

Article

Linking Forest Cover to Water Quality: A Multivariate Analysis of Large Monitoring Datasets

Delphine Brogna ^{1,*}, Adrien Michez ², Sander Jacobs ^{3,4}, Marc Dufrêne ⁵, Caroline Vincke ⁶ and Nicolas Dendoncker ¹

¹ University of Namur, Department of Geography, 61 rue de Bruxelles, 5000 Namur, Belgium; nicolas.dendoncker@unamur.be

² University of Liege, Gembloux Agro-Bio Tech, Biosystem Engineering research unit, 2 Passage des Déportés, 5030 Gembloux, Belgium; adrien.michez@ulg.ac.be

³ Research institute of nature and forest INBO, Research group Nature & Society, Kliniekstraat 25, 1070 Brussels, Belgium; sander.jacobs@inbo.be

⁴ Belgian Biodiversity Platform BBPF, Avenue Louise 231 Louizalaan, 1050 Brussels, Belgium

⁵ University of Liège, Gembloux Agro-Bio Tech, Biodiversity and Landscape group, UR TERRA, 2 Passage des Déportés, 5030 Gembloux, Belgium; marc.dufrene@ulg.ac.be

⁶ Université catholique de Louvain, Earth and Life Institute, Environmental Sciences, Croix du sud 2 L7.05.09, 1348 Louvain-la-Neuve, Belgium; caroline.vincke@uclouvain.be

* Correspondence: delphine.brogna@gmail.com

Academic Editor: Jiangyong Hu

Received: 17 December 2016; Accepted: 21 February 2017; Published: 1 March 2017

Abstract: Forested catchments are generally assumed to provide higher quality water. However, this hypothesis must be validated in various contexts as interactions between multiple land use and land cover (LULC) types, ecological variables and water quality variables render this relationship highly complex. This paper applies a straightforward multivariate approach on a typical large monitoring dataset of a highly managed and densely populated area (Wallonia, Belgium; 10-year dataset), quantifying forest cover effects on nine physico-chemical water quality variables. Results show that forest cover explains about one third of the variability of water quality and is positively correlated with higher quality water. When controlling for spatial autocorrelation, forest cover still explains 9% of water quality. Unlike needle-leaved forest cover, broad-leaved forest cover presents an independent effect from ecological variables and explains independently 4.8% of water quality variability while it shares 5.8% with cropland cover. This study demonstrates clear independent effects of forest cover on water quality, and presents a method to tease out independent LULC effects from typical large multivariate monitoring datasets. Further research on explanatory variables, spatial distribution effects and water quality datasets could lead to effective strategies to mitigate pollution and reach legal targets.

Keywords: physico-chemical water quality; land cover; land use; forests; agriculture; Wallonia; Belgium; monitoring network; water framework directive; redundancy analysis

1. Introduction

Water is the most essential component for the life of all beings [1,2]. However, freshwater systems and consequently human well-being are directly threatened by human activities [3–5]. In response to the global degradation of ecosystems and their services, water quality management is at the core of policies such as the European Water Framework Directive (Directive, 2000/60/CE) aiming at maintaining or restoring the chemical, physical and biological integrity of surface and groundwater bodies. Managing water quality is challenging and implies to deal with both point and non-point source

pollutions. As non-point source pollutions result from complex run-off and landscape interactions, they are more complex to identify than confined point source pollutions [6]. Land use and Land cover (LULC) are key landscape elements affecting water quality through their impact on non-point source pollution.

Previous studies attempting to address LULC impact on water quality broadly correlate urban and/or agricultural LULC with poor water quality either at the catchment or riparian scale. These represent water quality through several variables, but nitrate and phosphate, which are at the basis of eutrophication problems, are the most studied. More specifically, Álvarez-Cabria et al. [7] model three water quality variables (temperature, concentrations of nitrates and phosphates) in a watershed located in Spain. Their results show that nitrate and phosphate concentrations were mainly related to agricultural LULC and urban LULC, respectively. Chen et al. [8] show that urban land is the dominant factor influencing nitrogen, phosphorus and chemical oxygen demand in highly urbanized regions of a catchment located in eastern China but that agricultural land has the greatest influence on nitrogen and phosphorus in suburban and rural areas. Yu et al. [9] also show direct and indirect negative impact of urbanization and agricultural activities on water quality in an urban area of China. De Oliveira et al. [10] assess LULC effect on nitrate, total ammonia nitrogen, total phosphorus and dissolved oxygen in a Brazilian watershed. Their results (correlations) point out that urban areas and agriculture/pasture tend to worsen water quality while some models (i.e., nitrate and total phosphorus) were not valid. Hwang et al. [11] show that relationships between urban LU and water quality vary according to the degree of urbanization. These studies are often specific to one or few catchments, treat different water quality variables at different temporal and spatial scales but broadly show negative impact of urban and agricultural LULC on water quality. These results are often presented in opposition to forest cover associated with higher water quality.

Forest is one of the LULC that interacts the most with water, whether in terms of quantity or quality, and consequently has an impact on hydrological services—which can be grouped into water supply and water damage mitigation and viewed in terms of quantity, quality and timing [12–14]. Indeed, forests and forest soils alter, relative to other land uses and soil types, each of the five physical, chemical and biological functions involving the reception, processing and transfer of water [15]. Forest general impact on water quality, relative to other land uses, can be summarized as water with less sediments and water with fewer nutrients [15,16]. Several studies state that forested catchments tend to have more stable water quality conditions [17–19], but we did not come across any study that quantifies this impact on several pollutants simultaneously, on a large scale and over a relatively long period. In addition, one may question the validity of these hypothesized relationships under different latitudes, at different temporal and spatial scales, under various management types, according to different forest types (i.e., needle-leaved vs. broad-leaved forests). In addition, global changes affect water quality and quantity and question the assumption that studies from last century can be used to face future conditions [20]. Finally, Giri and Qiu [21] stress the importance of assessing the relationship between LULC and water quality (see also Chauhan and Verma [22]), pointing out that improving our understanding of these can help managing water quality in unmonitored watersheds but also that this knowledge can provide guidelines to watershed managers and policymakers in the land planning decision process.

In Wallonia (Belgium), few studies attempting to assess the impact of LULC and in particular forest cover on water quality were published. Some specific studies related to the subject exist as the assessment of variability of nitrate removal in riparian zones [23] or the in lab assessment of temperature, throughfall volume and ammonium cation deposition impact on soil solution nitrate concentrations, nitrous oxide emissions and numbers of ammonium oxidisers from a forest stand soil [24]. Regarding nitrates, Sohier and Degré present a hydrological model for evaluating the effectiveness of agricultural policy measures on nitrate concentration in surface and ground waters [25].

Different methods can be used to assess the relationship between water quality and LULC. Giri and Qiu [21] classified these into three categories: monitoring, hydrologic/water quality modeling and statistical modeling. Direct and real-time monitoring in the stream is expensive, time consuming,

and ineffective for larger areas. Hydrological and water quality models used in several studies [26–28] require vast amounts of data, are complex, costly and time consuming to calibrate, and therefore only applied on one single or few catchments. Statistical methods tend to be simpler, easier to apply, and more efficient than physically-based hydrologic/water quality models when observed data are limited in time and when datasets are covering many different catchments [29].

Regarding this context, this study applies multivariate statistical methods to mine a regional monitoring dataset from the highly managed and densely populated Walloon region (Belgium). It provides a quantification of forest cover effect on several physico-chemical variables simultaneously. More specifically, this paper:

- i. Analyzes the link between sub-catchments' LULC and the legal status of in stream water quality;
- ii. Quantitatively assesses the link between forest cover and nine water quality variables, verifying spatio-temporal variability;
- iii. Quantifies the independent effect of forest cover types (i.e., needle-leaved and broadleaved forests) on water quality relatively to effects of other LULC.

This study develops a novel approach, replicable in space and time, based on easily accessible public data (public monitoring network data directly linked to the Water Framework Directive and LULC data), and using powerful but simple statistical methods ran in open source statistical software (R stat, [30]).

2. Materials and Methods

2.1. Study Area

The study area is the southern region of Belgium (Wallonia) covering 16,902 km² (ca. 55% of Belgium's area, see Figure 1). We work on the publically managed river network and in particular on 362 water quality stations monitored for the Walloon Public Service (WPS [31], Figure 1).

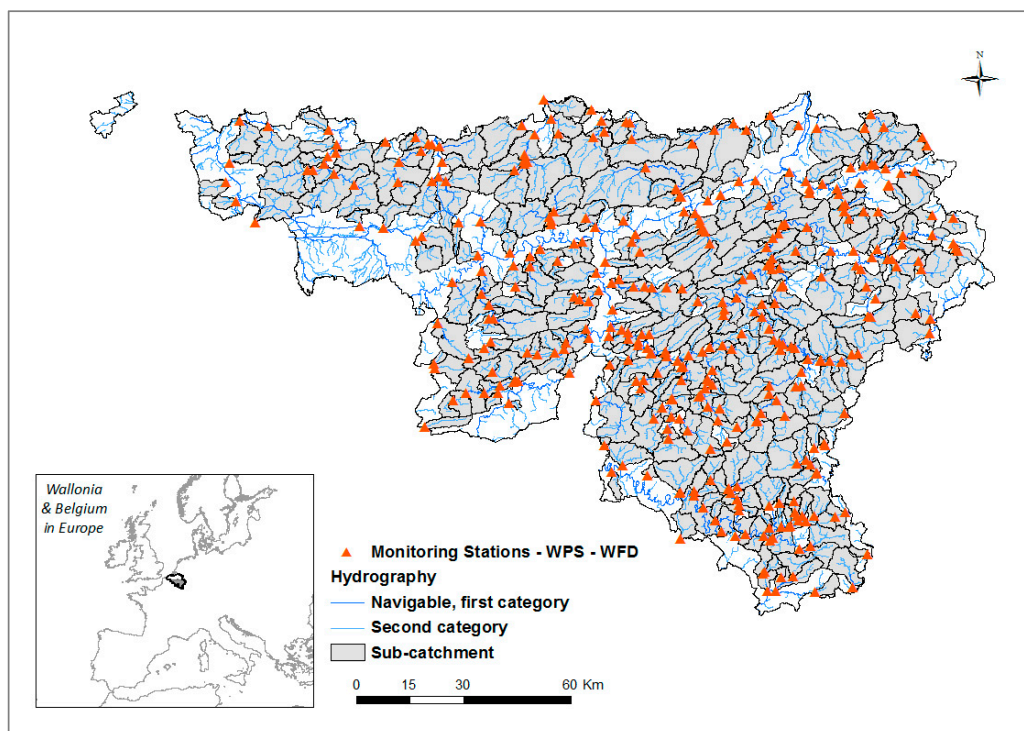


Figure 1. Water quality monitoring stations used in this study and sub-catchments.

Population density in Wallonia is 202 inhabitants per square kilometer and, with hardly less than 1% of the territory benefiting of a natural reserve status, all landscapes are mostly managed or perturbed by human activities. Main LULC are agriculture land (52%) comprising grassland (30%) and cropland (22%); forests (30%) split into needle-leaved (13%), broad-leaved (16%) and mixed forests (1%); and artificial surfaces (10%) (see LULC dataset, Section 2.4 and Figure 2). Cropland cover is mostly located in the North of the region and at a low elevation while forests (especially needle-leaved forests) are located in the South at a higher elevation. In the studied sub-catchments (see delineation method below), the most represented LULC classes are forests, grassland and cropland (see boxplot in Supplementary Materials Figure S1).

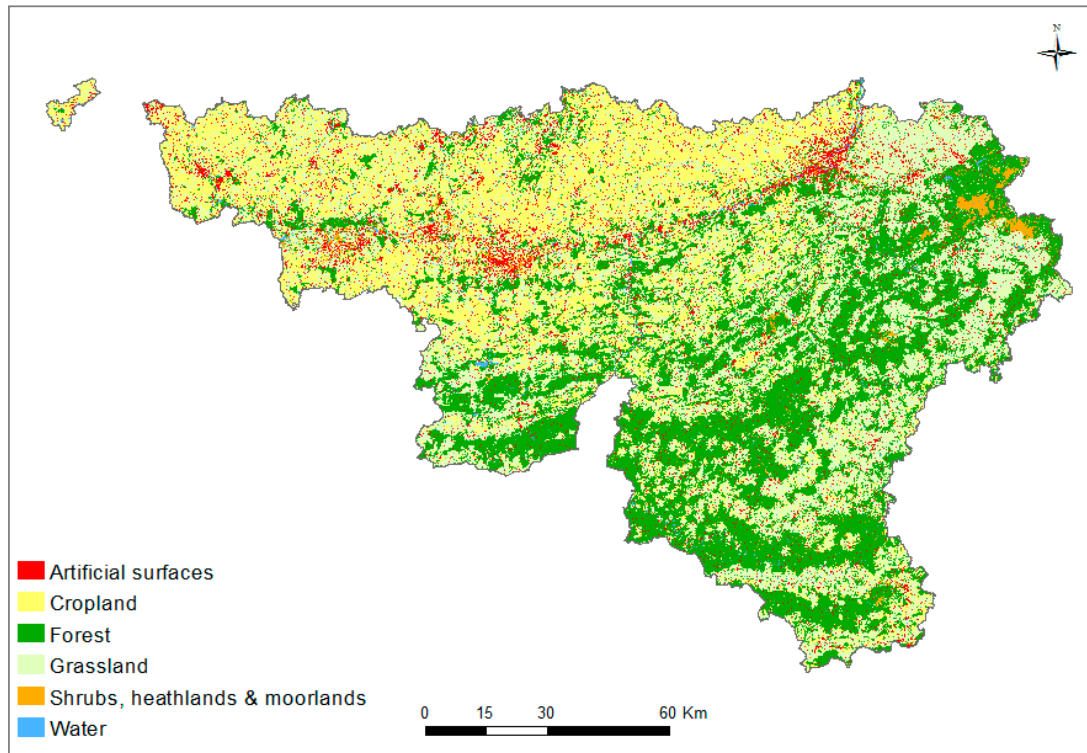


Figure 2. Land use and land cover (LULC) in Wallonia.

Intensive agriculture impacts water quality through the use of mineral fertilizer, in particular nitrogen (N) and phosphorus (P), causing eutrophication and drinking water quality degradation. Even if declining since 1990, inputs of nitrogen and phosphorus were still above the European average in 2001 [32,33]. Nitrogen still exceeded (about double) the European average in 2012 while phosphorus decreased to around half of the European average.

Regarding species composition, broad-leaved forests are largely dominated by oaks (*Quercus robur* and *Q. petraea*) and beech (*Fagus sylvatica*) but other species such as birch (*Betula pendula*), ash (*Fraxinus excelsior*), maple (*Acer pseudoplatanus*), and hornbeam (*Carpinus betulus*) are also present [34]. Needle-leaved forests are very intensively managed with the use of exotic species (mainly spruce (*Picea abies*) but also Douglas fir (*Pseudotsuga menziesii*), larches (*Larix* sp.), and pines (*Pinus sylvestris* and *P. nigra*)), in even-aged stands conducted with systematically clear-cuttings, and high drainage infrastructure on wet soils. In the last century, forests in Wallonia have suffered from large inputs of sulfur and nitrogen through acidic rainfalls. Even if this phenomenon has been declining since 1990, it affected forests during the studied decade [32].

2.2. Workflow

The overall methodology is illustrated in Figure 3 and described below. The first section describes the physico-chemical water quality dataset processing. Secondly, the delineation of sub-catchments and LULC characterization are explained. Finally, the methodology to assess the link between forest cover and water quality is presented. In this last part, we detail the following analyses: the link between forest cover and WFD standards, the quantification of the forest cover effects on water quality and spatio-temporal aspects, and, finally, the partitioning of the LULC effect on water quality between forest types and LULC and their shared effects while controlling for ecological gradient.

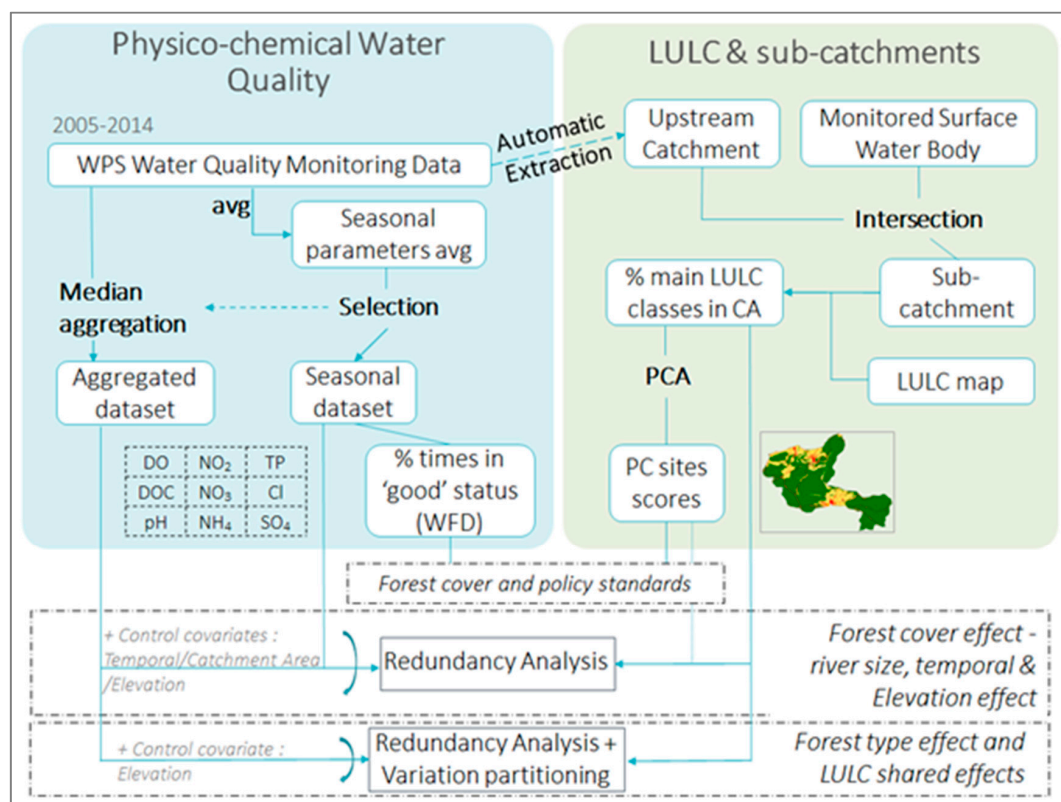


Figure 3. Schematic figure of the methodological approach to link forest cover and water quality. WPS: Walloon Public Service; avg: average; WFD: Water Framework Directive; PCA: Principal Component Analysis.

2.3. Physico-Chemical Water Quality

Physico-chemical water quality was studied through selected variables (Table 1) from the monitoring performed by the WPS. We processed data from 10 hydrological years (October 2004 to September 2014). The hydrological year—i.e., the annual base period—is temporally defined by the period from October to the following September and corresponds to the natural hydrological cycle [35].

We selected 362 stations from the monitoring network dataset of the WPS. The selected stations are characterized by an upstream catchment that can be extracted automatically from a digital elevation model. Consequently, artificial water bodies such as artificial canals hydrologically disconnected or crossing watersheds were excluded. Moreover, we excluded stations whose upstream catchments are partially located outside Wallonia. In order to maximize the number of multiple observations and still fit to seasonal vegetation development, we averaged water quality variables' values by season and by station. Following Brogna et al. [13], seasons were delineated according to vegetation development and rainfall distribution, splitting the year into a “non-vegetation season” (October–March) and a “vegetation season” (April–September). Among the water quality variables monitored by the

administration, variables were excluded: (1) if the Pearson correlation coefficient exceeded 80% to exclude highly redundant variables for multivariate analysis purposes [36]; or (2) if missing data exceeded 5% of the dataset (not to lose too many sampling dates in the multivariate table). Thus, 9 out of 16 variables were analyzed in this study (Table 1).

Table 1. Physico-chemical variables investigated.

Quality "Group"	Variable		Unit
Oxygen balance	Dissolved Oxygen	DO	(mgO ₂ /L)
	Dissolved Organic Carbon	DOC	(mgC/L)
Phosphorus	Total phosphorus	TP	(mgP/L)
Nitrogenous material	Ammonium	NH ₄	(mgN/L)
	Nitrites	NO ₂	
	Nitrates	NO ₃	
Acidification	pH	pH	
Mineralization	Chloride	Cl	(mg/L)
	Sulfate	SO ₄	(mg/L)

The seasonal dataset consists of a 9 variables \times 3793 observations (related to 362 stations) table. We excluded values that exceeded the 99th quantile as they are outliers representing incorrect values. We applied Log- or square-based transformations when improving the normality of variables' distribution was needed. We computed the percentage of times the monitoring station was classified as "good status" (i.e., good or high status according to WFD) and linked this legal status to land cover. This dataset was also used as input to study the link between forest cover and water quality and test the influence of the stream size effect and temporal (seasonal and year) effect.

Following the results of the temporal effect analysis, we built an aggregated dataset. We aggregated variables' values by station over the entire 10-year dataset using two aggregation function types: the 90th quantile (and 10th quantile for the dissolved oxygen variable) and the median values. As results were highly similar, we only present the median aggregation results. Again, we applied Log- or square-based transformations improving the normality of variables' distribution was needed. We used this dataset to quantify the independent effect of forest cover and forest cover types (i.e., needle-leaved and broadleaved forests) on water quality relatively to effects of other LULC.

2.4. Land Use and Land Cover Data

Different spatial units can be considered in order to study the impact of land cover on water quality study. Some authors consider the entire upstream catchment as the spatial unit of LULC reference [37]. Others use the riparian zone (i.e., a buffer around the stream) to characterize the land cover impacting water stream quality [38]. Both approaches have some drawbacks. Indeed, when considering the upstream catchment in non-spatial statistical methods, the same importance is given to points irrespective of their distance to the monitoring station ignoring processes such as the self-purification of the stream. On the other hand, Giri and Qiu [21] point out problems with the riparian zone approach such as the absence of a uniform way to define its width or the fact that this zone does not represent nor behave ideally in terms of hydrological variation in the landscape. Pratt and Chang [39] also state that while riparian land cover affect water quality, a wider contributing area must be included to take into account distant sources of pollutants.

In this study, the area associated with the monitoring station is the intersection of the upstream catchment station (automatically extracted from the ERRUISSOL digital terrain model <http://geoportail.wallonie.be>) with the monitored surface water body defined by the WPS. This spatial unit of reference is thus directly based on the Walloon surface water body delineation which is the basic unit area for water quality assessment in line with the WFD directive [40]. This water

body delineation already integrates different variables as catchment size and hydromorphological parameters. In addition, this choice overcomes some above-mentioned drawbacks while providing a clear “rule” to apply across the whole region. This spatial unit of analysis will be further referred to as “sub-catchment”.

We used the Top10VGIS land cover data set provided by the Belgian National Geographic Institute (NGI, www.ngi.be) to characterize the land cover of the sub-catchments. This vector data set covers the whole of Belgium. It is based on the NGI topogeographic data that classify LULC into 37 classes. For the purpose of our study, we either selected the original land cover classes as such or aggregated them to end up with seven classes of interest: needle-leaved forest, broad-leaved forest, cropland, grassland, artificial surfaces, water surfaces, shrubs-heathlands and moorlands. We computed percentages of these classes in each sub-catchment. Following Brogna et al. [13], we assumed that the evolution of the retained classes in the region was minor throughout the studied decade.

To relate the LULC to water quality, we intersected the Top10VGIS dataset with each sub-catchment. To control the natural correlation between percentages of LULC variables, we constructed independent LULC variables by running Principal Component Analysis (PCA) on the main land cover classes’ percentage in the sub-catchments (i.e., needle-leaved forest, broad-leaved forest, grassland and cropland). We ran these PCA on standardized data, and extracted the stations coordinates on first and second components as independent LULC variables (i.e., LULC1 and LULC2).

2.5. Coupling Forest Cover and Water Quality

First, we performed a preliminary analysis on the seasonal dataset to assess the link between the main LULC of the region and the percentage of time over the decade each monitoring station was classified as “good status” according to current standards [40]. This provided a comprehensive picture of the relationships between LULC and policy standards for each water quality variables group (see Table 1). A station was considered in “good status” if it was classified as such for every component of the water quality variables group (see Table 1).

Then, we quantified the relationship between forest cover and physico-chemical water quality by applying a redundancy analysis (RDA, see Legendre and Legendre [41], R package vegan [42]) on LULC independent variables resulting from PCA (see above). This multivariate analysis allows capturing the linear relationship between dependent variables (i.e., physico-chemical variables, referred to as WQ in equations) and a matrix of explanatory variables (i.e., sub-catchments land cover variables). This analysis thus quantifies the percentage of water quality variability explained by LULC variables [42]. It allows quantifying and excluding the variability explained by certain covariates (e.g., season, year, upstream catchment area). We ran these RDA on centered and scaled variables because of the heterogeneity of the water quality variables units.

Analysis of the seasonal dataset explains the WQ matrix by the linear combination of the main land cover classes of interest:

$$WQ \sim LULC1 + LULC2 \quad (1)$$

where WQ = matrix of physico-chemical measurements (see Table 1); and LULC1 and LULC2 = independent LULC variables derived from the land cover PCA (i.e., sites scores on the first two axes of the LULC PCA).

However, following the interpretation of this first RDA, we simplified both the analysis and the results’ reading by applying an RDA directly on the percentage of forest cover in the sub-catchments (see Section 3.2).

$$WQ \sim \text{Forest} \quad (2)$$

where WQ = matrix of physico-chemical measurements (see Table 1), and Forest = percentage of forest cover in the sub-catchment (i.e., the sum of needle-leaved, broad-leaved and mixed forest percentage).

We tested the effect of the river size, directly linked to the discharge, to verify our choice to work with pollutant concentrations rather than loads, by putting it as covariate in the partial RDA under the proxy of whole upstream catchment area.

$$WQ \sim \text{Forest} + \text{Condition (upstream catchment area)} \quad (3)$$

We also tested the impact of temporality while considering the season, year and their interaction as covariates.

$$WQ \sim \text{Forest} + \text{Condition (season} + \text{year} + \text{season} \times \text{year)} \quad (4)$$

Following the analysis of the temporal variability impact, we ran the same analysis (Equation (2)) on aggregated water quality values (median value by station over the 10 years).

As spatial autocorrelation often explains an important part of biological structures [43], we partially controlled it by using “elevation” as covariate when quantifying the forest cover effect on water quality (Equation (5)). Indeed, there is a strong continuous ecological gradient in Belgium, that is highly correlated to elevation [44,45]. Dufrière and Legendre [44] show that elevation, although not exceeding 700 m in Belgium, explains almost all the geographic structure of several ecological variables given their spatial autocorrelation.

$$WQ \sim \text{Forest} + \text{Condition (Elevation)} \quad (5)$$

Finally, we distinguished effects from needle-leaved and broad-leaved forest covers while capturing the shared effects between LULC classes and distinguishing those from the elevation effect. We therefore ran several RDA with elevation as covariate and a variation partitioning between most important LULC and elevation [41]. The latter method divides the explained variance of a dataset between partial effects (i.e., the proportion of variation explained by a particular variable) and shared effects (i.e., variation that cannot be attributed to one variable but is shared between two or more).

3. Results

3.1. Forest Cover Versus Legal Standards

Figure 4 shows the biplot of the PCA ran on the main land cover classes' proportions at the sub-catchment level. The first two components explain 80% of the variability of the data set. The first component (LULC1), which explains 47.6% of the dataset variability, opposes grassland and cropland on the one hand (negative side), and needle-leaved and broad-leaved forests on the other hand (positive side). The second component (LULC2) explains 32.4% of the dataset variability and is mostly based on an opposition between grassland (positive side), and broad-leaved forest and cropland classes (negative side). Figure 4 illustrates the percentage of times that each monitoring station was classified in “good status” throughout the studied decade following current WFD standards (seasonal values). Except for the “mineralization” group (i.e., sulfates and chlorides) for which all stations are 100% of the time in “good status”, we observe a clear gradient linked to LULC. Regarding nitrogenous material (i.e., nitrates, nitrites and ammonium), the majority of stations that, most of the time, do not reach the “good status” have high cropland or grassland cover. Moreover, very few stations with high cropland cover are classified as “good status”, even for part of the decade. The same trend—even if less strong—is observed for the total phosphorus standards. Regarding oxygen balance (i.e., dissolved oxygen and dissolved organic carbon), stations that, most of the time, do not reach the “good status” are linked to high cropland cover (with a relatively important grassland cover).

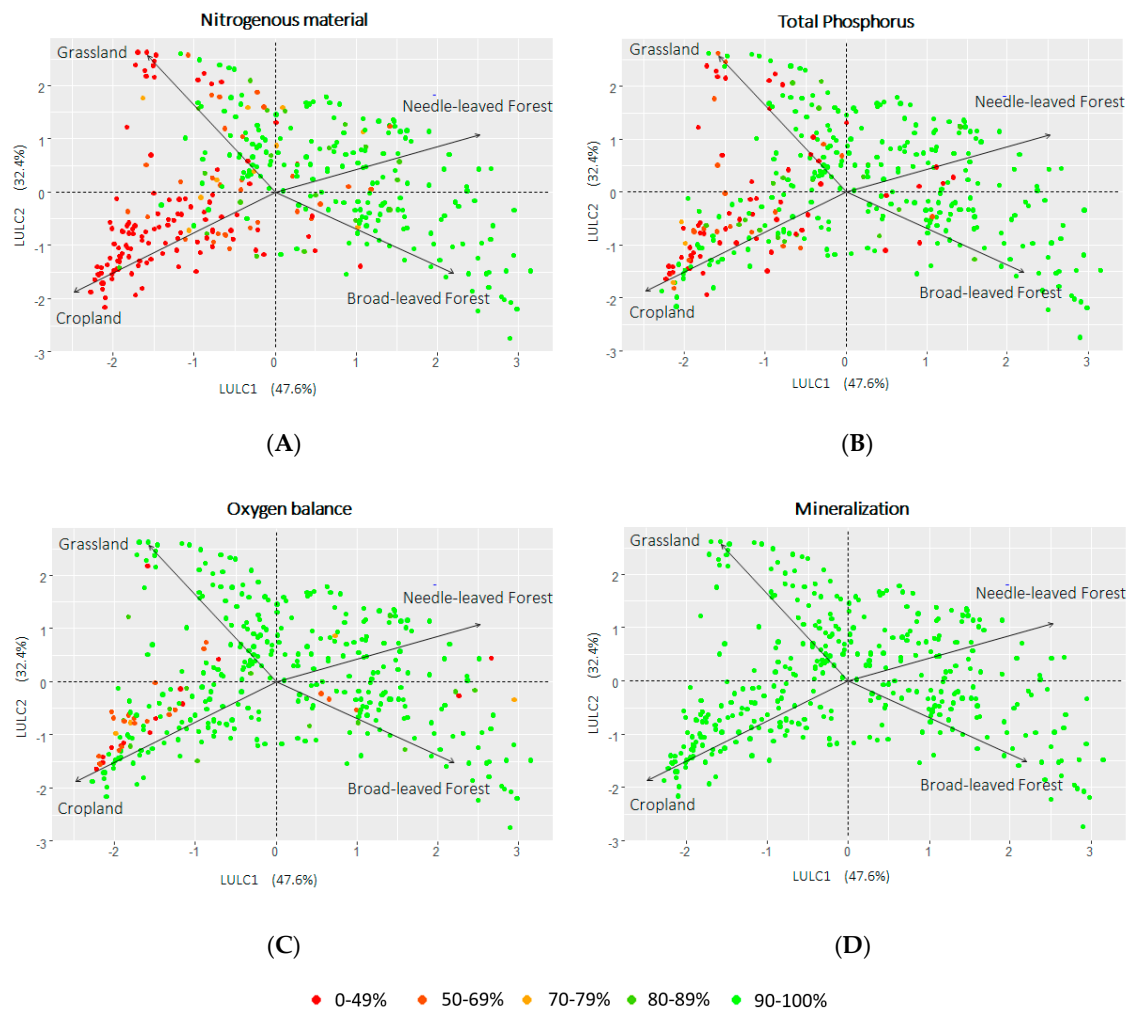


Figure 4. Biplot representing the monitoring stations from a PCA on the main land cover classes in studied sub-catchments. Colors represent the percentage of time the station was classified in “good status” according to current Water Framework Directive standards regarding: (A) the “nitrogenous material” group (i.e., nitrates, nitrites and ammonium); (B) total phosphorus; (C) the “oxygen balance” group (i.e., dissolved oxygen and dissolved organic carbon) and (D) the “mineralization” group (i.e., sulfates and chlorides).

3.2. Forest Cover Versus Multivariate Water Quality

Redundancy analysis on seasonal water quality values (Equation (1)) and LULC independent variables (see Figure 4) showed that the first constrained axis explains 43% of the water quality variability, and the second one explaining less than 1% of the water quality variability. This first axis is highly correlated to forest cover in sub-catchments. Consequently, following RDA in this study are directly based on forest cover percentage in sub-catchments. Table 2 (seasonal dataset) shows the percentage of variability explained by forest cover in these RDA either considering the forest cover only or removing some effects through the use of covariates. Temporal, river size and elevation effects have thus been removed successively. We applied permutation tests to each model produced in this study and all the presented models are significant.

Forest cover explains 29.6% of the seasonal physico-chemical water quality. The upstream catchment area effect is independent from the forest cover effect (shared effect of 0.3%) and explains 0.5% of the total variability of water quality. Temporal effect is also independent from the forest cover effect (shared effect of 1.2%) and explains 7% of total water quality variability. Consequently, we present below, and in the second part of Table 2, results from the aggregated dataset.

Table 2. Redundancy analysis results for seasonal dataset (October–March and April–September) and aggregated dataset (median value over the study period): variability (%) explained by forest cover, by covariates and shared effects.

Redundancy Analysis Covariate(s)	Variability Explained by Forest Cover (%)	Variability Shared between Forest Cover and Covariate(s) (%)	Variability Explained by Covariate(s) (%)
<i>Seasonal dataset</i>			
No covariates—[Equation (2)]	29.6	-	-
Covariate: upstream catchment area—[Equation (3)]	29.3	0.3	0.5
Covariates: season + year + season × year—[Equation (4)]	28.4	1.2	7
<i>Aggregated dataset</i>			
No covariates—[Equation (2)]	33.8	-	-
Covariate: Elevation—[Equation (5)]	9.3	24.5	12.9

Results of RDA (Equation (2)) on the aggregated dataset show that forest cover explains 33.8% of the water quality variability (see Table 2, aggregated dataset). Figure 5A–C presents these results in a factorial plan constituted by the constrained axis (RDA1) and the first residual axis (PC1).

Figure 5A shows the correlation circle of the RDA with “active” variables (water quality variables) and the constrained variable (forest cover percentage in sub-catchments). Forest cover is clearly linked to high stream water quality. Indeed, Dissolved Oxygen is positively correlated to forest cover while Ammonium, Nitrites, Total Phosphorus, Sulfates and Chloride concentrations are highly negatively correlated with forest cover. Dissolved Organic Carbon, pH, and Nitrates are also negatively correlated to forest cover but to a lesser extent. Figure 5B present the correlation circle from the RDA with both the constrained variable (i.e., forest cover) and “passive” variables (i.e., LULC classes and elevation). The constrained variable is obviously positively correlated with needle-leaved forest cover and broad-leaved forest cover in a less extent and inversely correlated to cropland cover. Grassland cover hardly contributes to the constrained axis. Figure 5C shows the position of stations in the same factorial plan. Elevation is also correlated with this constrained axis and RDA (see Table 2, aggregated dataset) shows that forest cover still explains 9.3% of water quality variability when the shared effect with elevation is removed. This is taken into account and refined in Section 3.3.

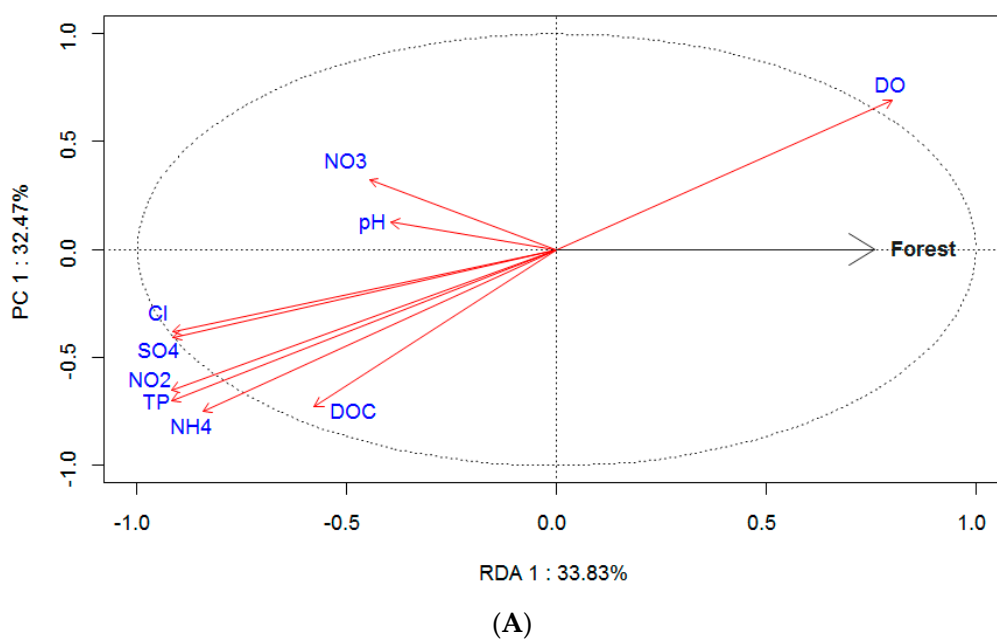


Figure 5. Cont.

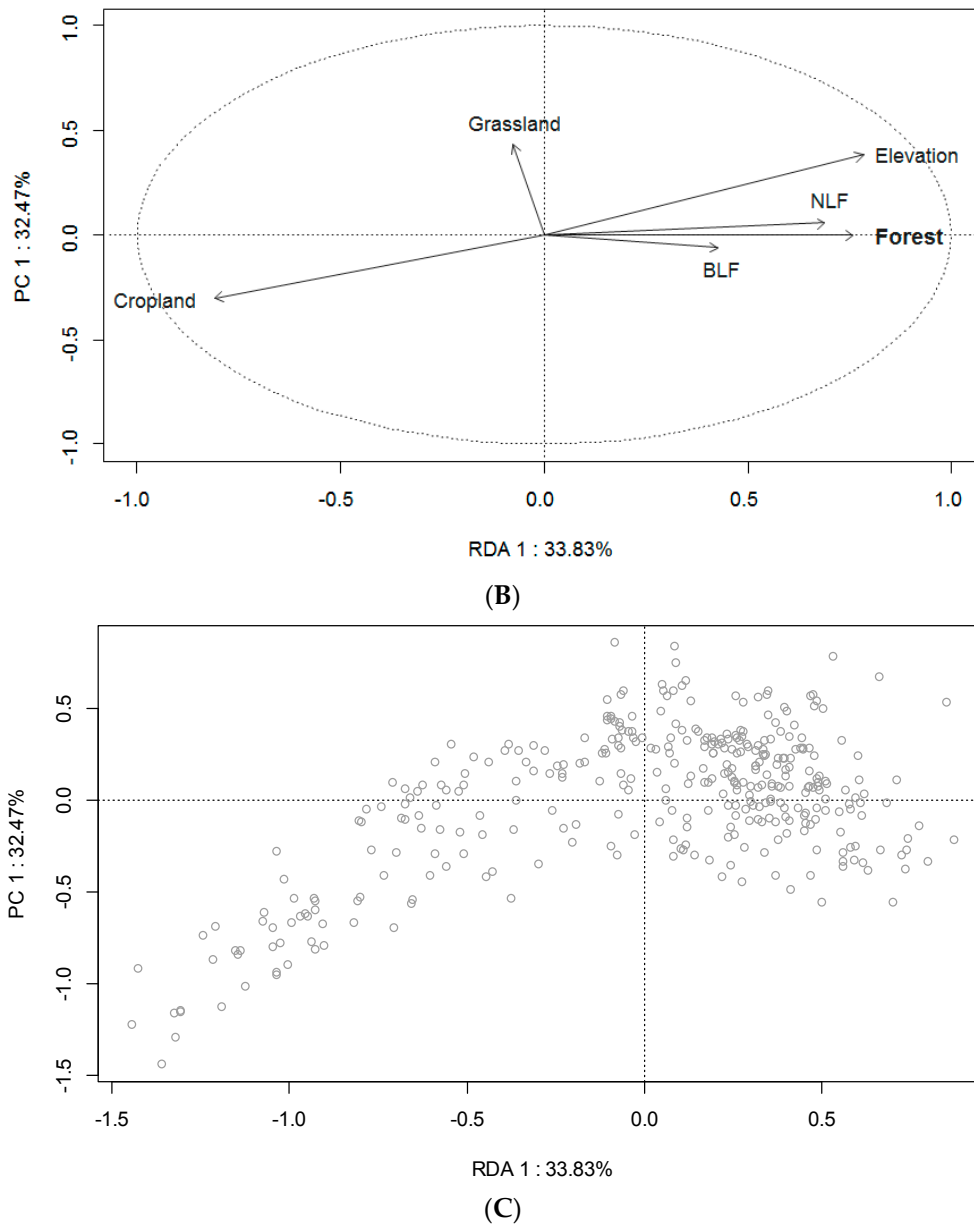


Figure 5. RDA results showing the link between forest cover and water quality. X-axis represents the constrained axis and y-axis, the first residual component. (A) Correlation circle of the RDA with “active” variables and the constrained variable (forest cover), with DO: Dissolved Oxygen, NO3: Nitrates, Cl: Chloride, SO4: Sulfates, TP: Total Phosphorus, NO2: Nitrites, NH4: Ammonium and DOC: Dissolved Organic Carbon. (B) Correlation circle of the RDA with constrained variable (forest cover) and “passive” variables (with NLF: Needle-leaved forest and BLF: Broad-leaved forest). (C) Stations location in the same plan.

3.3. Independent and Shared Effects of LULC Classes (Included Forest Type Distinction)

Results of RDA explaining water quality dataset with, on one the hand, each LULC class as constrained variable and, on the other hand, the same analysis but controlling for spatial autocorrelation (elevation as covariate) are presented in Table 3.

Table 3. Independent effect of main LULC (with distinction between forest types) on water quality and effect when removing shared effect with elevation, *F* and *p* values of the models.

Constrained variable—LULC (%) in sub-catchments	Independent Effect			with Elevation as Covariate		
	%	<i>F</i>	<i>p</i>	%	<i>F</i>	<i>p</i>
Needle-leaved Forest	29.3%	148.5	0.001 ***	2.3%	13.7	0.001 ***
Broad-leaved Forest	12.1%	49.4	0.001 ***	10.9%	75.2	0.001 ***
Cropland	38.7%	226.1	0.001 ***	7.9%	51.8	0.001 ***
Grassland	1.2%	4.5	0.013 ***	1.9%	11	0.001 ***

Note: ***: highly significant.

Regarding forest types effect, needle-leaved forest cover explains water quality much more (29.3% versus 12.1%) than broad-leaved forest but is highly correlated with elevation, as is cropland cover. Indeed, when removing shared effect of needle-leaved forest cover with elevation, the explained proportion drops from 29.3% to 2.3%. On the contrary, broad-leaved forest cover effect is stable and independent from the ecological gradient; this LULC class explaining the most (10.9%) when removing the elevation effect. As expected (see Figure 5B), grassland cover hardly explains water quality variability.

Variation partitioning (see Figure 6) splits the variation of water quality dataset into independent effects of LULC classes and elevation (i.e., the proportion of variation that is not shared with other variables) and shared interactions (i.e., interaction that cannot be attributed to a single class). Needle-leaved forest and broad-leaved forest independently explain 4.8 and 1.8% of the total water quality variability, respectively. Broad-leaved forest cover shares 5.8% of explained variability with cropland but only 0.8% with elevation. An important part of variability (21.3%) is shared by needle-leaved forest, cropland cover and elevation, whereas less than 1% is shared by broad-leaved forest, cropland cover and elevation.

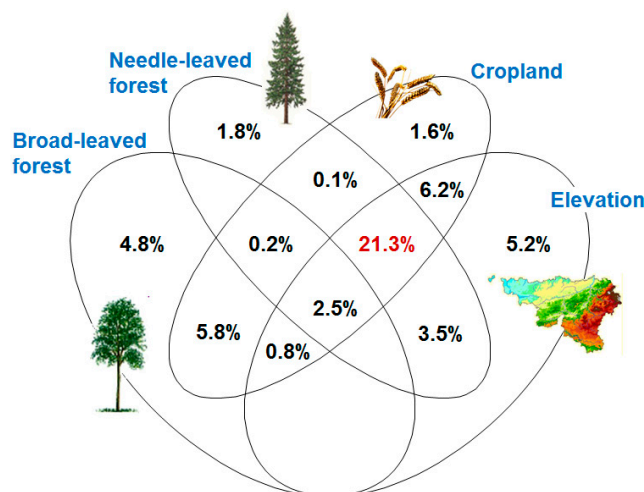


Figure 6. Venn diagram of the variance partitioning into main LULC classes and elevation. Figures are positively adjusted coefficients of determination and represent the variability explained by each subspace being either a single variable or shared effect between two or more variables.

4. Discussion

4.1. Forest Cover and Water Quality

This study assesses the link between forest cover and water quality data at the regional scale, by applying straightforward multivariate statistics on a large monitoring dataset. The analysis of sub-catchment's LULC and the legal water quality status of streams (objective (i)) shows that

sub-catchments with high forest cover tend to achieve “good status” over the studied decade more often than sub-catchments with high cropland and/or grassland covers. This is especially true for nitrogenous material and testifies that, despite the decrease in N input in agriculture since 1990, WFD target of “good status” is not yet fully reached. Sub-catchments with high grassland and cropland covers are also from far more polluted regarding phosphorus than forested sub-catchments. Nevertheless, stations with good phosphorus status are more frequent than in the case of nitrogenous materials. This analysis presents the advantage of providing an easily readable picture of the state of each station regarding water quality legal framework in relation with the LULC in its sub-catchment. This could provide a basis for further analyses to mitigate effects on water quality through improvement of catchment management. Indeed, some stations record unexpectedly high or low water quality in contrast to stations with similar LULC upstream signature. If an unexpected high water quality can be linked to particular practices or land planning measures, this could help managers to come up with locally relevant solutions.

Several insights were derived from the quantitative assessment of the link between forest cover and quality variables (objective (ii)). First, river size did not significantly affect the statistical relationship between LULC and water quality data. This allows aggregation of data from the entire monitoring network, with a high diversity of catchment sizes and thus discharges. It also justifies the use of concentrations instead of loads, unlike some authors such as de Oliveira et al. [10] suggest. This renders the methodology less complex and more easily replicable on a such large numbers of sampling stations. Second, seasonal and between-year effects were insignificant across the full decade. This entails that the link between LULC and water quality reflects a background “multi-pollutants” load that can be considered as temporally stable. A potential seasonal effect as observed in other studies on particular relationships between particular variables and LULC [7,8] might have been mitigated as the developed method treats several water quality variables together and as seasonal values are averaged values. For studies aiming to clearly focus on seasonal effects, we recommend specific and regular spatio-temporal sampling while focusing on homogeneous groups of pollutants regarding their seasonal variability (see e.g., Johnson et al. [46]). Analysis of the aggregated dataset over the studied decade showed that forest cover explains one third of the median water quality variability. Using elevation effect as a proxy for various environmental variables and as a mean for controlling spatial autocorrelation, we demonstrated an independent forest effect of 9.3%. Specifically, in this densely populated region with highly managed landscapes and forests, sub-catchments which are dominated by forest, and have lower agriculture and grassland cover, provide water with higher oxygen availability and lower concentrations of Ammonium, Nitrites, Nitrates, Total Phosphorus, Sulfates and Chloride and Dissolved Organic Carbon. This is confirming previous findings and reinforcing papers stating that forest cover is associated with higher water quality [17–19]. Main processes underlying these results are linked to the protection against erosion resulting in water with less sediments and fewer nutrients [15,16] which is also due to lower pressures compared to agricultural and urban land uses.

Finally, using the power of both a large monitoring dataset and multivariate statistics, we quantified the partial effect of forest cover types (i.e., needle-leaved and broadleaved forests) on water quality and shared effects with other LULC and environmental variables represented by the elevation variable (objective (iii)). This study empirically confirms a clear effect, independent from elevation, of broad-leaved cover on water quality (10.9%). The important effects of needle-leaved forest (29.3%) and cropland cover (38.7%) are largely shared with elevation and can therefore not be proven as independent effects. Regarding shared effect between different LULC classes and elevation, our analysis shows that broad-leaved forest and needle-leaved forest independently explain 4.8 and 1.8% of the total water quality variability respectively. Broad-leaved forest cover shares 5.8% of explained variability with cropland but only 0.8% with elevation. The biggest part of the explained variability is shared by needle-leaved forest cover, cropland cover and elevation (21.3%). A part of this variability is surely linked to forest cover effect but cannot be attributed to it.

4.2. Strengths and Limitations of the Study

Some limitations of this approach can be pointed out: (i) the methodological design and data used do not allow for isolating quantitatively a potential “active” effect of the forest (i.e., water purification per se) from the “passive” one being directly linked to the pressure degree of each LULC. (ii) The advantage of capturing the relationships between several water quality variables and forest cover implies that: (a) water quality variables with less frequent measurement during the studied decade could not be studied as they would have drastically reduced the number of observations; and (b) conclusions remain rather general and do not allow to discuss one particular variable in detail. However, we believe this approach has the following strengths: (i) it is based on public monitoring network data (several annual measurements for the processed pollutants) linked to the WFD and monitored in many other countries; (ii) as we did previously [13], we based this analysis on “real-life catchments” making conclusions more complex to draw but allowing for studying this phenomena at a regional scale and provide land planners with insights potentially contributing to a more sustainable resource management; and (iii) we applied a straightforward statistical approach, easier to apply and more efficient than physically based hydrologic/water quality models when observed data are limited in time but when datasets cover many different catchments. Furthermore, this statistical approach allows quantifying the effect of land cover on several pollutants concomitantly while controlling for autocorrelation and ecological factors, allowing for comparison between them. Finally, the statistical method was implemented in open source statistical software (R), which eases the replicability in other regions and/or through time.

5. Conclusions

Our study demonstrates significant effects of forest cover on water quality, disentangles independent and shared effects of correlated LULC categories while controlling for autocorrelation, and applies a method to mine large monitoring datasets. Capturing effects of land cover on several water quality variables at the same time from measured data allows for comparison between them. This contributes to validation and refining of the hypothesis that forests improve water quality.

Further research could focus on measuring direct LULC change impacts on water quality in study areas where more detailed and regularly updated land cover datasets are available. In addition, spatial information on land use practices could be further integrated into the analysis to enrich the interpretation. LULC effects can also be studied in more detail by taking into account spatial heterogeneity through landscape metrics studies, as some authors did [18,47,48], or, e.g., focusing on specific locations or types of forest such as riparian forests. Study of residuals and outliers could bring to light how catchment management can mitigate effects on water quality, as there are sub-catchments where water quality is unexpectedly high or low based on its LULC profile. Finally, similar analyses could be performed with biological water quality data, such as with macroinvertebrate [49] or diatom indices, which would bring complementary information.

The approach presented here, replicable in time and space, has a large application potential. First, it uses the publicly funded standard monitoring data linked to the WFD. Second, the analysis is based on “real-life” sub-catchments reflecting LULC heterogeneity and based on WFD water bodies, providing land planners and decision makers with directly applicable insights. Finally, we applied a straightforward statistical analysis, which are simpler, easier to apply, and more efficient than physically-based hydrologic/water quality models.

Supplementary Materials: The following is available online at www.mdpi.com/2073-4441/9/3/176/s1, Figure S1: LULC percentages in the sub-catchments with NL F needle-leaved forest, BL F: broad-leaved forest, S: Surfaces and MH: shrubs-heathlands and moorlands.

Acknowledgments: This work is connected to the Walloon Ecosystem Services platform (WALES) and the Belgian Ecosystem service community of practice (www.beescommunity.be). We acknowledge the Walloon Public Service (DGO2, Surface Water Department) for providing physico-chemical water quality datasets, the Belgian Geographic National Institute for providing land cover datasets.

Author Contributions: Delphine Brogna conceived the methodology, analyzed, interpreted the data and wrote the paper. Adrien Michez, Sander Jacobs, Marc Dufrêne, Caroline Vincke and Nicolas Dendoncker contributed to the methodology conception and the manuscript redaction. Adrien Michez, Sander Jacobs and Marc Dufrêne contributed to the data analysis and results interpretation. Nicolas Dendoncker and Caroline Vincke provide an academic supervision to the author.

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

References

1. United Nations Water. *A Post-2015 Global Goal for Water: Synthesis of Key Findings and Recommendations from UN-Water*; UN-Water: Genève, Suisse, 2014; p. 41.
2. Haddadin, M.J. Water Scarcity Impacts and Potential Conflicts in the MENA Region. *Water Int.* **2001**, *26*, 460–470. [[CrossRef](#)]
3. Vörösmarty, C.J.; McIntyre, P.B.; Gessner, M.O.; Dudgeon, D.; Prusevich, A.; Green, P.; Glidden, S.; Bunn, S.E.; Sullivan, C.A.; Liermann, C.R.; Davies, P.M. Global threats to human water security and river biodiversity. *Nature* **2010**, *467*, 555–561. [[CrossRef](#)] [[PubMed](#)]
4. Meybeck, M. Global analysis of river systems: From Earth system controls to Anthropocene syndromes. *Philos. Trans. R. Soc. B Biol. Sci.* **2003**, *358*, 1935–1955. [[CrossRef](#)] [[PubMed](#)]
5. Millennium Ecosystem Assessment (Program). *Ecosystems and Human Well-Being: Synthesis*; Island Press: Washington, DC, USA, 2005.
6. Carpenter, S.R.; Caraco, N.F.; Correll, D.L.; Howarth, R.W.; Sharpley, A.N.; Smith, V.H. Nonpoint Pollution of Surface Waters with Phosphorus and Nitrogen. *Ecol. Appl.* **1998**, *8*, 559–568. [[CrossRef](#)]
7. Álvarez-Cabria, M.; Barquín, J.; Peñas, F.J. Modelling the spatial and seasonal variability of water quality for entire river networks: Relationships with natural and anthropogenic factors. *Sci. Total Environ.* **2016**, *545–546*, 152–162.
8. Chen, Q.; Mei, K.; Dahlgren, R.A.; Wang, T.; Gong, J.; Zhang, M. Impacts of land use and population density on seasonal surface water quality using a modified geographically weighted regression. *Sci. Total Environ.* **2016**, *572*, 450–466. [[CrossRef](#)] [[PubMed](#)]
9. Yu, D.; Shi, P.; Liu, Y.; Xun, B. Detecting land use-water quality relationships from the viewpoint of ecological restoration in an urban area. *Ecol. Eng.* **2013**, *53*, 205–216. [[CrossRef](#)]
10. De Oliveira, L.M.; Maillard, P.; de Andrade Pinto, É.J. Modeling the effect of land use/land cover on nitrogen, phosphorous and dissolved oxygen loads in the Velhas River using the concept of exclusive contribution area. *Environ. Monit. Assess.* **2016**, *188*, 333. [[CrossRef](#)] [[PubMed](#)]
11. Hwang, S.-A.; Hwang, S.-J.; Park, S.-R.; Lee, S.-W. Examining the Relationships between Watershed Urban Land Use and Stream Water Quality Using Linear and Generalized Additive Models. *Water* **2016**, *8*, 155. [[CrossRef](#)]
12. Carvalho-Santos, C.; Honrado, J.P.; Hein, L. Hydrological services and the role of forests: Conceptualization and indicator-based analysis with an illustration at a regional scale. *Ecol. Complex.* **2014**, *20*, 69–80. [[CrossRef](#)]
13. Brogna, D.; Vincke, C.; Brostaux, Y.; Soyeurt, H.; Dufrêne, M.; Dendoncker, N. How does forest cover impact water flows and ecosystem services? Insights from “real-life” catchments in Wallonia (Belgium). *Ecol. Indic.* **2017**, *72*, 675–685.
14. Brauman, K.A.; Daily, G.C.; Duarte, T.K.; Mooney, H.A. The Nature and Value of Ecosystem Services: An Overview Highlighting Hydrologic Services. *Annu. Rev. Environ. Resour.* **2007**, *32*, 67–98. [[CrossRef](#)]
15. Neary, D.G.; Ice, G.G.; Jackson, C.R. Linkages between forest soils and water quality and quantity. *For. Ecol. Manag.* **2009**, *258*, 2269–2281. [[CrossRef](#)]
16. The Economics of Ecosystems and Biodiversity (TEEB). *The Economics of Ecosystems and Biodiversity Ecological and Economic Foundations*; Kumar, P., Ed.; Earthscan: London, UK; Washington, DC, USA, 2010.
17. Figuepron, J.; Garcia, S.; Stenger, A. Land use impact on water quality: Valuing forest services in terms of the water supply sector. *J. Environ. Manag.* **2013**, *126*, 113–121. [[CrossRef](#)] [[PubMed](#)]
18. Łowicki, D. Prediction of flowing water pollution on the basis of landscape metrics as a tool supporting delimitation of Nitrate Vulnerable Zones. *Ecol. Indic.* **2012**, *23*, 27–33. [[CrossRef](#)]
19. Tong, S.T.Y.; Chen, W. Modeling the relationship between land use and surface water quality. *J. Environ. Manag.* **2002**, *66*, 377–393. [[CrossRef](#)]

20. Vose, J.M.; Sun, G.; Ford, C.R.; Bredemeier, M.; Otsuki, K.; Wei, X.; Zhang, Z.; Zhang, L. Forest ecohydrological research in the 21st century: What are the critical needs? *Ecohydrology* **2011**, *4*, 146–158. [CrossRef]
21. Giri, S.; Qiu, Z. Understanding the relationship of land uses and water quality in Twenty First Century: A review. *J. Environ. Manag.* **2016**, *173*, 41–48. [CrossRef] [PubMed]
22. Chauhan, A.; Verma, S. Impact of Agriculture, Urban and Forest Land Use on Physico-Chemical Properties of Water a Review. *Int. J. Curr. Microbiol. Appl. Sci.* **2015**, *4*, 18–22.
23. Dhondt, K.; Boeckx, P.; Verhoest, N.E.C.; Hofman, G.; Van Cleemput, O. Assessment of temporal and spatial variation of nitrate removal in riparian zones. *Environ. Monit. Assess.* **2006**, *116*, 197–215. [CrossRef] [PubMed]
24. Carnol, M.; Ineson, P. Environmental factors controlling NO₃⁻ leaching, N₂O emissions and numbers of NH₄⁺ oxidisers in a coniferous forest soil. *Soil Biol. Biochem.* **1999**, *31*, 979–990. [CrossRef]
25. Sohler, C.; Degré, A. Modelling the effects of the current policy measures in agriculture: An unique model from field to regional scale in Walloon region of Belgium. *Environ. Sci. Policy* **2010**, *13*, 754–765. [CrossRef]
26. Lin, Z.; Anar, M.J.; Zheng, H. Hydrologic and water-quality impacts of agricultural land use changes incurred from bioenergy policies. *J. Hydrol.* **2015**, *525*, 429–440. [CrossRef]
27. Jomaa, S.; Jiang, S.; Thraen, D.; Rode, M. Modelling the effect of different agricultural practices on stream nitrogen load in central Germany. *Energy Sustain. Soc.* **2016**, *6*, 1–16. [CrossRef]
28. Carvalho-Santos, C.; Nunes, J.P.; Monteiro, A.T.; Hein, L.; Honrado, J.P. Assessing the effects of land cover and future climate conditions on the provision of hydrological services in a medium-sized watershed of Portugal. *Hydrol. Process.* **2015**. [CrossRef]
29. Wan, R.; Cai, S.; Li, H.; Yang, G.; Li, Z.; Nie, X. Inferring land use and land cover impact on stream water quality using a Bayesian hierarchical modeling approach in the Xitiaoxi River Watershed, China. *J. Environ. Manag.* **2014**, *133*, 1–11. [CrossRef] [PubMed]
30. R Core Team. *R: A Language and Environment for Statistical Computing*; R Foundation for Statistical Computing: Vienna, Austria, 2013.
31. Le Réseau de Surveillance de la Qualité des Eaux de Surface—Portail “Directive-Cadre sur l’Eau”—DGO3/SPW. Available online: <http://eau.wallonie.be/spip.php?article120> (accessed on 16 December 2016).
32. SPW-DGO3-Direction de l’Etat Environnemental. *Les Indicateurs clés de L’environnement Wallon 2014 (ICEW 2014)*; SPW: Namur, Belgique, 2014.
33. SPW-DGO3-Direction de l’Etat Environnemental. *Rapport Analytique sur L’état de L’environnement Wallon 2006–2007*; SPW: Namur, Belgique, 2007; 736p.
34. Alderweireld, M.; Burnay, F.; Pitchugin, M.; Lecomte, H. *Inventaire Forestier Wallon—Résultats 1994-2012*; SPW: Jambes, Belgique, 2015.
35. Gailliez, S. Estimation des Débits D’étiage pour des Sites non Jaugés. Application en Région Wallonne. Ph.D. Thesis, Université de Liège—Gembloux Agro-Bio Tech, Gembloux, Belgique, 2013.
36. Olsen, R.L.; Chappell, R.W.; Loftis, J.C. Water quality sample collection, data treatment and results presentation for principal components analysis—Literature review and Illinois River watershed case study. *Water Res.* **2012**, *46*, 3110–3122. [CrossRef] [PubMed]
37. Tu, J. Spatial variations in the relationships between land use and water quality across an urbanization gradient in the watersheds of northern Georgia, USA. *Environ. Manag.* **2013**, *51*, 1–17. [CrossRef] [PubMed]
38. Li, S.; Gu, S.; Tan, X.; Zhang, Q. Water quality in the upper Han River basin, China: The impacts of land use/land cover in riparian buffer zone. *J. Hazard. Mater.* **2009**, *165*, 317–324. [CrossRef] [PubMed]
39. 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] [PubMed]
40. SPW-DGO3-Direction de l’Etat Environnemental. *Deuxièmes Plans de gestion des Districts Hydrographiques—Document Général; Mise en Œuvre de la Directive-Cadre sur l’Eau (2000/60/CE)—Région wallonne*; SPW: Namur, Belgique, 2016.
41. Legendre, P.; Legendre, L. Chapter 11—Canonical analysis. In *Numerical Ecology*; Legendre, P., Legendre, L.F.J., Eds.; Elsevier: Amsterdam, The Netherlands, 2012; Volume 24, pp. 625–710.
42. Oksanen, J.; Blanchet, F.G.; Friendly, M.; Kindt, R.; Legendre, P.; McGlinn, D.; Minchin, P.R.; O’Hara, R.B.; Simpson, G.L.; Solymos, P.; et al. *Vegan: Community Ecology Package*. 2016. Available online: <https://r-forge.r-project.org/> (accessed on 8 February 2016).
43. Legendre, P. Spatial Autocorrelation: Trouble or New Paradigm? *Ecology* **1993**, *74*, 1659–1673. [CrossRef]

44. Dufrene, M.; Legendre, P. Geographic structure and potential ecological factors in Belgium. *J. Biogeogr.* **1991**, *18*, 257–266. [[CrossRef](#)]
45. Noirfalise, A. *Les Régions Naturelles de la Belgique*; No. 23; C.E.F.R.: Gembloux, 1988.
46. Johnson, L.; Richards, C.; Host, G.; Arthur, J. Landscape influences on water chemistry in Midwestern stream ecosystems. *Freshw. Biol.* **1997**, *37*, 193–208. [[CrossRef](#)]
47. Clément, F.; Ruiz, J.; Rodríguez, M.A.; Blais, D.; Campeau, S. Landscape diversity and forest edge density regulate stream water quality in agricultural catchments. *Ecol. Indic.* **2017**, *72*, 627–639. [[CrossRef](#)]
48. Amiri, B.J.; Nakane, K. Modeling the Linkage between River Water Quality and Landscape Metrics in the Chugoku District of Japan. *Water Resour. Manag.* **2008**, *23*, 931–956. [[CrossRef](#)]
49. Miserendino, M.L.; Pizzolon, L.A. Interactive effects of basin features and land-use change on macroinvertebrate communities of headwater streams in the Patagonian Andes. *River Res. Applic.* **2004**, *20*, 967–983. [[CrossRef](#)]



© 2017 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 (<http://creativecommons.org/licenses/by/4.0/>).