A Meta-Analysis on Nitrogen Retention by Buffer Zones

Elena Valkama,* Kirsi Usva, Merja Saarinen, and Jaana Uusi-Kämppä

Abstract

Buffer zones, established between agricultural fields and water bodies, are widely used as a measure to reduce N in surface runoff and groundwater. However, the literature indicates inconsistent results on the N removal efficiency of buffer zones between studies. We performed a weighted meta-analysis on the buffer zone effects on NO\textsubscript{3}–N and total N in surface runoff and groundwater by summarizing 46 studies published between 1980 and 2017. The overall effects of buffer zones were a 33% (−48 to −17%, n = 25) and 70% (−78 to −62%, n = 38) NO\textsubscript{3}–N reduction in surface runoff and in groundwater, respectively, compared with controls with no buffer zone. In addition, buffer zones reduced the total N in surface runoff by 57% (−68 to −43%, n = 16). The effects of buffer zones on N retention were consistent across continents and in different climates. Nitrogen retention increased with increasing initial N concentrations discharged from the source of pollution. According to a meta-regression, the N removal efficiency in surface runoff decreased in consort with increasing buffer zone age. Otherwise, the meta-analysis revealed no effects of buffer zone characteristics such as the width or species number (for grass buffer zones) on the N retention in surface runoff and groundwater. Unlike groundwater quality, which responded equally well regardless of the source of pollution, buffer zone type, or buffer zone age, surface water quality is more sensitive, and it might not be satisfactorily improved by tree buffer zones or aged buffer zones, or when the source of pollution originates from grass production fields.

Core Ideas

• We performed a global weighted meta-analysis by summarizing 46 studies (1980–2017).
• Buffer zones reduced NO\textsubscript{3}–N by 33% in surface runoff and by 70% in groundwater.
• Surface water is more sensitive to buffer zone characteristics than groundwater.
• The higher the initial N concentrations, the larger the N retention by the buffer zones.
• Meta-analysis revealed no effect of buffer zone width on N retention.

Buffer zones, established between agricultural fields and water bodies, are widely used for many purposes such as sustaining hydrological connectivity, maintaining habitat and ecological connectivity, and for biomass production and C sequestration, as well as for providing cultural services and recreational value for public (Stutter et al., 2012). The greatest attention is given to controlling the transport of anthropogenic inputs of eroded soil particles and nutrients into receiving waters. For these reasons, buffer zones have been considered a useful measure, serving among other purposes to reduce the N in surface water and groundwater. In Europe, the Water Framework Directive was launched in 2000 to ensure that all surface water bodies achieve at least a good ecological status. Each member state is mandated to develop a River Basin Management Plan for each river basin district. According to the European Environmental Bureau (2010), for some rivers, such as the Meuse (the Netherlands) and the Scheldt (Belgium), the installation of buffer zones are the only new measure included in the River Basin Management Plan.

The efficiency of buffer zones has been intensively studied for the last 30 yr, but the earliest studies were already conducted in the 1950s (Haddaway et al., 2018). The results have been summarized in qualitative and narrative reviews, demonstrating a large variability in N removal efficiency, from 20 to 100% (Vought et al., 1994; Fennessy and Cronk, 1997; Mander et al., 1997; Weng, 1999; Fischer and Fischenich, 2000; Lyons et al., 2000; Hawes and Smith, 2005; Dosskey et al., 2010; Pärn et al., 2012; Christen and Dalgaard, 2013; Sweeney and Newbold, 2014). It has also been noted that occasionally the buffers may become saturated and act as net N sources (Pärn et al., 2012).

The most frequently discussed question is what construes a sufficient buffer zone width. Fischer and Fischenich (2000) wrote: "Although many buffer strip width recommendations tend to be arbitrary or based on anecdotal information, the scientific literature is replete with recommendations for maintaining or improving water quality in a variety of different settings." For example, for low to moderate slopes, Hawes and Smith (2005) recommended buffer zones <10 m, where the most filtering occurs. Vought et al. (1994) confirmed that a buffer strip between

Abbreviations: CI, confidence interval.
and statistical models that take into account the distinct hierarchical structure of meta-analytic data (Gurevitch et al., 2018). The implementation of a meta-analysis requires special care, and its value may be greatly reduced by the use of inappropriate techniques. A pitfall in analyzing a number of independent studies lies in their methodological diversity. In addition, the study-specific sampling error variances are almost never identical across studies, violating the underlying assumptions of traditional statistical analysis (Gurevitch and Hedges, 1999).

The present study aimed to summarize global studies on the retention of NO$_3^-$–N and the total N in surface runoff and in groundwater by using a meta-analysis. We examined the source of variation in retention capacity, such as the design and duration of experiments, as well as the metrics of outcomes (concentrations and loads), N forms, climates and locations, sources of pollution, the N concentrations entering the buffer zones, soil texture, buffer zone vegetation and species number, as well as the buffer zone slope and width.

Materials and Methods

The Database

We found the articles by searching for keywords “buffer zones” or “buffer strips” or “filter strips” or “vegetative strips” or “riparian forest buffers” or “riparian zones” or “vegetated buffer strips” AND “nitrate” or “NO$_3^-$–N” or “nitrogen” or “nitrogen leaching” or “nitrate leaching” in the Web of Science Database, in addition to Scopus, and ScienceDirect. We also found journal articles in the reference lists of previously published articles.

To be included in the database, a study had to meet the following criteria:

1. A study was conducted in the field concerning natural or artificial runoff.
2. The sources of pollution were agricultural fields for grass or cereal production, natural pasture or feedlots.
3. The study had an appropriate control group without buffer zone: (a) control plots in surface runoff studies, which were generally arranged in randomized block design; (b) field edge (above buffer zone) in groundwater monitoring assessment along a vegetation transect; (c) control waterways in a paired watershed comparison studies (monitored in both surface and groundwater studies).
4. The buffer zone was nonfertilized.
5. The study assessed the buffer zone effects on NO$_3^-$–N or the total N in the surface runoff or on NO$_3^-$–N in the groundwater.
6. The NO$_3^-$–N or total N were recorded as either original data for each experimental year, or as a sample or replicate, or as means of treatment (i.e., with a buffer zone) and control (i.e., with no buffer zone) with SDs and sample sizes.

Data were extracted from tables and digitized from figures using the ImageJ 1.37 program (Schneider et al., 2012). However, ~200 studies from those located were not included in the database used in this study because means, SDs, or sample sizes for controls and buffer zone treatments were not possible to extract. In most articles, the results were reported in terms of the removal efficiency of the buffer zone (%), which made the calculation of effect sizes and performance of a meta-analysis impossible.
The final database used for this study consisted of 46 studies published between 1980 and 2017 in peer-reviewed scientific journals (Valkama et al., 2018; Supplemental Tables S1 and S2). Altogether, 33 studies were conducted in the United States, three in the Netherlands, two in Australia and Finland, and one in the United Kingdom, Italy, Canada, New Zealand, Kenya, and China. The major climates in this range of reported studies were humid continental and humid subtropical, but also Mediterranean and oceanic ( Köppen climate classification).

The sources of pollution included (i) tilled or nontilled cereals, mainly barley (Hordeum vulgare L.), oats (Avena sativa L.), corn (Zea mays L.), and soybeans (Glycine max (L.) Merr.), but also wheat (Triticum aestivum L.), fodder maize (Zea mays L.), and rice (Oryza sativa L.); (ii) grazed or nongrazed grasses, such as timothy grass (Phleum pratense L.), meadow fescue (Schedonorus pratensis (Huds.) P. Beauv.), perennial ryegrass (Lolium perenne L.), clover (Trifolium spp.), 'Coastal' bermudagrass (Cynodon dactylon (L.) Pers.), natural pasture; or (iii) simulated or natural feedlots.

The soils were mainly loam (silt loam, sandy clay loam, silty clay loam, sandy loam, and fine sandy loam), but also sand and clay. Manure or inorganic fertilizer was applied at rates of 50 to 470 kg N ha⁻¹. In eight studies, there was no information on manure or inorganic fertilizer applied.

In the majority of the studies, the buffer zones included one, two, or multiple species of grasses (timothy, meadow fescue, orchardgrass [Dactylis glomerata L.], tall fescue [Schedonorus arundinaceus (Schreb.) Dumort.], switchgrass [ Panicum virgatum L.], clover, perennial and annual [Lolium multiflorum Lam.] ryegrasses, Kentucky bluegrass [Poa pratensis L.], cane [Arundinaria Michx. spp.], bromes [Bromus L. spp.], Indiangrass [Sorghastrum nutans (L.) Nash], witchgrass [Panicum capillare L.], redtop [Agrostis gigantea Roth], birds’-foot trefoil [Lotus corniculatus L.], Japanese lawngrass [ Zoysia japonica Steud.], Coastal bermudagrass, reed canarygrass [ Phalaris arundinacea L.], barnyardgrass [ Echinochloa crus-galli (L.) P. Beauv.], sedges [Carex L. spp.], dogfennel [ Eupatorium capillifolium (Lam.) Small], Japanese stilgrass [ Microstegium vimineum (Trin.) A. Camus], spotted jep pye weed [ Euthrochium maculatum (L.) E.E. Lamont], aster [Aster L. spp.], Canadian anemone [Anemone canadensis L.], goldenrods [ Solidago spp.], and common bent [ Agrostis capillaris L.], and also one, two, or multiple species of deciduous trees (e.g., birch [Betula L. spp.]), alder [Alnus Mill. spp.], maple [Acer L. spp.], ash [ Fraxinus L. spp.], oak [ Quercus L. spp.], poplar [ Populus L. spp.], willow [Salix L. spp.], elder [Sambucus L. spp.], and others) and shrubs (cherry [Prunus virginiana L.], plum [Prunus americana Marsh.], and dogwood [Cornus L. spp.]). Eleven articles reported the effects of multispecies buffer zones, which were arranged into several zones, such as grass, shrub, and/or tree zones. In surface runoff studies, the buffer zone width varied from 2 to 30 m, but one study had a width of 113 m. In groundwater studies, buffer zone width varied from 2 to 300 m. Buffer zone slope varied from 0.9 to 20% (surface runoff) and from 0 to 15% (groundwater). If slope range was reported, the largest value was extracted for meta-regression.

Surface water samples were collected from natural or artificial runoffs by using lysimeters or runoff collectors. Shallow and deep groundwater samples were collected from natural rainfall by using wells, ceramic cups, or piezometers along transects in each vegetation type. A randomized (complete) block design was used mostly in studies on surface runoff. Other designs were also used (e.g., paired comparisons or monitoring assessment along vegetation transects including field edge [above the buffer zone], as in the majority of the groundwater studies). The duration of experiments of natural runoff varied from 1 to 10 yr (surface runoff), or from 1 to 15 yr (groundwater). The buffer zone age at the end of experiment was assessed as the sum of buffer zone age at the beginning of experiment and the duration of experiment, and it varied from 2 to 31 yr (for surface runoff studies) and from 2 to 49 yr (for the groundwater studies).

**Response and Explanatory Variables**

In surface runoff, concentrations (mg L⁻¹), or loads (kg ha⁻¹) of NO₃⁻N or total N were measured in 22 studies, resulting in 41 observations of natural (14 studies) or artificial (eight studies) runoff. For groundwater, mainly concentrations of NO₃⁻ N in natural runoff were measured in 25 studies, resulting in 38 observations. To explain the variation in the N changes due to the buffer zones, we included the categorical and continuous explanatory variables listed in Supplemental Table S3.

**Meta-analysis**

We used MetaWin 2.0 statistical software (Rosenberg et al., 2000) to carry out the meta-analysis. Quantitative meta-analysis involves calculating an effect size (i.e., the magnitude of the treatment effect) that can be averaged across independent studies, thereby yielding an overall mean effect size.

For the response variables, we calculated separate estimates of the response ratio (r) as an index of the effect size:

\[
 r = \frac{\overline{X}_{BZ}}{\overline{X}_C}
\]

where \( \overline{X}_{BZ} \) and \( \overline{X}_C \) represent the means for treatments with a buffer zone and with no buffer zone (i.e., the control groups), respectively, averaged for replicates or samples for short-term experiments, and over the duration (years) of long-term experiments. If a study reported the results of NO₃⁻N measurements in surface runoff and groundwater, both values were extracted. Similarly, the data were extracted for each site, source of pollution, and buffer zone type per study. However, if the results for several buffer zone widths or depths of wells were reported in a study, data were extracted for one randomly selected width or depth to avoid problem with non-independence of the effect sizes.

Since distribution of \( r \) is skewed, performing statistical analyses in the metric of the natural logarithm of \( r \) \( \ln(r) \) is preferable due to its much more normal distribution in small samples than that of \( r \) (Hedges et al., 1999):

\[
 \ln(r) = \ln \left( \frac{\overline{X}_{BZ}}{\overline{X}_C} \right)
\]

We calculated the variance of \( \ln(r) \) as

\[
 \nu_{\ln(r)} = \frac{(SD_{BZ})^2}{n_{BZ}(\overline{X}_{BZ})^2} + \frac{(SD_C)^2}{n_C(\overline{X}_C)^2}
\]

272 Journal of Environmental Quality
where SD\(_{BZ}\) and SD\(_{C}\) are the corresponding SDs for treatments with buffer zones and the control, respectively, and \(n_{BZ}\) and \(n_{C}\) are the replicates or samples for short-term experiments and the duration (years) for long-term experiments for treatments with buffer zones and the control, respectively.

We calculated the weighted mean of the ln(\(r\)) for all studies as

\[
\text{ln}(r) = \frac{\sum_{i=1}^{n} w_i \text{ln}(r_i)}{\sum_{i=1}^{n} w_i}
\]

where \(w_i\) is the weight for study \(i\), defined by the reciprocal of the sample variance (\(w_i = 1/v_i\)), \(\text{ln}(r_i)\) is the log response ratio for study \(i\), and \(n\) is the number of studies. Because the variance of the effect sizes is a function of the sample size (Eq. [3]), studies with a larger sample size had lower variances and received heavier weights. In computing means or regression coefficients, weighting improves the precision of the combined estimates and the power of the tests. Weighting also helps to ensure similar sampling distributions for different statistics (e.g., \(\chi^2\) test).

A random effects model served to combine estimates across the studies, assuming that studies in each subgroup do not share the same effect size. We used a bootstrap statistical method (Efron and Tibshirani, 1986) to generate bias-corrected 95% CIs around the ln(\(r\)) from 4999 iterations. To test whether ln(\(r\)) differed between the groups of categorical explanatory variables, we used the \(\chi^2\) test to examine the between-group heterogeneity (\(Q_B\)), as well as to check for possible intercorrelation between the variables. To study the effect of continuous explanatory variables, we ran weighted meta-regressions with ln(\(r\)) as the dependent variable and the continuous variables as independent ones. We also used the \(\chi^2\) test to examine model heterogeneity (\(Q_M\)), which describes the amount of heterogeneity explained by the regression models.

Except for meta-regression, we back-transformed the ln(\(r\)) values and reported them in the text and figures as percentage changes from the controls:

\[
\text{Response} \% = \{\exp[\text{ln}(r)] - 1\} \times 100%
\]

The buffer zone effects on the reduction of NO\(_3\)–N and total N were considered significantly different from the controls if the 95% CIs did not overlap with zero, and significantly different between the groups of explanatory variables if their 95% CIs did not overlap.

**Results**

**Overall Effects**

A histogram of the changes in NO\(_3\)–N and total N in surface runoff due to buffer zones in all studies indicates a wide range of responses from −90 to 55% compared with the controls with no buffer zone (Fig. 1a). The summarized effect of buffer zones was a 33% reduction in NO\(_3\)–N (95% CI = −48 to −17%, \(n = 25\)) and a 57% reduction in total N (95% CI = −68 to −43%, \(n = 16\)). A histogram of the changes in NO\(_3\)–N in groundwater, measured mainly as concentrations, showed a clear shift toward negative values, from −95 to −5%, with a summarized effect of a 70% (95% CI = −78 to −62%, \(n = 38\)) reduction compared with controls with no buffer zones (Fig. 1b).

In the surface runoff, the N changes due to buffer zones did not differ in the outcomes of studies measured as loads and concentrations (\(Q_B = 1.34\), df = 1, \(P = 0.329\)), or those conducted on artificial and natural runoffs (\(Q_B = 0.06\), df = 1, \(P = 0.828\)), or arranged into randomized (complete) block design and other design (\(Q_B = 0.69\), df = 1, \(P = 0.472\)). Similarly, the design of experiments in groundwater studies had no impact on the outcomes (\(Q_B = 0.06\), df = 1, \(P = 0.831\)). Therefore, we pooled the data for further analysis.

**Initial Nitrogen Concentrations**

The range of initial N concentrations and loads from the different sources of pollution without buffer zones appears in Table 1. Fields under cereal production, often tilled in the autumn, and...
feedlots with compact soils receiving a large amount of manure, impose a several times larger N contamination on the surface runoff and groundwater compared with fields under grass production.

A meta-regression indicated that the N retention by the buffer zone from surface runoff and groundwater increased with increasing N concentrations entering the buffer zone from the source of pollution (Fig. 2a and 2b). For example, with increasing N concentration from 0.1 to 25 mg L\(^{-1}\), the N retention of surface runoff increased from 8 to 45% (Fig. 2a) and of groundwater from 60 to 85% (Fig. 2b).

### Pollution source, Soil Texture, and Climatic Conditions

Variation in N retention of buffer zones concerning the source of pollution, soil texture, and compared among continents and climates are shown in Fig. 3 and Supplemental Table S4. For surface runoff, the main factor determining the variation in N removal efficiency was the source of pollution \((Q_m = 10.8, df = 2, P = 0.023)\). No buffer zone impact was found for the fields used for grass production, probably due to their initially low levels of pollution; however, double N retention was observed for fields used for cereal production and feedlots, which also had higher levels of pollution (Fig. 3a). In contrast, buffer zones improved groundwater quality to the same extent regardless of the source of pollution (Fig. 3a); moreover, concerning the same source of pollution, the groundwater quality clearly benefited more from buffer zones than the surface runoff.

Regardless of the soil texture, the N retention capacity of the buffer zone was similar for the surface runoff and for the groundwater (Supplemental Table S4), but again, the latter benefited more, as shown for loam soils (Fig. 3b). The effects of the buffer zones on the N retention for surface runoff or groundwater were similar for all the continents and climates (Fig. 3c and 3d). In North America or in humid subtropical climates, the groundwater quality benefited more than the quality of the surface runoff, whereas in Europe or in humid continental and other climates, there was a statistically nonsignificant trend.

### Buffer Zone Characteristics

Although a subgroup analysis did not reveal differences between the buffer zone types (Supplemental Table S4), it is clear that tree buffer zones reduced the amount of NO\(_3\)–N in groundwater, but not in surface runoff, since 95% CI overlaps with zero (Fig. 3e), whereas grass buffer zones and those with a mix of grass and trees reduced the N to the same extent in groundwater as well as in surface runoff. The factors “buffer zone type” and “number of species” were intercorrelated in surface water \((\chi^2 = 16.4, df = 4, P = 0.002)\) and groundwater studies \((\chi^2 = 10.1, df = 4, P = 0.026)\):

\[
\begin{array}{l}
\text{Source of} \\
\text{pollution} \\
\hline
\text{Concentrations} & \text{ Loads} & \text{Concentrations} \\
\hline
\text{Cereals} & 2.43–30.60 & 0.05–22.30 & 0.11–36.30 \\
\text{Grasses} & 0.06–3.74 & 0.24–4.57 & 0.22–12.80 \\
\text{Feedlots} & – & 0.04–80.20 & – \\
\end{array}
\]

† NO\(_3\)–N and total N.
‡ NO\(_3\)–N.

**Fig. 2.** The effect of initial N concentrations on N retention by buffer zones in (a) surface runoff and (b) groundwater. \(Q_m\) model heterogeneity; \(n\), number of observations and studies. The legend shows buffer zone types. The dashed line indicates the control without a buffer zone. For back-transformation of the log response ratio \([\ln(r)]\), see Eq. [5].

**Table 1.** The range of initial N concentrations and loads from the different pollution sources in surface runoff and groundwater without buffer zones.

- grass buffer zones mostly consisted of one or two species, whereas grass–tree and tree buffer zones mostly consisted of multiple species. Therefore, the effect of the species number on N retention was tested for grass buffer zones; however, no impact was found in surface runoff or groundwater (Fig. 3f).

Moreover, according to a meta-regression, the buffer zone efficiency in reducing NO\(_3\)–N and the total N in surface runoff decreased with increasing buffer zone age (Supplemental Fig. S1a). For example, a 22-yr-old buffer zone reduced its efficiency to zero, and older buffer zones seem to be a source of NO\(_3\)–N and total N in the surface runoff. In contrast, buffer zone aging had no effect on the retention of NO\(_3\)–N in groundwater (Supplemental Fig. S1b).

The meta-regressions showed no the effect of the buffer zone width (combined for all types of buffer zone) on the N removal efficiency for both surface runoff \((Q_m = 2.23, P = 0.135, n = 40; \text{Fig. 4a})\) and groundwater \((Q_m = 0.0384, P = 0.845, n = 38; \text{Fig. 4b})\). The observations with broad widths and corresponding N retentions are not shown in Fig. 4. They were 113 m (−88%).
for surface runoff and 130 (−71%), 150 (−69%), and 200 m (−56%) for groundwater, but they were included in the meta-regressions. Similarly, no relationships were found between N removal efficiency and the width of grass buffer zones (surface: \( P = 0.186 \), groundwater: \( P = 0.722 \)), grass–tree buffer zones (surface: \( P = 0.899 \), groundwater: \( P = 0.619 \)), or tree buffer zones (surface: \( P = 0.778 \), groundwater: \( P = 0.869 \)). Narrow buffer zones (<10 m) reduced the NO\(_3\)-N and total N equally as well as wider buffer zones did; however, it should be noted that there was enormous variation in the results (Fig. 4a and 4b).

The buffer zone slope in a range that was reported in the articles did not modify buffer zone efficiency for surface runoff (\( Q_{M} = 0.105, P = 0.747, n = 40; \) Fig. 4c) or groundwater (\( Q_{M} = 1.59, P = 0.206, n = 28; \) Fig. 4d). However, it should be noticed that steep (>10%) buffer zone slope was reported in only one study on groundwater in the entire database (Fig. 4d).

**Discussion**

This meta-analysis quantitatively summarized the results of 46 studies published over a period of >35 yr on the N retention capacity of buffer zones regarding surface runoff and groundwater. Although many relevant studies were omitted due to incomplete and poor reporting of the results, the number of studies included in the database was sufficient to perform a robust, weighted meta-analysis to calculate a summarized effect size across the studies, as well as the means for different categories of explanatory variables, and to determine the CIs around the means. To compare the meta-analysis results with the literature,
we have listed the existing statements and conclusions most frequently mentioned in several narrative reviews regarding the performance of different types of buffer zones and indicated agreements with the meta-analysis (Table 2). Some of the statements have been supported by the meta-analysis, while others have been rejected due to lack of evidence.

Surface Runoff vs. Groundwater

According to Mayer et al. (2007), subsurface N removal is more efficient than removal through surface flows. This is in accordance with this meta-analysis, which showed about a twice larger N retention by buffer zone concerning groundwater compared with that of surface runoff (Fig. 1). The removal of N in the subsurface flow is explained only partly by vegetation uptake, but the main mechanism for removal is usually denitrification, which occurs throughout the year and depends on an abundance of NO\textsubscript{3}\textsuperscript{−} and C sources, together with anaerobic conditions (Vought et al., 1994; Fennessy and Cronk, 1997). The denitrification rates measured in different studies varied between 9 and 70 kg N ha\textsuperscript{−1} yr\textsuperscript{−1} in riparian buffer zones (Pärn et al., 2012). When the NO\textsubscript{3}\textsuperscript{−} loading in riparian buffer zones is high, N\textsubscript{2}O is an important end product of denitrification. In these cases, N transformation by the buffer zones results in an unfavorable shift from water pollution to an increase in greenhouse gas emissions (Hefting et al., 2003).

Buffer zones are generally very effective in the removal of sediment, as well as sediment-associated nutrients and compounds from surface runoff (Dillaha et al., 1989; Gilliam et al., 1996). Since surface flows bypass zones of denitrification, the main mechanism of N removal is N uptake by plants, which is not active, however, during dormant periods or in older buffer zones (Fennessy and Cronk, 1997; Mander et al., 1997). Nitrogen uptake in riparian vegetation varies considerably depending on the aerobic conditions in the soil, the plant community, the natural disturbances, and the harvesting rates. Riparian meadow grasses and herbs normally accumulate 20 to 70 kg N ha\textsuperscript{−1} yr\textsuperscript{−1}, whereas riparian forests take up as much as 30 to 170 kg N ha\textsuperscript{−1} yr\textsuperscript{−1} (reviewed by Pärn et al., 2012).

To reduce the risk of N leaching from dormant buffer zone plants, Räty et al. (2010) recommend harvesting the aboveground biomass in the buffer zone. In addition, incoming soluble NO\textsubscript{3}\textsuperscript{−} ions in the runoff can be retained by physical retention (e.g., infiltration), microbial immobilization, chemical transformation (e.g., denitrification in wet weather conditions), and dilution (e.g., groundwater mixing) (Osborne and Kovacic, 1993; Polyakov et al., 2005; Dosskey et al., 2010; Pärn et al., 2012).

Initial Nitrogen Concentrations

Previously, it was reported that the capacity of buffer zones to retain nitrate N and total N in surface runoff and groundwater...
was related to the initial N concentration or load levels (Petersen et al., 1992; Haycock and Pinay, 1993; Lowrance et al., 1997; Mander et al., 1999; Mayer et al., 2007). Most reduction of total N in surface runoff occurred when the initial N concentration level was high (Petersen et al., 1992). Similarly, in groundwater, nitrate N retention by grass and tree buffer zones linearly correlated with load rates during the winter (Haycock and Pinay, 1993). This meta-analysis proved that the N retention in surface runoff and groundwater was linearly related to initial N concentrations entering the buffer zone as indicated by the meta-regressions (Fig. 2). Moreover, a subgroup analysis showed that, for grass production fields, discharging very low, similar to rainwater, N concentrations (<3.74 mg L⁻¹), into surface runoff, the N retention by buffer zones was only 15% and did not statistically differ from the control (Fig. 3a). In contrast, for cereal production fields, discharging high N concentrations (up to 30.6 mg L⁻¹), the N retention by buffer zones was 50%.

### Buffer Zone Characteristics

#### Buffer Zone Type

Woody plants are considered to be more efficient in the uptake of nutrients from deeper subsurface waters than grass vegetation, since they are generally much larger, taller, longer lived, and their stems grow more widely spaced than herbaceous plants (Dosskey et al., 2010). Therefore, many review papers stress that tree buffer zones have higher N removal efficacy than grass buffer zones (Fennessy and Cronk, 1997; Lyons et al., 2000; Hawes and Smith, 2005; Zhang et al., 2010), and buffers combining grass and trees tend to be more effective in NO₃⁻N retention than buffer zones consisting only of herbaceous vegetation (Christen and Dalgaard, 2013). However, this meta-analysis showed that the groundwater N retention capacity of buffer zones with trees and combined grass–tree buffer zones seemed to be only 10 to 15% larger than that of purely grass buffer zones, but the difference was not statistically significant (Fig. 3c). Thus, in terms of groundwater, the meta-analysis is in accordance with the reviews by Wenger (1999) and Mayer et al. (2007) claiming that there is no effect resulting from the buffer vegetation type on the N retention capacity. In contrast with all the reviews, the meta-analysis demonstrated that tree buffer zones were not effective in reducing N in surface runoff (Fig. 3c).

#### Buffer Zone Age

Nutrient removal efficiencies in buffer zones may also be affected by the age of the vegetation. Young forest stands, bushes, and wet grasslands achieved the most intensive nutrient removal due to intensive nutrient uptake by plants as they were in an active growth phase (Mander et al., 1997). This meta-analysis supported the statement concerning N retention in surface runoff, but not in groundwater, where the N removal mechanism relies mostly on denitrification.

#### Buffer Zone Width and Slope

The riparian buffer width is thought to be positively related to N removal efficacy, and the steeper the land within the buffer, the wider it needs to be to have time to slow the flow of water and absorb the pollutants within it (Fischer and Fischenich, 2000). As the buffer slope increases, the runoff speed increases, reducing the residence time of the runoff in the buffer and reducing the removal efficacy. There is evidence, however, that a slight slope facilitates runoff and encourages water flow across the buffer, increasing the removal efficacy (Wenger, 1999). Many researchers suggest that especially steep slopes (>10%) serve little value as a buffer, and they recommend excluding steeply sloped areas when calculating buffer widths (Hawes and Smith, 2005; Petersen and Vondracek, 2006). This meta-analysis did not support these statements, and the range of surface N removal efficiency for steeper slopes was as large as for lower slopes (Fig. 4c); however, in our database on groundwater, there was a lack of measurements for slopes >10% (Fig. 4d).

Previous reviews reported that buffer width alone explains 20 to 44% of the variation in surface N removal efficacy (Mayer et al., 2007; Zhang et al., 2010). In contrast, subsurface N removal efficiency was not related to the buffer zone width (Wenger, 1999; Mayer et al., 2007; Sweeney and Newbold, 2014). This meta-analysis found no relationships between buffer zone width and N removal efficiency regarding surface runoff or groundwater. Sweeney and Newbold (2014) noticed that N removal per unit width of buffer varied inversely with subsurface water flux, and the failure of simple correlations to detect a significant influence of the buffer width on removal efficiency can be explained by the large obscuring influence of water flux.

---

**Table 2. Statements and conclusions reported in reviews regarding the performance of buffer zones and results of the meta-analysis.**

<table>
<thead>
<tr>
<th>Water quality†</th>
<th>Statements and conclusions</th>
<th>Reviews</th>
<th>Meta-analysis results</th>
</tr>
</thead>
<tbody>
<tr>
<td>S, G</td>
<td>Groundwater &gt; surface</td>
<td>Mayer et al. (2007)</td>
<td>Support Fig. 1</td>
</tr>
<tr>
<td>S, G</td>
<td>More loads or concentrations than larger retention</td>
<td>Petersen et al. (1992), Haycock and Pinay (1993), Mander et al. (1997)</td>
<td>Support Fig. 2</td>
</tr>
<tr>
<td>S, G</td>
<td>Tree buffer zone = grass buffer zone</td>
<td>Wenger (1999), Mayer et al. (2007)</td>
<td>Support for groundwater in Fig. 3e</td>
</tr>
<tr>
<td>S, G</td>
<td>Tree buffer zone &gt; grass buffer zone</td>
<td>Fennessy and Cronk (1997), Lyons et al. (2000), Hawes and Smith (2005), Zhang et al. (2010)</td>
<td>Does not support Fig. 3e</td>
</tr>
<tr>
<td>S, G</td>
<td>Grass–tree buffer zone &gt; grass buffer zone</td>
<td>Christen and Dalgaard (2013)</td>
<td>Does not support Fig. 3e</td>
</tr>
<tr>
<td>S</td>
<td>Buffer zone efficiency is related to width</td>
<td>Mayer et al. (2007), Zhang et al. (2010)</td>
<td>Does not support Fig. 4a</td>
</tr>
<tr>
<td>G</td>
<td>Buffer zone efficiency is not related to width</td>
<td>Wenger (1999), Mayer et al. (2007), Sweeney and Newbold (2014)</td>
<td>Support Fig. 4b</td>
</tr>
<tr>
<td>S</td>
<td>Low slope &gt; steep slope (&gt;10%)</td>
<td>Hawes and Smith (2005)</td>
<td>Does not support Fig. 4c</td>
</tr>
<tr>
<td>S, G</td>
<td>Soil texture is important</td>
<td>Barling and Moore (1994), Vidon and Hill (2004)</td>
<td>Does not support Fig. 3b</td>
</tr>
<tr>
<td>S, G</td>
<td>Younger buffer zone &gt; older buffer zone</td>
<td>Mander et al. (1997)</td>
<td>Support for surface runoff in Supplemental Fig. S1a</td>
</tr>
</tbody>
</table>

† S, surface runoff; G, groundwater.
Soil Texture

The role of soil texture in the determination of the retention capacity of buffer zones seems not to be straightforward, and it is difficult to evaluate soil texture effects in a study. The type of soil affects how quickly water can be absorbed. Soils that are high in clay are less permeable and may have greater runoff. On the other hand, soils that are largely made up of sand may drain water so rapidly into the groundwater that roots are not able to effectively trap pollutants (Hawes and Smith, 2005). This meta-analysis showed that the N retention capacity for surface runoff and groundwater was the same regardless of soil texture (Fig. 3b). It should be noted that the majority of studies have been conducted on loam, whereas observations in clay soils originated from two studies conducted in Finland (Uusi-Kämppä and Palojärvi, 2006; Uusi-Kämppä and Jauhiainen, 2010). In a recent study of nine sites in Italy, Balestrini et al. (2016) concluded that, in addition to soil texture, there were several predictors of NO\textsubscript{3}–N removal capacity linked to the water residence time, such as the hydraulic conductivity and the slope of the riparian profile, together with the water table depth and soil organic C.

Conclusions

The meta-analysis clearly demonstrated that buffer zones more effectively reduced N in groundwater than in surface runoff, despite the large variation of results across the studies. The main source of variation in the buffer zone capacity to improve quality in the surface runoff, as well as in the groundwater, was the initial N concentrations discharging from the source of pollution, and the capacity increased with increasing N pollution. To improve surface runoff quality, buffer zones need to be established between water bodies and fields with cereal production or feedlots due to their higher environmental risk compared with grass production fields, for which buffer zones seem to be beneficial. Unlike groundwater quality, which responded equally well regardless of the source of pollution, buffer zone type, or age, surface water quality is more sensitive, and it might not be satisfactorily improved by tree buffer zones or aged buffer zones, or when the source of pollution is from grass production fields.

According to the results of this meta-analysis, we recommend buffer zones to be primarily used to reduce the NO\textsubscript{3}–N concentrations in groundwater and improve its quality. Buffer zones should also be used to protect surface water from N emissions, especially from agricultural fields with a high risk of N runoff, such as fields used for cereal production. The meta-analysis provides a solid base for developing more accurate environmental impact assessment methods (i.e., life cycle assessments), and further studies are needed for this kind of approach.

It should be noticed that a number of published experiments failed to report means, sample sizes, and variances for controls and buffer zone treatments, which make it impossible to include those studies in a meta-analysis that uses the weighted parametric statistical tests designed for meta-analysis. For example, there were significantly fewer studies outside of North America included in the meta-analysis, and also a lack of data for buffer zone slopes >10% in groundwater studies. The obvious solution was already proposed 20 yr ago by Gurvitch and Hedges (1999) to upgrade publication standards, alerting authors, reviewers, and editors so that papers are not published without the basic information necessary for readers to properly evaluate the results.

Supplemental Material

The supplemental material consists of a description of the database, a list of explanatory variables and their effects (Supplemental Tables S1–S4), Supplemental Fig. S1, and references included in the database.

Acknowledgments

This study was funded by the EU LIFE Programme as part of the project FRESHABIT—Towards integrated management of freshwater Nature 2000 sites and habitats (LIFE14 IPE/FI/023). We thank the two anonymous reviewers for their valuable comments.

References


