

Article Multi-Scale Analysis of the Dependence of Water Quality on Land Use Using Linear and Mixed Models

Johann Alexander Vera Mercado D and Bernard Engel *

Department of Agricultural & Biological Engineering, Purdue University, West Lafayette, IN 47906, USA; vera3@purdue.edu

* Correspondence: engelb@purdue.edu

Abstract: Land use influences water quality in streams at different spatial scales and varies in time and space. Water quality has long been associated with agricultural and urban land uses in catchments. The effects of developed, forest, pasture, and agricultural land use on nitrogen, nitrate, and nitrite (NNN); total phosphorus (TP); total suspended solids (TSS); chemical oxygen demand (COD); dissolved oxygen (DO) and total Kjeldahl nitrogen (TKN) concentrations and their sensitivity were quantified to spatial pattern differences. The linear mixed modeling framework was used to examine the importance of spatial extent on models with water quality parameters as the response variable and land use types as the predictor variable. The results indicated that land use categories on different water quality parameters were significant and dependent on the selected spatial scales. Land use exhibited a strong association with total phosphorus and total suspended solids for close reach distances. Phosphorus is not highly soluble, and it binds strongly to fine soil particles, which are transported by water via runoff. Nitrogen, nitrate, and nitrite, dissolved oxygen, chemical oxygen demand, and total Kjeldahl nitrogen concentrations were better predicted for further reach distances, such as 45 or 50 km, where the best model of nitrogen, nitrate, and nitrite is consistent with the high mobility of NO₃⁻.

Keywords: land use; linear mixed model; spatial scale effects; water quality

1. Introduction

The quality of water supply is usually related to land use within a catchment area as it determines the quality and quantity of runoff before and after a rainfall event [1]. Hence, most water pollution problems are caused by changes in the distribution of land use and its management within a catchment as human activities increase. The main factors behind the alteration of the hydrological system are changes in land cover and land management practices, causing changes in runoff as well as the quality of receiving water [2,3].

Watershed factors, such as land use, landscape, and physiographical characteristics, may have a direct impact on the quality of river water [4,5]. The shifts in urban land use and land cover influence spatial and temporal runoff trends, which in turn impacts surface water quality [6]. Changing land use and land management practices are therefore regarded as central factors in altering the hydrological system, causing changes in runoff and water quality [7].

Land use changes in agricultural, forest, marsh and developed areas also affect soil quality, nutrient fluxes, soil surface temperature and native species assemblies, and these changes can influence watershed hydrology by altering interception, drainage, evapotranspiration, and groundwater recharge levels, resulting in changes in timing and surface and river runoff amounts [8,9]. Land uses affect water quality, whether positive or negative; for example, in forests and other areas with different types and coverage of vegetative surfaces, most precipitation infiltrates into the soil rather than running off. As a result, stream flows are steady, and water quality is excellent. In developed zones with asphalt



Citation: Vera Mercado, J.A.; Engel, B. Multi-Scale Analysis of the Dependence of Water Quality on Land Use Using Linear and Mixed Models. *Water* 2021, *13*, 2618. https://doi.org/10.3390/w13192618

Academic Editor: Sophia Barinova

Received: 31 August 2021 Accepted: 20 September 2021 Published: 23 September 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https:// creativecommons.org/licenses/by/ 4.0/).



and structures, little precipitation infiltrates into the soil, causing high runoff, stream flows with high peaks and low flows in between, and reduced water quality.

Recent studies in landscape ecology, landscape planning, and urban design have examined the relationship between land use and water quality parameters at different scales [10–13]. Some studies have examined the impacts of land use/cover changes in buffer zones and whole catchments on river water quality of the rivers [14–16]. In a study conducted in Greece, Gamvroula et al. [17] used hydrochemical analysis, statistical analysis and a GIS database to explain the main factors and mechanisms that control the distribution of major and trace elements in groundwater. Other studies have used sub-basins to compare water quality parameters in areas dominated by different land use types [18,19].

The composition of land use is typically related to water quality within a watershed, and landscape arrangements can be responsive water quality predictors [12]. Geographic Information System (GIS) technology and landscape ecology have helped quantify land use configurations and soil properties [12,20]. Land use metrics are sensitive to change with spatial patterns, and scale dependence is referred to as a spatial extent [19]. Land use itself has little impact on the nature of river water; however, categories of land use may reflect the spatial variety of anthropogenic activities responsible for producing and transporting pollutants [19]. For example, agricultural land uses are sources of sediments, nutrients, and pesticides. On the other hand, forest land works as a barrier to minimize erosion, thus reducing sediments in water bodies, and storing water pollutants or removing them. Riparian buffers can substantially reduce nutrients, surface sediments, and the flow of groundwater through deposition, adsorption, and denitrification processes [21].

Three spatial scales have been frequently used to link landscape variables to stream water quality, including reach, catchment and riparian [22,23]. A distance of several hundred km usually describes "reach" from the outlet. The reach size varies based on the size and shape of the watershed and the location of the sampling station. The riparian zone is often defined as the buffer with a similar or larger distance from the stream compared to reach extending to the entire upstream area above the study site. The catchment is the whole watershed above the sampling station or interest point [24]. However, the findings have not reliably determined which spatial dimension is the best indicator of water quality. Pollutants from agriculture or urban areas may be dissolved in the stream or stored in the soil before reaching the sampling station, providing an incorrect measure of the relationship between the land cover and water quality parameters [25]. Some studies suggested that the entire catchment was critical in determining the effects of anthropogenic activities or the quality of water [11,26,27]. Other studies have found that land use at the reach or riparian buffer better accounted for water quality variability [28–30]. Thus, more research is needed to quantify the impact of different land use scales on water quality in order to improve water resource management [12].

The primary nutrients are nitrogen and phosphorus that, in excessive amounts, pollute lakes, streams, and wetlands in the United States [31]. Farmers use many sources of nitrogen for crop production, and nitrogen is a critical nutrient for plant growth, but too much nitrogen in the water can lead to eutrophication of streams and lakes. High concentrations of nitrate in drinking water due to fertilizer, domestic effluent, sewage disposal, and industrial discharges are toxic to humans [32–34]. Phosphorus results in overproduction of algae, which obstruct lakes and reservoirs. Nutrient sources in agricultural lands include chemical fertilizers, livestock manure, and septic systems. Nutrient sources in urban areas are fertilizers used on lawns, domestic detergents, gardens, pet waste, and discharge from sewage systems [35].

The watersheds selected for the study are located within the Wabash River watershed. The Wabash River in Indiana is the largest free-flowing tributary of the Ohio River. It has a drainage basin area of 85,500 km², draining most of Indiana [36]. The mean population density in the watershed was 62.7 people/km² in 2000 [37]. Approximately 5% of the Wabash River basin is urban [37]. The Wabash Basin has about 2.5 million people using

1,900,000 m³ of surface water and groundwater each day. About 1 million people on farms and in small towns use 650,000 m³ each day, mostly groundwater [37]. Agricultural products that come from the basin include corn, soybeans, wheat, hay, vegetables, fruit, dairy products, livestock, and forest products. The hydraulic conditions of the Wabash River have been modified due to development within the basin, construction of lakes, and flood protection projects adjacent to the streams and rivers to protect the various communities.

The Wabash River has been steadily polluted since the settlement period, largely due to agricultural growth and increased human impacts. These past patterns appeared to be reversed in 1984, when an immediate and dramatic change was established in the middle west, the Wabash fishery, which continues [36,38]. In a study of riparian wetlands in Indiana streams, Gammon proposed that these improvements were due to slow but accumulated point-source decrease in the biological oxygen demand loading of the river [38]. Several negative impacts of agricultural activities persist, including silting, rapid drainage due to the tiling of fields, and fertilizer and pesticide inputs. Loss of wetlands and natural riparian zones has increased flooding, and deforestation has increased riverbank erosion, sediments, and contaminants from agricultural fields and urban areas have contributed to the Gulf of Mexico "Hypoxic Zone".

Regression models have been used to determine the relationship between land use and water quality parameters. A comprehensive study that investigated the relationship of nitrate concentrations in streams to row crop land use in Iowa found that the nitrate concentration decreases with watershed size [39]. In another study, ordinary least squares regression models were applied to examine the spatially varying relationship between landscape metrics and topography with water quality parameters [23]. Traditional regression models assume that "random variables" must be independently and equally distributed; however, these assumptions are not frequently satisfied in environmental studies [39].

Linear Mixed Models (LMM) are a natural extension of classical linear regression models to allow for the incorporation of random effects. LMM provide the ability to handle non-probability-based sampling schemes and spatial correlation structures between samples [40,41]. The linear mixed model has been used to estimate water quality using turbidity and discharge [41], water quality, and landscape characteristics at watershed scales [34,42].

The objective of this research is to examine the relationship between land-use and water quality for Wabash River watershed sub-basins. The scenarios explored are reach distance, buffer width and catchment using Linear Regression Modeling and LMM using as fixed effects the land use, and as Random Effects the Site, Year and Site*Year, addressing the following questions: 1. At what spatial scale does landscape heterogeneity act to influence water quality, and 2. How do spatial and temporal variation in land use within and across watersheds influence water quality pollutants in the basins inside of the Wabash River watershed. The model results can provide insight for categorizing watersheds and offer information that might assist in improving local land use and water quality management.

2. Study Area

The Wabash River and its tributaries generally exhibit moderate water quality; the Wabash River is the largest southward-flowing tributary of the Ohio River and has a watershed area of 85,340 km² [35]. Decades of consuming and developing land in the Wabash River Watershed have degraded the quality of its waters. By 1960, 48% of Indiana was subject to artificial drainage [43]. The topography ranges from 98.76 m (324 feet) above sea level in the southwest corner of the state at the mouth of the Wabash River to 383.14 m (1257 feet) in the extreme east-central portion of the state. The average monthly temperature in Indiana ranges from -1 °C in January to around 24 °C in July [43]. Precipitation varies from 92 to 112 cm of rainfall from north to south, and 41 to 187 cm of snow, and snowfall contributes 5–18 cm of average annual precipitation [44].

2.1. Watershed Characteristics

About two-thirds of the Wabash River watershed is devoted to agricultural cropland, and an additional 8.2% is in pasture or grassland. Forests or woodlands constitute only 13.5% of the basin's land area [36]. The study focuses on 17 sub-basins inside the Wabash Basin (Figure 1). Percentage of developed land in a watershed was defined as the sum of NLCD (National Land Cover Database) low-, medium-, high-intensity, and open space land-cover classes. The deciduous forest as forest, pasture/hay as pasture, cultivated crops as agriculture, and rest as the sum of open water, barren, shrubland, and herbaceous (Figure 2). The land cover percentages were calculated within watersheds.



Figure 1. Sub-basins inside of Wabash River Watershed.



Figure 2. Diagram of some spatial scales including (**a**–**d**) reach from the outlet, (**e**) entire watershed, and (**f**–**h**) riparian buffers for watershed WAW030_0022.

Cultivated crops account for 79.76% of land uses inside of the Wabash Basin mostly for corn and soybean, 8.88% are developed areas in general, 6.14% are forest, and pasture constitutes only 2.96% of the sub-basin average areas in 2001 (Table 1). The "Path" column in Table 1 shows the longest flow path within each sub-basin. The sub-basin areas of the study sites vary several orders in magnitude (85–1334 km²). Land use composition changes across watersheds, such as developed land varies between 5.5 to 14.7%, forest 1.8 to 25.4%, pasture 0.0 to 8.5% and agriculture land 60.2 to 88.3% [45], with highest differences in agriculture land and the smallest in pasture land in 2001 (Table 1).

		Watershed Area	Maximum Path Length	Land Use (%)			
Ν	Station	(km ²)	(km)	Developed	Forest	Pasture	Agriculture
1	WAW030_0022	239	45.6	6.2	4.1	0.8	87
2	WTI150-0011	1334	107.6	7.1	6.7	1.7	82.4
3	WAW010-0063	501	68.9	7.7	2.2	0	88.3
4	WAW040-0001	190	34.2	12.1	2.01	0.13	81.5
5	WDE050-0002	707	87.4	6.76	3.71	1.48	86.41
6	WLV160-0001	459	67.3	5.5	8.1	5.4	79.6
7	WWU080-0002	324	50.2	8.1	1.8	0.4	88.2
8	WWU100-0001	484	58.9	14.3	6.25	4	72.28
9	WWU120-0007	303	36.9	14	6.2	8.5	68.4
10	WSU020-0003	733	61.6	8.4	2.95	1.79	85.01
11	WED090-0004	1223	138.3	14.7	6.2	3.3	74.2
12	WEF020-0002	314	53.9	6.4	4.5	3	85.3
13	WEU030-0011	646	102.3	7.4	25.4	5.9	60.2
14	WWU010-0006	85	31.3	8	5.5	2.29	82.76
15	WMI020-0002	339	43.5	6.7	5	2.8	84.5
16	WSA010-0002	223	32	8.3	9.4	2.4	78.5
17	WUW040-0005	791	74.3	9.3	4.3	6.5	71.4
	Average	509	64.4	8.88	6.14	2.96	79.76

Table 1. Land use percentages [45] in 2001 for the study watersheds at the water quality stations.

2.2. Data

The Digital Elevation Model (DEM), expressed as a 30-m raster from USGS, was used to delineate upstream watersheds using water quality station locations. The National Hydrological Dataset Plus [46] was used to characterize the stream network. The NHD dataset was used to improve hydrologic boundary delineation [47]. Water quality data [48] from 2001 to 2016 were obtained from the United States Environmental Protection Agency's STORET (Table 1), and they were aggregated yearly using the arithmetic mean. Land cover percentages were summarized from the National Land Cover Database (NLCD), and a raster dataset developed from 30-m data from 2001, 2006, 2011, and 2016. After watershed delineation, the maps were overlaid with the NLCD in ArcGIS, and land-cover class percentages were extracted and calculated for each sub-watershed area for the years 2001 (Table 1), 2006, 2011 and 2016, and the land-cover percentages of the missing years were calculated using linear extrapolation.

Riparian buffers were defined as the width of n meters on each side of the stream center extended to the entire upstream above all water quality stations [12]. Reach distance areas were defined as the area from the quality station upstream using the flow-length concept, which was calculated using the Flow Length tool in ArcGIS to calculate the flow path to the outlet within the basin at 2, 5, 10, 25, 35, 40, 45, and 50-km constrained by watershed boundaries (Figure 2a–e). The flow length is calculated using the eight-directional algorithm described by [49] O'Callaghan and Jenson [50] and implemented in the ARC/INFO GIS system (ESRI, Redlands, CA, USA). Riparian buffers were delineated using NHD stream information, and buffer analysis in ArcGIS was used to create areas at 200, 500, and 1000-m from streams (Figure 2f–h); watershed boundaries further constrained all buffer analyses.



The watershed WAW030_0022 is shown in Figure 3, and land uses for the reach distance at 25 km. Agriculture covers most of the reach distance at 25 km, and the remaining area of the figure shows the riparian buffer at 100 m.

Figure 3. Schematic diagram of three spatial scales: 1. Reach (left side), 2. Riparian (right side), 3. Watershed (brown), for WAW030_0022.

Table 2 shows the land use percentages for different reach distances in the year 2001 for watersheds WAW030-0022, WDE050-0022, and WSU020-0003. At water quality station WAW030-0022, forest area percentage decreases from a 2 km to 45 km reach distance; the 45 km reach distance has a similar forestry land use percentage value to that for the catchment. At water quality station WSU020-0003, forest decreases from a reach distance of 2 km to the catchment, and developed increases from the reach distance of 10 to 50 km (Table 2).

Table 2. Land use percentages [45] for different reach distances at the water	quali	ty stations in the	year 2001
---	-------	--------------------	-----------

Ν	Station	Scale (km)	Developed (%)	Forest (%)	Pasture (%)	Agriculture (%)	Other (%)	Area (%)
1	WAW030-0022	watershed	6.2	4.1	0.8	87	1.9	100
		2	6.5	22.6	1	67.44	2.39	100
		10	10.31	7.32	1.14	78.36	2.87	100
		25	6.84	5.14	1.12	84.85	2.05	100
		35	6.37	4.45	0.98	86.21	1.99	100
		40	6.21	4.18	0.87	86.86	1.88	100
		45	6.17	4.05	0.84	87.05	1.89	100
5	WDE050-0002	watershed	6.76	3.71	1.48	86.41	1.64	100
		2	2.89	34	12.55	44.23	6.33	100
		10	5.1	17.6	4.76	69.94	2.63	100
		25	7.3	5.79	1.87	83.36	1.68	100
		35	6.58	5.33	1.24	85.24	1.61	100
		40	6.42	5.22	1.5	85.32	1.54	100
		45	6.25	5.17	1.63	85.31	1.64	100
		50	6.13	4.91	1.64	85.67	1.65	100
10	WSU020-0003	watershed	8.4	2.95	1.79	85.01	1.85	100
		2	8.28	24.7	1.09	60.11	5.87	100
		10	6.18	5.17	0.66	84.2	3.79	100
		25	7.32	5.09	2.87	81.93	2.79	100
		35	7.42	4.38	3.19	82.52	2.49	100
		40	8.87	4.25	2.97	81.52	2.39	100
		45	9.9	4.13	2.61	81.04	2.32	100
		50	9.21	3.58	2.14	82.93	2.14	100

7 of 19

Six water quality parameters (Table 3), between 2001 and 2016 were selected for the 17 study watersheds based on availability and importance to human and aquatic life. Chemical oxygen demand (COD) and dissolved oxygen (DO) were selected, as they are related to the estimation of the probability of algae bloom and habitat quality for fish and other aquatic species. Total solids (TS) were chosen as they can be used as a quantitative measure of aesthetics as suspended sediments in streams make the water appear cloudy [25]. Nitrogen, nitrate, and nitrite (NNN) and total phosphorus (TP) are generally considered direct measures of human activity in an area due to their presence in fertilizers, vehicle emissions, and impervious surfaces [25]. Total Kjeldahl Nitrogen (TKN) is the total concentration of organic nitrogen and ammonia.

Table 3. Main characteristics of study drainage areas and their water quality [48].

Ν	Quality Station	Area (km ²)	COD (mg/L) (Mean \pm sd)	DO (mg/L) (Mean \pm sd)	NNN (mg/L) (Mean \pm sd)	TP (mg/L) (Mean \pm sd)	TKN (mg/L) (Mean \pm sd)	TSS (mg/L) (Mean \pm sd)	# Obs
1	WAW030_0022	239	15.39 ± 4.05	9.82 ± 0.59	4.07 ± 1.07	0.14 ± 0.05	0.65 ± 0.17	52.71 ± 58.47	185
2	WTI150-0011	1334	21.41 ± 2.67	10.29 ± 0.99	7.66 ± 8.68	0.08 ± 0.019	0.85 ± 0.11	13.41 ± 4.93	188
3	WAW010-0063	501	20.06 ± 2.95	9.07 ± 0.63	3.98 ± 1.1	0.12 ± 0.03	0.84 ± 0.11	20.23 ± 6.83	180
4	WAW040-0001	190	17.86 ± 4.33	9.24 ± 0.58	6.27 ± 1.17	0.55 ± 0.17	0.92 ± 0.24	34.03 ± 32.02	192
5	WDE050-0002	707	16.19 ± 4.65	10.66 ± 0.77	4.42 ± 1.10	0.12 ± 0.07	0.64 ± 0.22	39.15 ± 43.70	187
6	WLV160-0001	459	14.13 ± 4.01	10.41 ± 0.54	3.16 ± 0.88	0.10 ± 0.03	0.57 ± 0.15	35.13 ± 14.93	190
7	WWU080-0002	324	16.04 ± 2.95	9.73 ± 0.47	5.23 ± 1.27	0.09 ± 0.03	0.64 ± 0.11	25.77 ± 14.49	184
8	WWU100-0001	484	13.56 ± 3.60	9.35 ± 0.41	2.50 ± 0.32	0.08 ± 0.02	0.49 ± 0.11	22.17 ± 10.51	199
9	WWU120-0007	303	16.28 ± 3.79	9.81 ± 0.39	3.36 ± 0.71	0.11 ± 0.02	0.67 ± 0.07	29.86 ± 31.51	190
10	WSU020-0003	733	16.70 ± 3.82	9.83 ± 0.44	4.55 ± 1.45	0.19 ± 0.08	0.70 ± 0.18	39.34 ± 29.82	188
11	WED090-0004	1223	13.59 ± 2.95	10.01 ± 0.27	2.61 ± 0.40	0.13 ± 0.03	0.55 ± 0.12	30.93 ± 25.98	195
12	WEF020-0002	314	12.92 ± 3.50	9.76 ± 0.37	4.66 ± 0.87	0.09 ± 0.03	0.52 ± 0.1	22.42 ± 12.41	195
13	WEU030-0011	646	17.36 ± 3.32	9.64 ± 0.44	2.78 ± 0.59	0.13 ± 0.04	0.70 ± 0.13	51.40 ± 34.31	191
14	WWU010-0006	85	16.73 ± 3.90	10.38 ± 0.61	2.70 ± 0.68	0.12 ± 0.03	0.62 ± 0.13	21.13 ± 10.89	187
15	WMI020-0002	339	19.93 ± 4.42	9.67 ± 0.59	4.28 ± 0.99	0.20 ± 0.06	0.90 ± 0.21	42.45 ± 26.35	194
16	WSA010-0002	223	28.62 ± 5.46	9.27 ± 0.71	4.55 ± 1.06	0.46 ± 0.14	1.39 ± 0.31	69.79 ± 42.05	169
17	WUW040-0005	791	43.18 ± 8.70	9.72 ± 0.87	4.97 ± 0.57	0.45 ± 0.14	2.47 ± 0.71	107.57 ± 77.8	138

2.3. Spatial Analysis

To determine the association between land use and water quality at each water quality station site, the riparian buffers and reach areas from the outlet were used to extract land use information from the land use raster (Figure 2). Using GIS tools, areas were determined for each land cover within each scenario (a riparian buffer area or reach distance area) divided by the entire covered area to derive the percentage of the total area covered by each riparian buffer or reach distance area. Land use map features for years 2001, 2006, 2011, and 2016 were extracted, cutting out data using the area boundaries.

2.4. Statistical Analysis

Average yearly water quality data were calculated for all the stations to reduce the impact of missing values. The normality of the dependent variables was checked using the one-sample Kolmogorov–Smirnov (KS) test. All water quality measurements were found to lack normality with 95% confidence. Consequently, the dependent variables were transformed using the Box–Cox technique to find adequate power transformation (or logarithmic) that produces a normal distribution with homoscedastic residuals.

The relationship between land use and water quality inside the Wabash River Watershed sub-basins was examined using multiple linear regression (MLR) and LMM. The MLR models were fitted to evaluate the relationship between a response (i.e., single water quality parameter) and predictors (i.e., land use percentages). The adjusted coefficient of multiple determination (adjusted R-squared) was used to measure the predictive power of the MLR model with a penalty for model complexity [22,51,52]. The standard partial regression coefficients (B) have been introduced to demonstrate the relative significance of the various response predictors, which are independent of the data units [22,53,54]. A predictor with a bigger B has a better correlation with the response. Separate stepwise MLR tests for each reach distance scale are determined for a given water quality parameter. The final MLR model with stepwise variable selection was selected based on specific criteria: (1) the model had the highest *adjusted* R^2 among all the reaches, buffers, and catchments; (2) the significance of the coefficients of the model and predictors are both at the 95% confidence level ($p \le 0.05$). Subsequently, the goodness-of-fit of the final MLR models was evaluated by scattering and simple linear regression analysis of observed versus predicted values [14].

A set of LMM was used to establish the relationship between land use proportions and water quality measures. The Linear Mixed Model is beneficial in spatial sciences as it accounts for the spatial autocorrelation between samples [41]. The spatial autocorrelation was used since there are multiples catchments, and there is a temporal autocorrelation because the data are between 2001 and 2016. To quantify the spatial variation in the relationships between individual land uses and water quality parameters, the linear mixed model was used. All analyses were conducted in R. For each water quality parameter, 192 univariate models were computed, with each water quality parameter as the response variable, and each of the four land uses types, at each of the eight spatial extents, as the predictor variable. Separate models were developed for each water quality parameter, where 4 land uses × 8 spatial extents = 32 models times 6 water quality parameters = 192 models. Because each of the rivers within regions could vary in mean water quality parameter concentrations (i.e., a random intercept) and/or the rivers within regions could vary in the response of river water quality parameters to differences in land uses (i.e., a random slope).

The analysis was conducted in two steps. Firstly, the amount of water quality parameter variation occurring among regions was quantified. If there was a significant (p < 0.05) amount of among scale variation, each area varies in its mean river water quality parameter concentration (i.e., a random intercept). Secondly, the potential of the effects of the land uses on river water quality parameters vary regionally and yearly. By creating and comparing univariate models, we avoided any problems with multicollinearity in the data (which are likely across spatial extents and land uses).

A random term accounts for the inherent spatial variation of observations belonging to the same location "Site," as well as a temporal-dependent random effect that is meant to account for extraordinary weather events for a given year "Year" and an additional random term that is the combination of year and site effect. Below, the parameter estimates for the fixed effects are presented as land use percentages, such as % Developed, % Forest, % Agriculture, and % Pasture. Formulating the predictive model for water quality parameters, the linear mixed-effects model in Equation (1) is.

$$y = X\beta + Zu + \varepsilon \tag{1}$$

where *y* is the response vector (water quality parameter), β represents all fixed effects; *u* the random effects, *X* design matrix for the fixed effects "land uses," *Z* design matrix for the random effects "site and year"; *u* and ε are independent. Random effects encompass variation among individuals [55]. The nlme package in R was used for the above statistical analysis.

The adjusted R squared, Equation (2), was used to adjust the R-value to the number of explanatory terms in the model relative to the number of data points where p is the total number of explanatory variables in the model (not including the constant term) and n is the size of the sample. The adjusted R-squared values were used to compare the explanatory power of models that contain different numbers of data. The adjusted R-squared value can be negative, and it is always lower than the R-squared value.

$$\acute{R^2} = 1 - \left(1 - R^2\right) \frac{n-1}{n-p-1}$$
⁽²⁾

3. Results and Discussion

Developed and pasture land show that there is not much difference in land proportion in reach distance greater than 25 km (Figure 4a,c). In contrast, for forest, there is a difference among reach distances for all reach distances (Figure 4b), and agriculture land has the most significant differences among reach distances (Figure 4d), varying from 0.1 to 89%. Agricultural land presents the smallest values at 2 and 10 km in some sites, where developed land or forest have a high percentage, indicating that the urban areas are closer to watershed outlets and near rivers. Developed land shows outliers at the 2 km reach distance; it shows that some of the sites are mainly developed land use at a 2 km reach distance. Forest land use shows a decreasing percentage from closer reach distances to further reach distances (Figure 4b). In contrast, agricultural land use shows an increasing proportion of land use from closer reach distances to further reach distances (Figure 4d).

Figure 4. Boxplots of the percentages of the four land uses at different reach distances for study watersheds: (**a**) developed, (**b**) forest, (**c**) pasture, and (**d**) agriculture.

The concentrations of nutrients in areas of mixed land use were found lower than in agricultural or urban areas but higher than those in undeveloped areas [56]. The land use spatial distribution varies, influencing water quality. In order to determine at what distance the heterogeneity of the landscape best explains water quality parameters, the fixed effects (developed, forest, pasture, and agriculture land use) were calculated for reach distances varying from 2, 10, 25, 35, 40, 45, and 50-km; riparian buffers of 200, 500, and 1000-m and basin, and an LMM with site, year, and site and year as random effects were fitted to each of them.

The adjusted R-squared (R^2) shows that the explanatory power of heterogeneity of landscape has the best model of NNN at a reach distance of 45 km (Figure 5a). In contrast, the results across riparian buffers did not show a relationship between NNN and land use with $R^2 < 0.03$ (Figure 6). The NNN was best explained by land use quantified within the entire watershed [57]. Comparing results for the reach distance of 45 km with those for the entire watershed (Figure 5a), the best model for the entire watershed has an R^2 of 0.41 for LMM with year as random effects, while for 45 km, it has a R^2 of 0.64, showing

50km

50km

50k n

M. share

W.shed

W.shed

that the total NNN variance is best explained by the land uses at 45 km with year and site as random effects (Figure 5a). Figure 7 shows the parameter estimates of the best fitting model where fixed effects were calculated at 45 km and the corresponding p-values.

Figure 5. The adjusted R² of multiple regression models describing the strong associations between land use and water quality across spatial scales and different models (**a**) NNN, (**b**) TP, (**c**) TSS, (**d**) COD, (**e**) DO, and (**f**) TKN.

Figure 6. The adjusted R² of multiple regression models describing the strong associations between land use and water quality across riparian scales (200 m, 500 m, and 100 m) and different models, with the water quality parameters NNN, TP, TSS, COD, DO, and TKN.

Figure 7. Significant coefficients (significance level = 0.1) in the mixed effect models describing the associations between land use and water quality for Total Nitrogen.

Figure 4 shows the boxplot of the percentages of the land uses at different reach distances. Based on Figure 4, the model estimates at 45 km are valid for developed, forest, pasture and agriculture percentages varying from 4.4 to 17%, 1 to 28%, 1 to 9% and 57 to 88%, respectively. Figure 8a visualizes the changes in the mean NNN based on the model estimates, for the plausible percentages of forest, developed, and agriculture of the best model (Figure 5a). More specifically, we looked at NNN changes when developed increases and agriculture decreases within one level with forest and pasture constant: forest high area of 27.9% (blue line in Figure 8a) and forest low area of 1.8% (orange line in Figure 8a). Figure 8a shows that as developed increases and agriculture decreases, the mean NNN decreases, and NNN has higher values in low forest conditions (orange line) compared to the high forest (blue line). Nitrogen losses from Indiana watersheds typically range from 4

to 20 Kg N/ha/yr, depending on precipitation; in the streams, inorganic N loads are 90% or more NO₃-N [58]. In intensively farmed areas, fertilizer use and drainage modifications appear to interact to produce large NO₃⁻ N losses through tiles in the cornbelt region [59]. In channelized headwater streams, precipitation and solutes are drained rapidly, bypassing riparian buffer areas with Indiana's vast network of subsurface tile drains [60]. Nitrogen in the form of NO₃⁻ is extremely mobile in soil where nitrate degradation by leaching is a physical process in which soluble NO₃⁻ moves below the root zone with soil water. The best model of NNN at a reach distance of 45 km is consistent with the high mobility of NO₃⁻.

Figure 8. Changes in the mean water quality parameters at different reach distances. (**a**) NNN at 45 km, (**b**) TP at 10 km, (**c**) TSS at 10 km, (**d**) COD at 50 km, (**e**) DO at 45 km, and TKN at 50 km.

The R² indicates that the explanatory strength of landscape heterogeneity shows the best model of TP at a reach distance of 10 km (Figure 5b). This result is similar for riparian buffers with an R² of 0.79 for LMM with year and site as random effects for 200, 500 and 1000 m (Figure 6). The results are similar to those of Liu et al. (2017), who found that the TP concentration was most correlated to land use in closer reach distances. Phosphorus reaches water in agricultural settings. Phosphorus is not highly soluble, and it moves slowly. It binds strongly to fine soil particles such as clay-sized material (<2 microns), and they are transported by flow through artificial tile drainage. Midwest agricultural land has tile drainage, where phosphorus moves in the form of ions of phosphate PO₄⁻³ (Royer et al., 2006), which is consistent with the result that the best TP model is at a reach distance of 10 km.

The parameter estimates (Figure S1b in Supplemental Materials) show the best fitting model where fixed effects were calculated at 10 km and the corresponding *p*-values. Based on Figure 4, the model estimates are valid for developed, forest, pasture and agriculture percentages varying from 3 to 50%, 4 to 44%, 0.1 to 18% and 14 to 87%, respectively. Figure 8b shows changes in the mean TP based on the model estimates, for the plausible percentages of forest, developed, and agriculture in the best model (Figure 5b). More specifically, we looked at TP changes when developed increases and agriculture decreases within two levels of forest and pasture: pasture high 0.18 (orange line in Figure 8b) and pasture low 0.003 (blue line in Figure 8b). Figure 8b shows that as developed increases and agriculture decreases, mean TP decreases, and TP has lower values in high pasture conditions (orange line) compared to low pasture (blue line). According to Figure 8b, as forest increase from 0.10 to 0.50 (blue line), TP decreases from 0.76 to 0.14 mg L⁻¹.

The adjusted R-squared for TSS shows that the explanatory strength of landscape heterogeneity has the best model at a 10 km reach distance (Figure 5c). Riparian scales show lower adjusted R² values (Figure 6). The parameter estimates show the best fitting model with year-site as a random effect where fixed effects were calculated at 10 km and the corresponding *p*-values (Figure S1c in Supplemental Materials). Based on Figure 4, the model estimates are valid for developed, forest, pasture and agriculture percentages varying from 3 to 50%, 4 to 45%, 1 to 18% and 14 to 87%, respectively. Figure S1c in Supplemental Materials shows the changes in the mean TSS based on the model estimates, for the plausible percentages of forest, developed, and agriculture in the best model. In particular, we looked at TSS changes when developed, and forest increases and agriculture decreases within two levels of pasture and forest: pasture 0.01 (orange line in Figure 8c), mean TSS decreases from 90.82 to 33.97 mg L⁻¹; and pasture 0.18 (blue line in Figure 8c), shows a decreasing mean TSS from 44.44 to 20.57 mg L⁻¹, additionally when forest decreases, mean TSS decreases too.

The R² suggests that the explanatory power of heterogeneity of the landscape has the best model for COD at a reach distance of 50 km (Figure 5d). COD for riparian buffers shows similar results with the highest R^2 of 0.62 (Figure 6). Figure S1d in Supplemental Materials shows the parameter estimates of the best fitting model where fixed effects were calculated at a 50 km reach distance and the corresponding *p*-values. Based on Figure 4, the model estimates are valid for developed, forest, pasture and agriculture percentages varying from 4 to 16%, 2 to 29%, 1 to 8% and 55 to 89%, respectively. Figure S1d in Supplemental Materials shows the changes in the mean COD based on the model estimates, for the plausible percentages of forest, developed, and agriculture in the best model (Figure 5d). Specifically, we analyzed shifts in COD when developed, and agriculture increases within two levels of forest and pasture: pasture high 0.08 (orange line in Figure 8d) and pasture low 0.01 (blue line in Figure 8d). The COD variation was better explained by closer reach distances with site as a random effect which is likely due to the proximity to urban areas of water quality stations, where daily inputs, such as urban stormwater runoff, including plant debris, sewage, waste, gasoline and motor oil, heavy metals, fertilizer, and pesticides are common reasons for higher COD. Figure 8d shows that as developed and agriculture increases and forest decreases, mean COD decreases, and COD has lower values in high pasture proportion condition (orange line) compared to low pasture (blue line). According to Figure 8d, the highest COD of 24.76 mg L^{-1} is when developed and agriculture decreases, and forest increases in the blue line; in contrast, the lowest COD 7.70 mg L^{-1} is when developed and agriculture increases, and forest decreases in the orange line.

The R^2 shows that the explanatory strength of landscape heterogeneity has the best model for DO at a reach distance of 45 km (Figure 5e); however; riparian has similar R² with values of 0.62 with year and site as random effects for 200, 500 and 1000 m (Figure 6). The result is in contrast with the findings of [16], who found that land uses for closer reach distance better describe the variation in water quality due to its higher heterogeneity. DO values are affected in streams by temperature, organic discharges, and fertilizer runoff from farms and lawns that changes with hydrology fluctuations, which are explained by the random model with year and site as random effects. Figure S1e in Supplemental Materials shows the parameter estimates of the best fitting model where fixed effects were calculated at 45 km and the corresponding *p*-values. Based on Figure 4, the model estimates are valid for developed, forest, pasture and agriculture percentages varying from 4 to 17%, 2 to 28%, 1 to 9% and 57 to 88%, respectively. The changes in the mean DO base shows on the model estimates (Figure S1e in Supplemental Materials), for the plausible percentages of forest, developed, pasture, and agriculture in the best model (Figure 5e). In particular, we looked at DO changes when developed and agriculture increase and forest decreases within two levels of forest and pasture: pasture high 0.09 (orange line in Figure 8e) and pasture low 0.01 (blue line in Figure 8e). Figure 8e shows that as developed and agriculture increases and forest decreases, the mean DO decrease having the lowest DO of 8.02 mg L^{-1} in high pasture condition (orange line). According to Figure 8e, the highest DO of 10.30 mg L^{-1} is when developed and agriculture decrease, and forest increases in the blue line.

The adjusted R-squared suggests that the explanatory power of heterogeneity of the landscape has the best model for TKN at a distance of 50 km (Figure 5f); similarly, the highest R^2 of 0.68 (Figure 6) was found for riparian buffers for TKN at 1000 m. The parameter estimates of the best fitting model where fixed effects were calculated at 50 km and the corresponding *p*-values (Figure S1f in Supplemental Materials). Based on Figure 4, the model estimates are valid for developed, forest, pasture, and agriculture percentages varying from 4 to 16%, 2 to 29%, 1 to 8%, and 55 to 89%, respectively. Figure S1f in Supplemental Materials shows the changes in the mean TKN based on the model estimates, for the plausible percentages of forest, developed, and agriculture in the best model (Figure 5f). Specifically, we analyzed shifts in TKN when developed, and agriculture increases within two levels of forest and pasture: pasture high 0.08 (orange line in Figure 8f) and pasture low 0.01 (blue line in Figure 8f). Figure 8f shows that as developed and agriculture increase and forest decreases, mean TKN increases, and TKN has higher value in low pasture condition (blue line) compared to low pasture (orange line). According to Figure 8f, the lowest TKN 0.43 mg L^{-1} is when developed and agriculture decrease and forest increases in the orange line; in contrast, the highest TKN 0.79 mg L^{-1} is when developed and agriculture increase and forest decreases in the blue line.

Relationship between Land Use and Water Quality

The study was designed to understand the influence of land use on water quality throughout different reach distances and riparian buffers over a long period (2001–2016). Because land uses are not randomly distributed throughout the basins, changing reach distances and riparian buffer widths of the contributing area alters the relative proportion of different land uses, the perceived dominant land use, and the observed rate of land-use change with downstream distance. Comparing the whole catchment with reach distances and riparian buffers, the whole catchment land use distribution appears more similar to each other, than they do to smaller reach distances or riparian buffer land use distributions. The study was not intended to predict particular water quality values at individual points over the years, as the design approach was tailored to spatial relationships. Therefore, this

analysis is useful in determining what reach distance and riparian buffer, the heterogeneity of the landscape, best explains water quality parameters. This study found that the LMM best explained the variance of mean NNN and DO at a reach distance of 45 km; for mean TP and TSS, the best reach distance is 10 km; and for mean COD and TKN, the best reach distance is 50 km.

Changing the reach distances or riparian buffer widths would also affect the relationship between land use and water quality. For example, Figure 7 shows that NNN has a high R^2 at 40 km and 45 km reach distances, indicating that at those reach distances, the influence of landscape configuration is a better predictor of water quality than other reach distances.

As the developed proportion increases, the agricultural proportion decreases, and the NNN, TP, and TSS means reduce as shown in Figure 8a–c, respectively. This result is beneficial because it shows that TP has a strong relationship with agricultural and developed land use [2]. As developed and agricultural proportion increases, the COD and DO mean decrease as shown in Figure 8d,e, respectively, and the TKN mean increases as shown in Figure 8f. The amount of nitrogen generated from agricultural lands was found about seven times higher than from impermeable urban areas and nine times higher than produced in pervious urban areas [2]. In a study with land use change, researchers found that when developed area increases, COD concentration decreases [61].

Forested land has been widely identified as having an essential role in improving the quality of surface water [62,63]. Similar findings were found in this study, where increasing forest land proportion decreases mean NNN, TP, TSS, and TKN, which may occur due to its role in the conservation of water, prevention of soil erosion and the filtration of overland pollutants [1].

Pasture land played a dominant role in influencing TN and TP in the Jinsha River [18]. This research had similar findings, where increasing pasture proportion increases the mean TP, TSS, and TKN, as shown in Figure 8b,c,f, respectively. Pasture land has a better relationship with water quality for further reach distances, likely due to the averaging out of the effects of local heterogeneity, that is a widely observed phenomenon in ecological trends [42,64]. As pasture proportion decreases, the mean COD and DO increase, as shown in Figure 8d,e, respectively. Pasture land was found inversely proportional to COD [61]. Pasture can benefit water quality by reducing the influx of nutrients entering into streams.

Increasing agricultural land increases the TP mean, which agrees with previous findings [2,65,66], suggesting that surface runoff and cropland degradation is the principal cause of phosphorus in streams [57]. Nitrogen and phosphorus losses in surface runoff are also more significant in agricultural areas than land with natural vegetation, where phosphorus losses are transported by sediment eroded from cropland [67]. This study showed (Figure 8a) that NNN increases when agricultural land proportion increases. The result is consistent with nitrogen fertilizer use in the corn belt area, which increases nitrogen inputs to watersheds [68].

4. Conclusions

This study showed that for developed and pasture land use percentages, there is not much difference among study watershed reach distances. The 2-km reach distance has the most outliers in developed and pasture land. In contrast, forest shows differences among reach distances. Agricultural land has the most significant differences among reach distances, varying from 0.1 to 89%. The developed and pasture land uses show that there is likely not a difference among reach distances, while in contrast, forest and agriculture are likely to have differences among reach distances, and the data are more dispersed at 2 km than at longer reach distances. This study revealed that closer reach distances have more dispersed data, due to the area size and heterogeneity, in contrast to farther distances, which experienced less dispersed data and more uniformity in land use.

The study demonstrated the efficacy of the linear mixed model to examine the importance of spatial extent on models with river water quality parameters. The adjusted R^2 improved from 0.36 to the highest adjusted R² of 0.79 for NNN at a reach distance of 45 km. The best model for TP is at a reach distance of 10 km, where the model could explain 82% of the variance, and the highest adjusted R² of 0.26 for TSS is at the same reach distance. The reason that TP and TSS have a higher adjusted R² at the same distance is likely since TP adheres to sediment particles and is likely to be absorbed on sediments. The lowest mean TSS of 20.57 mg L⁻¹ was found at lower agricultural land use percentages; in contrast, the highest mean TSS of 90.82 mg L⁻¹ was found at high agricultural land use percentage. The results of the study suggest that NNN, DO, COD and TKN are better predicted for land uses at further reach distances, as opposed to TP and TSS that are better predicted for land uses at closer reach distances.

This study found that LMM with year and site as random variables better explained the NNN, TP, TSS, COD, DO, and TKN water quality parameters variance than linear models. Year and site help explain part of the variation among years and sites for the water quality parameter means. For TP, the linear model explained 20% of the variance at a 10-km reach distance, compared to the LMM with year × site that explained 82% of the variance at a 10-km reach distance. Comparing these results with riparian buffer results, the linear model explained 0.05% of the variance at 10 km, compared to the LMM with year × site that explained 80% of the variance. However, riparian buffer plays a significant role in governing water quality parameter transport and retention, while the linear model and LMM explained less of the variance. In summary, this study found that the land use predicted NNN, TP, TSS, COD, DO, and TKN water quality parameters better using reach distances than buffers.

In general, the use of linear mixed model in the monitoring of water quality has outlined many advantages of its use and how the effects of the fitted models can be used to classify catchments and provide information on the factors that can help optimize local land use and manage water quality. Further research will be required to refine and define more spatial scales, and use of high spatial resolution digital maps to aid improve modeling and temporal sampling will likely better explain the dynamic existence of the relationship between land use and water quality.

Supplementary Materials: The following are available online at https://www.mdpi.com/article/10 .3390/w13192618/s1, Figure S1. Significant coefficients (significance level = 0.1) in the mixed effect models describing the associations between land use and water quality. (a) TP, (b) TSS, (c) COD, (d) DO, (e) TKN, Table S1. Random coefficients values for LMM at reach distance of 40 km with Site as Random effects.

Author Contributions: Writing original draft, J.A.V.M.; methodology, J.A.V.M.; formal analysis, J.A.V.M. and B.E.; editing and reviewing, B.E.; conceptualization, J.A.V.M.; and supervision, B.E. Both authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: Not applicable.

Acknowledgments: We are thankful to Farzin Shamloo for statistical guidance provided during this work and also would like to thank to the Colombian government for the Colciencias Scholarship administrated by Minciencias.

Conflicts of Interest: The authors declare no conflict of interest.

References

- Haidary, A.; Amiri, B.J.; Adamowski, J.; Fohrer, N.; Nakane, K. Assessing the impacts of four land use types on the water quality of wetlands in Japan. *Water Resour. Manag.* 2013, 27, 2217–2229. [CrossRef]
- Tong, S.; Chen, W. Modeling the relationship between land use and surface water quality. J. Environ. Manag. 2002, 66, 377–393.
 [CrossRef] [PubMed]
- 3. Aboelnour, M.; Gitau, M.W.; Engel, B.A. Hydrologic response in an urban watershed as affected by climate and land-use change. *Water* **2019**, *11*, 1603. [CrossRef]
- 4. Meynendonckx, J.; Heuvelmans, G.; Muys, B.; Feyen, J. Hydrology and earth system sciences effects of watershed and riparian zone characteristics on nutrient concentrations in the River Scheldt Basin. *PLoS ONE* **2006**, *10*, e53163. [CrossRef]

- 5. Tromboni, F.; Dilts, T.E.; Null, S.E.; Lohani, S.; Ngor, P.B.; Soum, S.; Hogan, Z.; Chandra, S. Changing land use and population density are degrading water quality in the lower mekong basin. *Water* **2021**, *13*, 1948. [CrossRef]
- 6. Wilson, C.; Weng, Q. Assessing surface water quality and its relation with urban land cover changes in the Lake Calumet Area, Greater Chicago. *Environ. Manag.* 2010, 45, 1096–1111. [CrossRef] [PubMed]
- 7. Mander, Ü.; Kull, A.; Tamm, V.; Kuusemets, V.; Karjus, R. Impact of climatic fluctuations and land use change on runoff and nutrient losses in rural landscapes. *Landsc. Urban Plan.* **1998**, *41*, 229–238. [CrossRef]
- Aboelnour, M.; Engel, B.A. Application of remote sensing techniques and geographic information systems to analyze land surface temperature in response to land use/land cover change in greater Cairo Region, Egypt. J. Geogr. Inf. Syst. 2018, 10, 57–88.
 [CrossRef]
- 9. Baker, T.J.; Miller, S.N. Using the soil and water assessment tool (SWAT) to assess land use impact on water resources in an East African watershed. *J. Hydrol.* **2013**, *486*, 100–111. [CrossRef]
- 10. Alberti, M.; Booth, D.; Hill, K.; Coburn, B.; Avolio, C.; Coe, S.; Spirandelli, D. The impact of urban patterns on aquatic ecosystems: An empirical analysis in Puget lowland sub-basins. *Landsc. Urban Plan.* **2007**, *80*, 345–361. [CrossRef]
- 11. King, R.S.; Baker, M.E.; Whigham, D.F.; Weller, D.E.; Jordan, T.E.; Kazyak, P.F.; Hurd, M.K. Spatial considerations for linking watershed land cover to ecological indicators in streams. *Ecol. Appl.* **2005**, *15*, 137–153. [CrossRef]
- 12. Shi, P.; Zhang, Y.; Li, Z.; Li, P.; Xu, G. Influence of land use and land cover patterns on seasonal water quality at multi-spatial scales. *Catena* **2017**, *151*, 182–190. [CrossRef]
- 13. Woli, K.P.; Nagumo, T.; Kuramochi, K.; Hatano, R. Evaluating river water quality through land use analysis and N budget approaches in livestock farming areas. *Sci. Total Environ.* **2004**, *329*, 61–74. [CrossRef]
- 14. Amiri, B.; Nakane, K. Entire catchment and buffer zone approaches to modeling linkage between river water quality and land cover-a case study of Yamaguchi Prefecture, Japan. *Chin. Geogr. Sci.* 2008, *18*, 85–92. [CrossRef]
- 15. Shen, Z.; Hou, X.; Li, W.; Aini, G.; Chen, L.; Gong, Y. Impact of landscape pattern at multiple spatial scales on water quality: A case study in a typical urbanised watershed in China. *Ecol. Indic.* **2015**, *48*, 417–427. [CrossRef]
- 16. Uriarte, M.; Yackulic, C.B.; Lim, Y.; Arce-Nazario, J.A. Influence of land use on water quality in a tropical landscape: A multi-scale analysis. *Landsc. Ecol.* 2011, 26, 1151–1164. [CrossRef]
- 17. Gamvroula, D.; Alexakis, D.; Stamatis, G. Diagnosis of groundwater quality and assessment of contamination sources in the Megara basin (Attica, Greece). *Arab. J. Geosci.* 2013, *6*, 2367–2381. [CrossRef]
- 18. Huang, J.; Zhan, J.; Yan, H.; Wu, F.; Deng, X. Evaluation of the impacts of land use on water quality: A case study in the Chaohu lake basin. *Sci. World J.* 2013, 2013, 329187. [CrossRef]
- 19. Zhang, S.; Fan, W.; Li, Y.; Yi, Y. The influence of changes in land use and landscape patterns on soil erosion in a watershed. *Sci. Total Environ.* **2017**, *574*, 34–45. [CrossRef] [PubMed]
- 20. Alexakis, D.E.; Bathrellos, G.D.; Skilodimou, H.D.; Gamvroula, D.E. Land suitability mapping using geochemical and spatial analysis methods. *Appl. Sci.* 2021, *11*, 5404. [CrossRef]
- Ma, X.; Xu, J.; Luo, Y.; Prasad Aggarwal, S.; Li, J. Response of hydrological processes to land-cover and climate changes in Kejie watershed, south-west China. *Hydrol. Process.* 2009, 23, 1179–1191. [CrossRef]
- 22. Ding, J.; Jiang, Y.; Liu, Q.; Hou, Z.; Liao, J.; Fu, L.; Peng, Q. Influences of the land use pattern on water quality in low-order streams of the Dongjiang River basin, China: A multi-scale analysis. *Sci. Total Environ.* **2016**, *551–552*, 205–216. [CrossRef]
- Sun, Y.; Guo, Q.; Liu, J.; Wang, R. Scale effects on spatially varying relationships between urban landscape patterns and water quality. *Environ. Manag.* 2014, 54, 272–287. [CrossRef] [PubMed]
- 24. Allan, D. Landscapes and riverscapes: The influence of land use on stream ecosystems. *Annu. Rev. Ecol. Evol. Syst.* 2004, 35, 257–284. [CrossRef]
- Pratt, B.; Chang, H. Effects of land cover, topography, and built structure on seasonal water quality at multiple spatial scales. J. Hazard. Mater. 2012, 209–210, 48–58. [CrossRef]
- 26. Schilling, K.E.; Jha, M.K.; Zhang, Y.K.; Gassman, P.W.; Wolter, C.F. Impact of land use and land cover change on the water balance of a large agricultural watershed: Historical effects and future directions. *Water Resour. Res.* **2008**, 45. [CrossRef]
- Williams, M.R.; King, K.W.; Macrae, M.L.; Ford, W.; Van Esbroeck, C.; Brunke, R.I.; English, M.C.; Schiff, S.L. Uncertainty in nutrient loads from tile-drained landscapes: Effect of sampling frequency, calculation algorithm, and compositing strategy. *J. Hydrol.* 2015, 530, 306–316. [CrossRef]
- 28. Collins, K.E.; Doscher, C.; Rennie, H.G.; Ross, J.G. The effectiveness of Riparian 'Restoration' on water quality-a case study of Lowland Streams in Canterbury, New Zealand. *Restor. Ecol.* **2013**, *21*, 40–48. [CrossRef]
- 29. McMillan, S.K.; Tuttle, A.K.; Jennings, G.D.; Gardner, A. Influence of restoration age and Riparian vegetation on reach-scale nutrient retention in restored urban streams. *JAWRA J. Am. Water Resour. Assoc.* **2014**, *50*, 626–638. [CrossRef]
- 30. Sahu, M.; Gu, R.R. Modeling the effects of riparian buffer zone and contour strips on stream water quality. *Ecol. Eng.* **2009**, *35*, 1167–1177. [CrossRef]
- Peterson, F.; Risberg, J. Nutrients: Phosphorus, Nitrogen Sources, Impact on Water Quality-A General Overview; Minnesota Pollution Control Agency: Saint Paul, MN, USA, 2008.
- Chappell, N.A.; Jones, T.D.; Tych, W. Sampling frequency for water quality variables in streams: Systems analysis to quantify minimum monitoring rates. *Water Res.* 2017, 123, 49–57. [CrossRef] [PubMed]

- Kumar, M.; Puri, A. A review of permissible limits of drinking water. *Indian J. Occup. Environ. Med.* 2012, 16, 40–44. [CrossRef] [PubMed]
- 34. Teixeira, Z.; Marques, J.C. Relating landscape to stream nitrate-N levels in a coastal eastern-Atlantic watershed (Portugal). *Ecol. Indic.* **2015**, *61*, 149–1056. [CrossRef]
- 35. Pyron, M.; Neumann, K. Hydrologic alterations in the Wabash River watershed, USA. *River Res. Appl.* **2008**, *24*, 1175–1184. [CrossRef]
- 36. Gammon, J. The Wabash River Ecosystem; Cinergy Corporation: Cincinnati, OH, USA, 1998; ISBN 0253212723.
- 37. White, D.; Johnston, K.; Miller, M. Ohio River Basin; Academic Press: Cambridge, MA, USA, 2005.
- 38. Gammon, J. The status of riparian wetlands in west-central Indiana streams. Proc. Indiana Acad. Sci. 1994, 103, 195–214.
- Bonansea, M.; Rodriguez, M.C.; Pinotti, L.; Ferrero, S. Using multi-temporal Landsat imagery and linear mixed models for assessing water quality parameters in Río Tercero reservoir (Argentina). *Remote Sens. Environ.* 2015, 158, 28–41. [CrossRef]
- 40. Lark, R.M.; Cullis, B.R. Model-based analysis using REML for inference from systematically sampled data on soil. *Eur. J. Soil Sci.* **2004**, *55*, 799–813. [CrossRef]
- 41. Lessels, J.S.; Bishop, T.F.A. Estimating water quality using linear mixed models with stream discharge and turbidity. *J. Hydrol.* **2013**, *498*, 13–22. [CrossRef]
- 42. Buck, O.; Niyogi, D.K.; Townsend, C.R. Scale-dependence of land use effects on water quality of streams in agricultural catchments. *Environ. Pollut.* **2004**, 130, 287–299. [CrossRef]
- 43. Kumar, S.; Merwade, V.; Kam, J.; Thurner, K. Streamflow trends in Indiana: Effects of long term persistence, precipitation and subsurface drains. *J. Hydrol.* 2009, 374, 171–183. [CrossRef]
- 44. Karns, D.R.; Pyron, M.; Simon, T.P. *Proceedings of the Indiana Academy of Science*; St. Mary-of-the-Woods College: Arlington, IN, USA, 2006; Volume 1.
- Multi-Resolution Land Characteristics (MRLC) Consortium | Multi-Resolution Land Characteristics (MRLC) Consortium. Available online: https://www.mrlc.gov/ (accessed on 11 May 2020).
- 46. Schmadel, N.M.; Harvey, J.W. NHD Plus; USGS: Reston, VA, USA, 2020.
- 47. Nejadhashemi, A.P.; Shen, C.; Associate, R.; Wardynski, B.J.; Assistant, R.; Mantha, P.S.; Lawrence, D.L. Evaluating the Impacts of Land Use Changes on Hydrologic Responses in the Agricultural Regions of Michigan and Wisconsin. In Proceedings of the 2010 ASABE Annual International Meeting Sponsored by ASABE, Pittsburgh, PA, USA, 20–23 January 2010.
- 48. US EPA. Water Quality Data (WQX); U.S. Environmental Protection Agency: Virginia, DC, USA, 2017.
- 49. O'Callaghan, J.F.; Mark, D.M. The extraction of drainage networks from digital elevation data. *Comput. Vision Graph. Image Process.* **1984**, *28*, 323–344. [CrossRef]
- 50. Jenson, S.K.; Domingue, J.O. Extracting topographic structure from digital elevation data for geographic information system analysis. *Photogramm. Eng. Remote Sens.* **1988**, *54*, 1593–1600.
- Basu, N.B.; Destouni, G.; Jawitz, J.W.; Thompson, S.E.; Loukinova, N.V.; Darracq, A.; Zanardo, S.; Yaeger, M.; Sivapalan, M.; Rinaldo, A.; et al. Nutrient loads exported from managed catchments reveal emergent biogeochemical stationarity. *Geophys. Res. Lett.* 2010, 37, L23404. [CrossRef]
- 52. Legendre, P.; Legendre, L. Numerical Ecology Ch 6-Multidimensional qualitative data. *Dev. Environ. Model.* **2012**, *24*, 337–424. [CrossRef]
- 53. Bring, J. How to standardize regression coefficients. Am. Stat. 1994, 48, 209–213. [CrossRef]
- 54. McDonald, J.H. Handbook of Biological Statistics: Introduction; Sparky House Publishing: Baltimore, MD, USA, 2009.
- 55. Bolker, B.M. Ecological Models and Data in R.; Princeton University Press: Princeton, NJ, USA, 2008; ISBN 9780691125220.
- 56. Martin, J.D.; Wentz, D.A.; Richards, K.D.; Waite, I.R. Nutrients from Atmospheric and Urban Sources, Fertilization, and Livestock Wastes Can Contribute to Excessive Algal Growth in Streams; USGS: Reston, VA, USA, 2017.
- 57. Liu, J.; Zhang, X.; Wu, B.; Pan, G.; Xu, J.; Wu, S. Spatial scale and seasonal dependence of land use impacts on riverine water quality in the Huai River basin, China. *Environ. Sci. Pollut. Res.* **2017**, *24*, 20995–21010. [CrossRef]
- David, M.B.; Gentry, L.E.; Kovacic, D.A.; Smith, K.M. Nitrogen balance in and export from an agricultural watershed. *Environ. Qual.* 1997, 1038–1048. [CrossRef]
- Royer, T.V.; David, M.B.; Gentry, L.E. Timing of riverine export of nitrate and phosphorus from agricultural watersheds in Illinois: Implications for reducing nutrient loading to the Mississippi River. *Environ. Sci. Technol.* 2006, 40, 4126–4131. [CrossRef] [PubMed]
- 60. Hill, A.R. The environmental impacts of agricultural land drainage. J. Environ. Manag. 1976, 4, 251–274.
- 61. Effendi, H.; Sabila, M.F.; Setiawan, Y. Correlation between water quality and land use change in ciliwung watershed. *Nat. Environ. Pollut. Technol.* **2018**, *17*, 139–144.
- 62. Dosskey, M.G.; Vidon, P.; Gurwick, N.P.; Allan, C.J.; Duval, T.P.; Lowrance, R. The role of Riparian vegetation in protecting and improving chemical water quality in streams. *J. Am. Water Resour. Assoc.* **2010**, *18*, 46, 261–277. [CrossRef]
- 63. Ding, J.; Jiang, Y.; Fu, L.; Liu, Q.; Peng, Q.; Kang, M. Impacts of land use on surface water quality in a subtropical river basin: A case study of the dongjiang river basin, Southeastern China. *Water* **2015**, *7*, 4427–4445. [CrossRef]
- 64. Wiens, J.A. Spatial scaling in ecology. Funct. Ecol. 1989, 3, 385. [CrossRef]
- 65. Fuchs, J.W.; Fox, G.A.; Storm, D.E.; Penn, C.J.; Brown, G.O. Subsurface transport of phosphorus in riparian floodplains: Influence of preferential flow paths. J. Environ. Qual. 2009, 38, 473–484. [CrossRef] [PubMed]

- 66. Hoorman, J.; Hone, T.; Sudman, T.; Dirksen, T.; Iles, J.; Islam, K.R. Agricultural impacts on lake and stream water quality in Grand Lake St. Marys, Western Ohio. *Water Air Soil Pollut.* **2008**, *193*, 309–322. [CrossRef]
- 67. Timmons, D.R.; Holt, R.F. Nutrient losses in surface runoff from a native prairie. J. Environ. Qual. 1977, 6, 369–373. [CrossRef]
- 68. Cao, Y.; Williams, D.D.; Williams, N.E. Data transformation and standardization in the multivariate analysis of river water quality. *Ecol. Appl.* **1999**, *9*, 669–677. [CrossRef]