$ (1)

Let $X(i) = [x_1(i), x_2(i), \ldots, x_m(i)]$ be the $m$ exploratory spatial variables for the sub-basin associated with sampling location $i$ for $i = 1 \ldots L$. $L$ is the number of sampling locations. Similarly, let $S(i, t) = [s_1(i,t), s_2(i,t), \ldots, s_n(i,t)]$ be the $n$ exploratory seasonal (temporal) variables for the sub-basin $i$ and Julian day $t$, for $t = 1, 2 \ldots Q$. $Q$ is the last Julian day of field sampling visits. Each variable set $X(i)$ were split in three categories: land cover (e.g., percent cropped area, percent forested), soil properties (e.g., percent silt, percent organic matter), and geomorphology (e.g., sub-basin area, average slope). The spatial variables were derived from each drainage area. The seasonal variables $S(i, t)$ were split into two categories: hydrologic characteristics (e.g., $SW$: soil water content, $PET$: potential evapotranspiration, $GW$: ground water volume) and weather inputs (e.g., point specific precipitation and temperatures). The observed DOC, $y(i, t)$ corresponded with all sampling locations $i$, and the Julian days $t$ of field visits.
The observed values $y(i, t)$ for all $i$s and $t$s were fitted to eq. 1 for estimating the parameters $c, \alpha_1, \alpha_2, \ldots, \alpha_m, \beta_1, \beta_2, \ldots, \beta_n$. Statistical metrics used to evaluate relative importance and/or inclusion of variables into the models were $p$-value, coefficient of determination ($R^2$), and F values reflecting the overall significance of each regression model. An alpha of 0.05 was used as the threshold to determine if any one variable was statistically meaningful. The analysis included all exploratory variables (in varying combinations) within each category (temporal, spatial, physical, and biological). Our objective was to identify which variables were significant at various temporal scales (i.e., daily, monthly, or annually). Accordingly, all associated data had to be averaged when moving to a more course temporal scale. For example, ~30 daily hydrologic variables were averaged to yield a single monthly value. Similarly, ~365 daily hydrologic variables were averaged to yield an annual average. Ultimately, a variety of variables were fit to both raw and averaged (monthly and yearly) DOC observations to obtain the desired coefficients of the linear model.

One of our research objectives was to evaluate the appropriate spatial scale (i.e., contributing hydrologic areal extent) for quantifying the inherent relationship between a variety of independent variables and the variation of observed DOC concentrations from cropland areas. Here, we designated stream segment $i$ and sub-basin $i$ as the immediate stream segment and sub-basin in which the sampling location $i$ resides. With this designation, this study tested two spatial scale scenarios: 1st-order extended sub-basins and non-extended drainage areas. All 1st-order extended sub-basins met one of the following two criteria: (1) all sub-basins associated with sampling sites within a 1st-order stream, such as sampling locations 4, 6, 7, 9, 10, 11, 13, 16, 18, 19, and 20, or (2) all sub-basins associated with sampling sites within a 2nd or 3rd order stream if they had the same dominant land cover as all their
upper-tributary sub-basins. Sampling locations 5, 15, and 17 met this second criterion and were
designated as extended drainage sub-basins. For example, sub-basin 5 resides in a 2nd-order stream and
its dominant land cover is identical to that of its upper tributary sub-basin, which in this case was sub-
basin 4 (1st-order). Thus, the modelled drainage area associated with sampling location 5 included both
the areal extent of the upstream 1st-order sub-basin 4 and its immediate 2nd order sub-basin 5.
Therefore, the extended sub-basins in this project consisted of 14 sub-basins: 4, 6, 7, 9, 10, 11, 13, 16,
18, 19, 20, 5, 15, and 17. The remaining 7 sample locations (1, 2, 3, 8, 12, 14, and 21) are associated to
the non-extended drainage areas.

RESULTS AND DISCUSSION

Our overarching research objective was to establish a theoretical understanding of the relationship
and significance between a variety of independent, spatiotemporal variables and observed DOC
concentrations in cropland areas. We present pertinent results and related discussion within four distinct
sections: 1) Spatial disparity of mean annual DOC, 2) Seasonal/temporal patterns from the observations
for both croplands and forested lands, 3) Variables useful to estimate spatial and temporal distributions
of DOC concentrations in streamflow throughout the year, and 4) A quantitative analysis of mean
monthly DOC trends.

1. Spatial disparity of mean annual DOC

DOC concentrations ranged from 5.5 to 10.5 mg/L, and these values varied greatly across the
drainage sub-basins of the primary study area. Our results revealed that the DOC observations averaged
over the entire year for each sampling location were largely controlled by four spatial variables for the
1st-order extended areas. Specifically, a resultant linear regression model (eq. 2) explained more than
82% ($R^2_{adj} = 0.74$) of mean annual DOC, $\bar{y}(i)$, the concentrations observed across 14 sampling locations: 4, 5, 6, 7, 9, 10, 11, 13, 15, 16, 17, 18, 19, and 20 (Figure 3A).

$$f(X) = 0.52crp + 1.19frt + 0.52wetl + 18.74dum - 63.07$$ (2)

Where the first three variables (i.e., $crp$, $frt$, $wetl$) were percent of crops, forest, and wetland, and $dum$ was the binary dummy variable useful for separating crops and mixed land covers (crop = 1, mixed = 0). All coefficients for each of the four variables and the y-intercept were significant with respect to the observed DOC data (p-values << 0.05). The overall model explained a significant portion of variation in DOC values ($F = 10.22$, $p$-value = 0.0021). The dummy variable used in this modeling effort played a role as a constraining condition. The dummy variable expanded the capability of the linear regression for these complex scenarios of mixed land cover composition.

The positive coefficient for the variable $crp$ shown above indicates a positive correlation between DOC concentrations and crop density ($crp$), as one might expect given that crop density and the resulting residues usually increase DOC loadings (Tian et al., 2013). However, inspection of the 3-D graph in Figure 3B clearly contradicts such a positive relationship, for the highest DOC values were clearly associated with the low $crp$ values. This negative correlation seemed to be regulated by a geomorphological feature, $Ln(A)$ which is the logarithm of the drainage areas.

The percent of crop land cover ($crp$) is indeed positively correlated to $Ln(A)$, which was not included in eq. 2 for crop dominant areas because the $p$-value was slightly larger than the 0.05 threshold ($p = 0.0638$). Moreover, further inspection of the 3-D graph in Figure 3B also revealed that both $crp$ and $ln(A)$ have an inverse relationship to mean annual DOC. This result demonstrated the relative importance of accounting for geomorphological parameters ($A$) that appear to act at the scale of
individual sampling locations. Interestingly, $\ln(A)$ was included in a linear model (eq. 3), once the $p$-value threshold was relaxed to 90 percent ($p$-values < 0.1).

$$f(X) = 0.43crp + 1.03frt + 0.41wetl + 16.01dum - 0.45ln(A) - 46.09$$ (3)

Inclusion of the variable $\ln(A)$ demonstrates the viability of our spatially explicit modelling approach in eq. 3. This spatial variable was helpful in estimating mean annual DOC concentrations in runoff from croplands with varying levels of crop percentages across a larger region. It is logical to assume that its negative coefficient reflects the hydrologic processes occurring on and within the relatively larger and flatter (i.e., equated to higher percentage of crops) sub-basin. The alteration of hydrologic processes would result in the reduction of DOC concentrations because of the longer transport and residence time, allowing for greater soil carbon adsorption and infiltration. The second model (eq. 3) explained DOC observations better (88%, $R^2_{adj} = 0.69$, $p=0.001$) than the model (eq. 2) which did not directly reference this geomorphologic variable, drainage area.

The predictive power of our regression results (i.e., eq. 2 and eq. 3) highlight the viability of modelling mean annual DOC concentrations at the individual sub-basin level. Verification of model performance at the individual sub-basin level is indeed a necessary step before coupling multiple sub-basins into a more integrative model for use at a larger or regional scale. In other words, integrative modelling can indeed estimate DOC concentrations within high order stream segments by cumulatively aggregating or mixing DOC loadings from modelled upstream sub-basins (Tian et al. 2002). This integrative and hydrologically cumulative approach appears to better approximate the reality of how the many interconnected watershed components exchange materials from 1st-order to higher-order downstream receiving waters.
Both our statistical analyses and encouraging modelling outputs of DOC dynamics at the annual scale advances the scientific communities understanding of how anthropogenic agricultural land use changes and related management practices can help to explain spatial disparities of DOC concentrations originating from croplands. Our results also highlight the need for comprehensive in-situ sampling campaigns across different ecoregions and climatic zones to help fine tune regional and global models. The variables identified in eqs. 2 and 3 are extractable from satellite images, which helps bridge the gap from more localized modelling efforts to more regional and global spatial scales that take advantage of remote sensing technologies. Thus, this investigation verified the feasibility of using easily-obtainable satellite image data to detect DOC dynamics and to study the impacts of human activities and management policies on both land sustainability and the ecology of freshwater habitats.

2. *Seasonal DOC trends from crop and forest lands*

Strong seasonal patterns in DOC concentrations were inherent in our data, and the associated driving factors became readily apparent from the analysis of daily and monthly DOC export rates. The patterns in Figure 4A describe the DOC observations averaged in each month from all sampling locations that were in the 1st-order extended sub-basin scenario. Sampled peak DOC concentrations were not associated with runoff events, because we purposefully established our sampling periods to occur between rainfall and storm runoff events. Accordingly, modeled DOC concentration peaks were the result of seasonal phenomenon and not just a single episodic event. Generally, snowpack elevates soil water content during late winter and early spring (i.e., middle of March). It has been reported that high soil moisture content enhances certain saturation-dependent metabolic processes and biogeochemical reactions (Leakey et al. 2006) as well as increased accumulation of DOC in the topsoil (Ågren et al. 2006).
The seasonal peak observed in March corresponds well with baseflow resulting from snowmelt which acts to flush DOC stored in the topsoil created by these saturation-dependent metabolic processes and biogeochemical reactions. The high soil-water content (i.e., soil moisture) for the primary study site in late winter and early spring were predicted quite well by the simulation results of the SWAT hydrologic model (Figure 5). The SWAT model also accurately predicted the high soil water content associated with winter snowpack and its subsequent spring melting. The seasonal accuracy of the SWAT model stands to reason, given the promising calibration performance (NSE = 0.66, R² = 0.88) when referencing daily flow records for 38 years. As such, a slight spike in DOC concentrations (i.e., March) originating from these organic-rich croplands appeared to be triggered by the infiltration of snowmelt and elevated soil water content, coupled with the with flat topographic characteristics of the mixed drainage sub-basins (blue curve in figure 4A). The mixed drainage sub-basins have relatively less area for crop residue accumulation across multiple years, due to their relatively large percentage of forest and wetland land covers. Rich crop residue biomass is indeed the dominant anthropogenic source of geochemically active and transformable DOC. Spring is a critical time of the year with respect to benthic and shallow water habitat in relation to biological communities, aquacultural production, and water quality to public health (Lehosmaa et al. 2018; Puczko et al. 2018), which corresponds to spikes in DOC concentrations.

In contrast, an early spring DOC peak was not observed (Figure 4B) from the steep forested lands of the second study site despite similar weather conditions. Traditionally, farming practices for both sweet corn and dent corn leave most plant residues in the field, which greatly increases soil organic matter (Guo et al. 2018; Motavalli et al. 1992; Oberle and Keeney 1990). These accumulated residues enrich soil organic biomass more so than the leaf-litter fell from forest canopies in autumn (Du et al.
In addition, the increased slopes associated with these non-cropped, forested hills encouraged direct surface runoff with limited water infiltration through the soil profile. Less infiltration results in less DOC flux movements which corresponds to the lack of such a DOC peak in early spring for these forested locations (Figure 4B).

Other DOC peaks matched well between the two study sites (crop versus forest) from May to December (correlation > 0.65). We feel confident that these parallel DOC export curves are seasonal effects rather than of individual storm events since the curves were derived from a large set of multi-year, independent field measurements (i.e., 50 field visits). DOC spikes in May were triggered by increased soil temperatures followed by elevated ground water volumes (Figure 4). It seems logical that soils at depth take a longer time to warm up than the land surface temperature. Thus, the flux of DOC stored in such soil profiles were delayed until May at-depth soil temperatures began to rise (Figure 4A). High ground water levels and low soil water content were exhibited from the late April to early May as simulated via SWAT (i.e., black curve in Figure 5). Further examination of Figure 5 revealed that higher DOC export for croplands compared to forested areas was partially due to the crop farming practices, namely tillage in preparation for spring seeding. The optimum corn planting period for much of Michigan is from the beginning to middle of May, and soil manipulations would inherently rise as fields are prepped for planting. Soil temperature combined with the elevated soil moistures associated with summer rains may affect the DOC peaks in August for both study sites. Higher temperatures are associated with higher rates of microbial activities and faster turnover of DOC in the Carbon cycle of organic soils (Bowering et al. 2020; Haaland and Mulder 2010).
The mean monthly DOC concentration for croplands were more than 3-fold greater (Figure 4A) compared with those for forestlands (Figure 4B) throughout the year. The monthly precipitation and temperature patterns between the two study sites were very comparable. Therefore, the excessive soil DOC production rates from croplands were caused by anthropogenic effects in terms of crop residue accumulation and related break-down. Our previous study’s preliminary mesocosm experiments revealed that corn plants have a slower metabolic process for transforming foliar organic matter into the dissolved form when compared to deciduous (D) and evergreen (E) foliage (Li et al. 2018). Figure 6 illustrates how DOC production rate changed for the same applied biomass (260-g dry biomass), the same number of incubation days, and two different temperatures (L:20 °C or H:25°C) for these three vegetation types. Figure 6 illustrates that DOC transformation (i.e., foliar to dissolved) rates were often 2.5 times faster for forest foliage as compared to sweet corn foliage. Intuitively, the amount of organic biomass of croplands must therefore be 5-6 times higher than that of forested lands to result in the 3-fold greater DOC concentrations. Clearly both farming and crop residue practices cause increased DOC concentrations originating from croplands as opposed to climate or temperature variations.

3. Variables attribute to spatial and temporal DOC variations in croplands

Seven variables were identified as being statistically significant in driving DOC variation (eq. 5) in reference to finer spatial and temporal scales. Three out of these seven variables were related to seasonal effects that acted to quantify combined hydrologic and climatic properties: soil water content ($sw$ in mm), ground water volumes ($gw$ in mm), and potential evapotranspiration ($pet$ in mm). The remaining four variables were related to the spatial characteristics of any given sub-basin (i.e., $crp$, $frt$, $wetl$ and $dum$). Note that these spatial variables are the same as those found to be significant when analyzing
mean annual DOC dynamics. Specifically, for an area draining to sampling location \( i \), and at a particular Julian day, \( t \), seasonal variable set \( S \) and spatial variable set \( X \) are:

\[
X(i) = [\text{crp}(i), \text{frt}(i), \text{wetl}(i), \text{dum}(i)] \quad \text{for all } i = 4, 5, 6, 7, 9, 10, 11, 13, 15, 16, 17, 18, 19, 20
\]

\[
S(i, t) = [\text{sw}(i, t), \text{gw}(i, t), \text{pet}(i, t)] \quad \text{for all } t = \text{Julian day when had a field sampling}
\]

For the purpose of statistical data exploration, the linear regression was based on daily data \( X(i) \) for all \( i \) and \( S(i, t) \) for \( i = \text{gauge} \) and all \( t \) (Figure 5). The seasonal data is specified for the sub-basin that contained the USGS gauge station, since all \( S(i, t) \) are highly correlated (Correlations \( \geq 0.8 \)). Given these circumstances, \( S(\text{gauge}, t) \) changes with Julian days and does not change spatially for each sampling location. This linear regression analysis (eq. 4) explained approximately 54% of \( y(i, t) \), DOC concentrations for all samples (crop and non-crop, \( R^2_{\text{adj}} = 0.503 \), \( N=100 \)) as displayed in Figure 7 (both yellow and blue points). For crop dominant drainage sub-basins, the model explained approximately 50.44% \( R^2_{\text{adj}} = 0.46 \), \( N=63 \), yellow triangle points).

\[
f(X, S) = 0.46\text{crp} + 0.95\text{frt} + 0.51\text{wetl} + 15.1\text{dum} + 1.25\text{pet} + 0.046\text{sw} - 1.68\text{gw} - 59.96 \quad (4)
\]

All seven variables in eq. 4 tested significant (p-value \( << 0.05 \)). The coefficients of eq. 4 likely indicate several responsive processes. The coefficients of the spatial variables have similar value ranges to those associated with eq. 3, albeit with the exclusion of the variable \( \text{Ln}(A) \). The \( \text{Ln}(A) \) variable with eq. 3 was functionally replaced by the ground water variable \( \text{gw} \) in this finer spatiotemporal-scale statistical analysis (eq. 4). The negative coefficient for \( \text{gw} \) indicates an inverse relationship to DOC observations. Logically, the presence of significant amounts of ground water is likely strongly correlated to the size of the drainage area. The presence of relatively high groundwater amounts within larger sub-
basins likely acts to dilute soil-profile DOC concentrations. These four spatial variables help to inform what, where, and how land surface characteristics impact the variation of riverine DOC.

The three seasonal/temporal variables indicate that the timing (when) of increased DOC concentrations is inherently related to the potential harm of the receiving aquatic ecosystem if DOC exceeds ecosystem capacities. The coefficients of the three seasonal variables are nearly of equal importance in this model. The hydrologic properties represented by these seasonal variables were simulated well by the SWAT model. The modelled data values marked with light blue vertical lines in Figure 7 are shown with their associated Julian days of DOC sample collection. The higher snowmelt modelled with SWAT is consistent with the idea that these meltwaters infiltrate the soil and increase soil water content ($sw$) and thus carbon transport in the early months of the year (Figure 5).

The resultant coefficients and listed significant variables reflect underlying scientific rationales rather than pure empirical analysis. The results indeed provide substantial information and knowledge about the future development of a non-linear analytical model for exploring spatial and temporal variations of DOC export. This statistical analysis of the variation of daily DOC concentrations also implicated the necessity of extending the anticipated non-linear analytical model to incorporate seasonal trends that would ultimately be helpful in analyzing the long-term anthropogenic effects on receiving water DOC concentrations. For this purpose, we expand the statistical analysis to mean monthly DOC concentrations in the following section.

4. A quantitative analysis of mean monthly DOC trend

Modelled daily DOC concentrations were averaged for each month from all 10 sampling locations in which drainage sub-basins were crop dominated, 1st-order extended sub-basins. The mean monthly
modelled DOC concentrations explained about 81% \( (R^2 = 0.81) \) for the observation data set 1 (Ob1 in Figure 8) for six months (i.e., Jan., Mar., Apr., May, Nov., and Dec.). The months related to Ob1 were chosen because they represented data generated with more frequent field visits. The observation data set 2 (Ob2) were collected during summer months (i.e., Jun, Jul., Aug., and Sep.) and were limited in sample frequency. Therefore, the model explained less (49%) in correspondence analysis for the Ob2 months. In general, the model explained approximately 56% \( (R^2 = 0.56) \) of the observations (Ob1 and Ob2) throughout the year.

Monthly intervals are appropriate to both hindcast or forecast long-term impacts of climate change and human activity on carbon-cycling processes. Our analysis suggests that monthly DOC concentrations are indeed quantifiable and can provide an ideal baseline for developing next-generation analytical models. The month-based, baseline model can provide a platform for including event-based DOC fluctuations (Qiao et al. 2017). Such monthly, next-generation models can then be expanded to incorporate both the spatial and temporal variables influencing DOC dynamics at sub-basin scales. Ultimately, such models can integrate biological and physical processes to help identify where, when, and how DOC source and transport mechanisms respond to climatic and anthropogenic processes.

CONCLUSIONS

Our research reported several analysis results of DOC cycling processes from croplands to receiving waters. The spatiotemporal distributions of DOC concentrations were observed from a cropland representative of the Midwestern United States. The main contributions of our research efforts are summarized below.
The elevated DOC export rates in early spring were in response to snowpacks and snowmelt processes over croplands with corn plant residue accumulation.

Baseflow plays an important role in driving seasonal changes in DOC concentrations across spatial scales.

The 1st-order extended drainage sub-basins are indeed an appropriate spatial scale for quantifying the inherent relationship between independent variables and DOC concentrations in cropland areas.

Integrating an appropriate LULC reclassification, hydrological processes, and systematic analysis of geospatial and temporal features is indeed a viable approach for understanding DOC dynamics from agricultural landscapes with lowland hydrology.

Acknowledgements

This study is supported by a collaborative NSF Grant (1025547, PI: Q. Yu; 1025546, PI: Y.Q. Tian). We would like to thank Ashley Blunden (Lowran) for completing the bulk of the field sampling as part of her graduate work at CMU and Donald Uzarski for making his laboratory services available to process our DOC samples. All supporting data have been listed in a pdf file that can be downloaded at: http://people.se.cmich.edu/tian2y/onlinedata.pdf
REFERENCES

Ågren, A., Haei, M., Köhler, S., Bishop, K., & Laudon, H. (2010). Regulation of stream water dissolved organic carbon (DOC) concentrations during snowmelt; the role of discharge, winter climate and memory effects. *Biogeosciences*, 7, 2901-2913

Ågren, A., Jansson, M., Ivarsson, H., Bishop, K., & Seibert, J. (2008). Seasonal and runoff-related changes in total organic carbon concentrations in the River Öre, Northern Sweden. *Aquatic Sciences*, 70, 21-29

Aitkenhead-Peterson, J.A., McDowell, W.H., & Neff, J.C. (2003). Sources, production, and regulation of allochthonous dissolved organic matter inputs to surface waters. *Aquatic ecosystems* (pp. 25-70): Elsevier

Arnold, J.G., & Fohrer, N. (2005). SWAT2000: current capabilities and research opportunities in applied watershed modelling. *Hydrological Processes: An International Journal*, 19, 563-572

Bhattacharya, R., & Osburn, C.L. (2020). Spatial patterns in dissolved organic matter composition controlled by watershed characteristics in a coastal river network: The Neuse River Basin, USA. *Water Research*, 169, 115248

Bowering, K.L., Edwards, K.A., Prestegaard, K., Zhu, X., & Ziegler, S.E. (2020). Dissolved organic carbon mobilized from organic horizons of mature and harvested black spruce plots in a mesic boreal region. *Biogeosciences*, 17, 581-595

Butman, D., & Raymond, P.A. (2011). Significant efflux of carbon dioxide from streams and rivers in the United States. In, *Nature Geoscience* (pp. 839-842)

Clark, J.M., Lane, S.N., Chapman, P.J., & Adamson, J.K. (2007). Export of dissolved organic carbon from an upland peatland during storm events: Implications for flux estimates. *Journal of Hydrology*, 347, 438-447

Confesor Jr, R.B., & Whittaker, G.W. (2007). Automatic Calibration of Hydrologic Models With Multi-Objective Evolutionary Algorithm and Pareto Optimization 1. *JAWRA Journal of the American Water Resources Association*, 43, 981-989

Du, X., Zhang, X., Mukundan, R., Hoang, L., & Owens, E.M. (2019). Integrating terrestrial and aquatic processes toward watershed scale modeling of dissolved organic carbon fluxes. *Environmental pollution*, 249, 125-135

Dusek, J., Vogel, T., Dohnal, M., Barth, J.A., Sanda, M., Marx, A., & Jankovec, J. (2017). Dynamics of dissolved organic carbon in hillslope discharge: Modeling and challenges. *Journal of Hydrology*, 546, 309-325

Ellis, E.C., Klein Goldewijk, K., Siebert, S., Lightman, D., & Ramankutty, N. (2010). Anthropogenic transformation of the biomes, 1700 to 2000. *Global ecology and biogeography*, 19, 589-606

Fu, B., Merritt, W.S., Croke, B.F., Weber, T.R., & Jakeman, A.J. (2019). A review of catchment-scale water quality and erosion models and a synthesis of future prospects. *Environmental modelling & software*, 114, 75-97

Futter, M., Löfgren, S., Köhler, S., Lundin, L., Moldan, F., & Bringmark, L. (2011). Simulating dissolved organic carbon dynamics at the Swedish integrated monitoring sites with the integrated catchments model for carbon, INCA-C. *Ambio*, 40, 906-919

Gómez - Gener, L., Hotchkiss, E.R., Laudon, H., & Sponseller, R.A. (2021). Integrating Discharge - Concentration Dynamics Across Carbon Forms in a Boreal Landscape. *Water Resources Research*, 57, e2020WR028806
Guo, R., Qin, W., Jiang, C., Kang, L., Nendel, C., & Chen, Q. (2018). Sweet corn significantly increases nitrogen retention and reduces nitrogen leaching as summer catch crop in protected vegetable production systems. *Soil and Tillage Research*, 180, 148-153

Haaland, S., & Mulder, J. (2010). Dissolved organic carbon concentrations in runoff from shallow heathland catchments: effects of frequent excessive leaching in summer and autumn. *Biogeochemistry*, 97, 45-53

Heppell, C.M., Binley, A., Trimmer, M., Darch, T., Jones, A., Malone, E., Collins, A.L., Johnes, P.J., Freer, J.E., & Lloyd, C.E. (2017). Hydrological controls on DOC: nitrate resource stoichiometry in a lowland agricultural catchment, southern UK. *Hydrology and Earth System Sciences*, 21, 4785-4802

Holgerson, M.A., & Raymond, P.A. (2016). Large contribution to inland water CO2 and CH4 emissions from very small ponds. In, *Nature Geoscience* (pp. 222-226)

Jones, J., Knowlton, M., Obrecht, D., & Cook, E. (2004). Importance of landscape variables and morphology on nutrients in Missouri reservoirs. In, *Canadian Journal of Fisheries and Aquatic Sciences* (pp. 1503-1512)

Jones, J.R., & Knowlton, M.F. (2005). Suspended solids in Missouri reservoirs in relation to catchment features and internal processes. In, *Water Research* (pp. 3629-3635)

Kellerman, A.M., Arellano, A., Podgorski, D.C., Martin, E.E., Martin, J.B., Deuerling, K.M., Bianchi, T.S., & Spencer, R.G. (2020). Fundamental drivers of dissolved organic matter composition across an Arctic effective precipitation gradient. *Limnology and Oceanography*, 65, 1217-1234

Kelly, V., Stets, E.G., & Crawford, C. (2015). Long-term changes in nitrate conditions over the 20th century in two Midwestern Corn Belt streams. *Journal of Hydrology*, 525, 559-571

Leakey, A.D., Uribelarrea, M., Ainsworth, E.A., Naidu, S.L., Rogers, A., Ort, D.R., & Long, S.P. (2006). Photosynthesis, productivity, and yield of maize are not affected by open-air elevation of CO2 concentration in the absence of drought. *Plant physiology*, 140, 779-790

Lehosmaa, K., Jyväsjärvi, J., Ilmonen, J., Rossi, P.M., Paasivirta, L., & Muotka, T. (2018). Groundwater contamination and land drainage induce divergent responses in boreal spring ecosystems. *Science of the Total Environment*, 639, 100-109

Li, J., Yu, Q., Tian, Y.Q., & Boutt, D.F. (2018). Effects of landcover, soil property, and temperature on covariations of DOC and CDOM in inland waters. *Journal of Geophysical Research: Biogeosciences*, 123, 1352-1365

Lindström, G., Pers, C., Rosberg, J., Strömqvist, J., & Arheimer, B. (2010). Development and testing of the HYPE (Hydrological Predictions for the Environment) water quality model for different spatial scales. *Hydrology research*, 41, 295-319

Millie, D.F., Weckman, G.R., Pigg, R.J., Tester, P.A., Dyble, J., Wayne Litaker, R., Carrick, H.J., & Fahrenstiel, G.L. (2006). MODELING PHYTOPLANKTON ABUNDANCE IN SAGINAW BAY, LAKE HURON: USING ARTIFICIAL NEURAL NETWORKS TO DISCERN FUNCTIONAL INFLUENCE OF ENVIRONMENTAL VARIABLES AND RELEVANCE TO A GREAT LAKES OBSERVING SYSTEM 1. *Journal of phycology*, 42, 336-349

Monaghan, R., Wilcock, R., Smith, L., Tikkisetty, B., Thorrold, B., & Costall, D. (2007). Linkages between land management activities and water quality in an intensively farmed catchment in southern New Zealand. In, *Agriculture, Ecosystems & Environment* (pp. 211-222)

Motavalli, P., Bundy, L., Andraski, T., & Peterson, A. (1992). Residual effects of long-term nitrogen fertilization on nitrogen availability to corn. *Journal of production agriculture*, 5, 363-368
Oberle, S., & Keeney, D. (1990). Factors influencing corn fertilizer N requirements in the northern US corn belt. *Journal of production agriculture, 3*, 527-534.

Olaoye, I.A., Confesor, R.B., & Ortiz, J.D. (2021). Effect of projected land use and climate change on water quality of Old Woman Creek Watershed, Ohio. *Hydrology, 8*, 62.

Olsson, M.T., Erlandsson, M., Lundin, L., Nilsson, T., Nilsson, Å., & Stendahl, J. (2009). Organic carbon stocks in Swedish Podzol soils in relation to soil hydrology and other site characteristics. *Silva Fennica, 43*, 209-222.

Pers, C., Temnerud, J., & Lindström, G. (2016). Modelling water, nutrients, and organic carbon in forested catchments: a HYPE application. *Hydrological Processes, 30*, 3252-3273.

Puczko, K., Zieliński, P., Jusik, S., Kołakowska, A., & Jekatieryczuk-Rudczyk, E. (2018). Vascular plant and bryophyte species richness in response to water quality in lowland spring niches with different anthropogenic impacts. *Environmental monitoring and assessment, 190*, 1-16.

Qiao, H., Tian, Y.Q., Yu, Q., Carrick, H.J., Francek, M., & Li, J. (2017). Snowpack enhanced dissolved organic carbon export during a variety of hydrologic of events in an agricultural landscape, Midwestern USA. In, *Agricultural and Forest Meteorology* (pp. 31-41).

Raymond, P.A., & Saiers, J.E. (2010). Event controlled DOC export from forested watersheds. *Biogeochemistry, 100*, 197-209.

Rudolph, J.C., Arendt, C.A., Hounshell, A.G., Paerl, H.W., & Osburn, C.L. (2020). Use of geospatial, hydrologic, and geochemical modeling to determine the influence of wetland-derived organic matter in coastal waters in response to extreme weather events. *Frontiers in Marine Science, 7*, 18.

Sela, S., Woodbury, P., Marjerison, R., & van Es, H. (2019). Towards applying N balance as a sustainability indicator for the US Corn Belt: realistic achievable targets, spatio-temporal variability and policy implications. *Environmental Research Letters, 14*, 064015.

Sharratt, B., Benoit, G., & Voorhees, W. (1998). Winter soil microclimate altered by corn residue management in the northern Corn Belt of the USA. *Soil and Tillage Research, 49*, 243-248.

Singh, S., Inamdar, S., & Mitchell, M. (2015). Changes in dissolved organic matter (DOM) amount and composition along nested headwater stream locations during baseflow and stormflow. *Hydrological Processes, 29*, 1505-1520.

Spencer, R.G., Aiken, G.R., Dornblaser, M.M., Butler, K.D., Holmes, R.M., Fiske, G., Mann, P.J., & Stubbins, A. (2013). Chromophoric dissolved organic matter export from US rivers. *Geophysical Research Letters, 40*, 1575-1579.

Stedmon, C.A., Markager, S., Søndergaard, M., Vang, T., Laubel, A., Borch, N.H., & Windelin, A. (2006). Dissolved organic matter (DOM) export to a temperate estuary: seasonal variations and implications of land use. *Estuaries and Coasts, 29*, 388-400.

Tian, Y.Q., Gong, P., Radke, J.D., & Scarborough, J. (2002). Spatial and temporal modeling of microbial contaminants on grazing farmlands. *Journal of Environmental Quality, 31*, 860-869.

Tian, Y.Q., Wang, D., Chen, R.F., & Huang, W. (2012). Using modeled runoff to study DOC dynamics in stream and river flow: A case study of an urban watershed southeast of Boston, Massachusetts. *Ecological Engineering, 42*, 212-222.

Tranvik, L. (2014). Carbon cycling in the Arctic. In, *Science* (pp. 870-870).
Vetsch, J.A., & Randall, G.W. (2002). Corn production as affected by tillage system and starter fertilizer. *Agronomy Journal, 94*, 532-540

Vlahos, P., Chen, R.F., & Repeta, D.J. (2002). Dissolved organic carbon in the Mid-Atlantic Bight. *Deep Sea Research Part II: Topical Studies in Oceanography, 49*, 4369-4385
This study is supported by a collaborative NSF Grant (1025547, PI: Q. Yu; 1025546, PI: Y.Q. Tian).
Author Contributions

1. Research concept, design, and question
   All authors

2. Funding:
   Yong Tian, Qian Yu, Hunter Carrick

3. Material preparation, and data collection
   Yong Tian, Qian Yu, Hunter Carrick, Brain Becker, and Remegio Confesor

4. Data analyses
   Yong Tian, Qian Yu, and Anderson, C. Olivia

5. Manuscript preparation and Final Manuscript Approval
   All authors
Figure 1, The primary study area and landcover. It consists of 21 sampling locations established at sub-basin outlets of the drainage sub-basins. Black polygons are the drainage sub-basins. The blue vector data are the streams. (Color needed)

Figure 2, Original national land cover classification codes (NLCD) and three newly formed classes.

Figure 3, A) the linear regression estimation against the DOC observations, and B) both Crop% and Ln(A) have inverse correlations to DOC concentrations, yet the correlation between Crop% and Ln(A) is positive.

Figure 4, A): Mean monthly DOC concentrations for Crop and mixed land covers over the primary study site. Note that February DOC concentrations were interpolated by using the average value of adjacent months. B): Monthly mean DOC concentrations of five sub-basins in the Connecticut River Watershed collected in years 2011-2016. Note that February and July data were interpolated by averaging data from the jacent months. In general, mean monthly DOC from croplands in the Chippewa River Watershed are highly correlated (0.65) to that from forested area in the Connecticut River Watershed.

Figure 5, SWAT modelled daily PET, SW, SN, GW and SM for sub-basin 49 for year 2013. The sub-basin 49 contains the USGS gage station. The dashed light blue vertical lines indicate the sampling date. (PET: Potential Evapotranspiration, SW: Soil Water content, GW: ground water, and Snowmelt: Snow Liquid Equivalent).
Figure 6, DOC production rates of three vegetation types: corn plant (AH/AL), broad leaf liters (DH/DL) and conifer leaf liters (EH/EL). The L stands for low temperature (20°C) and H is the high temperature (25°C).

Figure 7, the modelled (Y axis) versus observed (X axis) DOC concentrations for all data points collected over enhanced-area landscapes (N=100), and for crop dominant drainage areas (yellow triangle points, N = 63).

Figure 8, Modelled (mod in red) versus observed mean monthly DOC concentrations for crop dominant drainage areas over 1st-order drainage sub-basins. Observation data set 1 (Ob1) is for Jan, Mar, Apr, May, Oct, Nov, and Dec, R2=0.81. Observation data set 2 (Ob2) is for Jun, Jul, Aug, and Sep, R2 = 0.49. For combined Ob1 & Ob2 in all 12 months together, R2 > 0.56.

Table 1, Land Use Composition (%) in the area (Cells) draining to each sample location (SID). Crop%: percent of crop coverage, Fort%: Forest, Wetl%: wetlands, Dev%: developed, and wat%: Water body.
Figure 2, The primary study area and landcover. It consists of 21 sampling locations established at sub-basin outlets of the drainage sub-basins. Black polygons are the drainage sub-basins. The blue vector data are the streams.
Figure 3, Original national land cover classification codes (NLCD) and three newly formed classes.
Figure 4, A) the linear regression estimation against the DOC observations, and B) both Crop% and Ln(A) have inverse correlations to DOC concentrations, yet the correlation between Crop% and Ln(A) is positive.
Figure 5, A): Mean monthly DOC concentrations for Crop and mixed land covers over the primary study site. Note that February DOC concentrations were interpolated by using the average value of adjacent months. B): Monthly mean DOC concentrations of five sub-basins in the Connecticut River Watershed collected in years 2011-2016. Note that February and July data were interpolated by averaging data from the adjacent months. In general, mean monthly DOC from croplands in the Chippewa River Watershed are highly correlated (0.65) to that from forested area in the Connecticut River Watershed.
Figure 6, SWAT modelled daily PET, SW, SN, GW and SM for subbasin 49 for year 2013. The subbasin 49 contains the USGS gage station. The dashed light blue vertical lines indicate the sampling date. (PET: Potential Evapotranspiration, SW: Soil Water content, GW: ground water, and Snowmelt: Snow Liquid Equivalent).
Figure 7, DOC production rates of three vegetation types: corn plant (AH/AL), broad leaf liters (DH/DL) and conifer leaf liters (EH/EL). The L stands for low temperature (20°C) and H is the high temperature (25°C).
Figure 8, the modelled (Y axis) versus observed (X axis) DOC concentrations for all data points collected over enhanced-area landscapes (N=100), and for crop dominant drainage areas (yellow triangle points, N = 63).
Figure 9, Modelled (mod in red) versus observed mean monthly DOC concentrations for crop dominant drainage areas over 1st-order drainage subbasins. Observation data set 1 (Ob1) is for Jan, Mar, Apr, May, Oct, Nov, and Dec, $R^2=0.81$. Observation data set 2 (Ob2) is for Jun, Jul, Aug, and Sep, $R^2 = 0.49$. For combined Ob1 & Ob2 in all 12 months together, $R^2 > 0.56$. 
Table 1, Land Use Composition (%) in the area (Cells) draining to each sample location (SID).
Crop%: percent of crop coverage, Fort%: Forest, Wetl%: wetlands, Dev%: developed, and wat%: Water body.

| ID | Samples # | DOC Mg/L | Cells   | Ln(Cells) | Crop% | Forest% | Wetland% | Developed% | Hay% | Water% |
|----|-----------|----------|---------|-----------|-------|---------|----------|------------|------|--------|
| 1  | 7         | 6.899    | 259437  | 12.47     | 64    | 16      | 13       | 5          | 3    | 0      |
| 2  | 7         | 5.864    | 1402613 | 14.15     | 38    | 29      | 18       | 8          | 4    | 2      |
| 3  | 6         | 6.009    | 1435616 | 14.18     | 37    | 16      | 15       | 29         | 2    | 1      |
| 4  | 6         | 7.300    | 26314   | 10.18     | 78    | 3       | 12       | 7          | 0    | 0      |
| 5  | 6         | 6.548    | 61578   | 11.03     | 80    | 3       | 11       | 6          | 0    | 0      |
| 6  | 6         | 5.739    | 38201   | 10.55     | 81    | 3       | 8        | 7          | 0    | 0      |
| 7  | 6         | 6.503    | 176446  | 12.08     | 75    | 4       | 14       | 6          | 1    | 0      |
| 8  | 7         | 6.263    | 1915023 | 14.47     | 39    | 18      | 26       | 14         | 3    | 1      |
| 9  | 6         | 5.516    | 113370  | 11.64     | 86    | 2       | 3        | 8          | 1    | 0      |
| 10 | 7         | 6.006    | 86448   | 11.37     | 87    | 2       | 6        | 5          | 0    | 0      |
| 11 | 6         | 5.890    | 166695  | 12.02     | 75    | 5       | 13       | 7          | 0    | 0      |
| 12 | 6         | 6.351    | 46173   | 10.74     | 26    | 24      | 33       | 12         | 5    | 0      |
| 13 | 6         | 6.274    | 41969   | 10.64     | 44    | 12      | 24       | 17         | 1    | 1      |
| 14 | 10        | 5.508    | 1079820 | 13.89     | 30    | 36      | 20       | 6          | 5    | 3      |
| 15 | 10        | 6.166    | 966081  | 13.78     | 27    | 37      | 21       | 6          | 6    | 3      |
| 16 | 10        | 8.736    | 327472  | 12.70     | 16    | 42      | 26       | 5          | 8    | 3      |
| 17 | 9         | 7.159    | 550234  | 13.22     | 20    | 40      | 24       | 5          | 7    | 4      |
| 18 | 10        | 6.972    | 62698   | 11.05     | 38    | 34      | 20       | 5          | 4    | 0      |
| 19 | 4         | 10.057   | 13910   | 9.54      | 66    | 10      | 15       | 6          | 3    | 0      |
| 20 | 8         | 9.704    | 91906   | 11.43     | 26    | 39      | 24       | 4          | 7    | 0      |
| 21 | 10        | 8.318    | 244838  | 12.41     | 59    | 17      | 16       | 5          | 3    | 0      |