

USE OF NONLINEAR REGRESSION FOR ESTIMATING EMISSION OF NITROGEN TO ESTONIAN RIVERS

Anatoli Vassiljev, Ivar Annus, Katrin Kaur, Nils Kändler

Tallinn University of Technology, Ehitajate tee 5, Tallinn 19086, Estonia

Abstract

A statistical model, MESAW, was modified to estimate non-linear diffuse source emission coefficients. This model uses regression for estimation of source emission coefficients for different land use or soil categories and retention coefficients for pollutants in a river basin or lakes. In this study, the emission coefficients were calculated based on a non-linear dependency between the emission and the land cover percentage in sub-basins. Results show that exponent of non-linear emission coefficients for different land covers vary from 2 to 6 times between single years. Maximal unit area loads gained in this study coincide well with results obtained in previous similar investigations indicating that the new approach on defining the emission coefficients is reliable.

Key words: nutrients, emission coefficient, diffuse sources, MESAW, drained peat soils, water quality

1. INTRODUCTION

Reduction of riverine nutrient loads is one of the major environmental management goals in the Baltic Sea region and thus also in Estonia. However, recent data analyses of Estonian rivers indicate that nitrogen concentrations have increased in some rivers (Iital et al. 2010) despite the decrease of fertilizer usage (Vassiljev & Blinova 2012). An increase in nitrogen concentrations has even been detected in watersheds with very low human activity (Iital et al. 2010). Hoffmann et al. (2000) assumed that increases of pollution by nutrients may also be caused by wide-scale melioration. Some authors (Heikkinen 1994; Kløve 2001; Kløve et al. 2010) have reported that drainage of peat soils leads to decomposition of peat and increases fluxes of nutrients to watercourses. Drainage of peatlands results in peat oxidation and significantly changes their physical and chemical properties (Litaor et al. 2008; Verhoeven & Setter 2010). This can result in high nitrate-nitrogen concentrations in the pore water of drained peatlands that is caused by the aeration of peat and subsequent mineralization and nitrification of organic N (Tiemeyer et al. 2007). Vassiljev & Blinova (2012) observed an increase of nitrogen concentrations in some Estonian river basins with high percentage of drained peat soils. Mineralization of nitrogen from peat soils is regarded as another possible source of nutrients in Europe (Eurostat 2011). In Vassiljev et al. (2016) a statistical model MESAW was used to show that unit-area loads from drained peat areas can be 2.3 times higher than from arable lands. Emission coefficients in that study were assumed to be constants. The main objective of this paper was to introduce nonlinear dependencies of emission coefficients on the percentage of total area of sub-basins.

2. CASE STUDIES AND METHODOLOGY

In the current study, a statistical approach was used to estimate the emission coefficients of nitrogen from various diffuse source categories. Emission coefficients were estimated for the whole Estonian territory (Fig. 1).



Figure 1. Map of the study area

A statistical model MESAW proposed by Grimvall & Stålnacke (1996) was used for the source apportionment and retention estimates of nitrogen. The model is suitable for source apportionment, especially for areas with a dense network of water quality monitoring sites, as shown by Lidèn et al. (1998), Vassiljev et al. (2008) and Stålnacke et al. (2015). The model approach uses non-linear regression for simultaneous estimation of source strength (i.e. export/emission coefficients to surface waters) for the different land use and/or soil categories and retention coefficients for pollutants in a river basin or lakes. There are four major steps in the procedure: (1) estimation of mean annual riverine N loads for each year at each water quality monitoring site, (2) subdivision of the entire drainage basin into sub-basins according to the upstream area of the water quality monitoring site, (3) derivation of statistics on land use, lake area, point source emissions and other relevant data for each sub-basin, (4) using a general non-linear regression expression with loads at each sub-basin as the dependent/response variable and sub-basin characteristics as covariates/explanatory variables (Lidèn et al. 1998; Stålnacke et al. 2015). In the MESAW model, the load at the outlet of an arbitrary sub-basin can be estimated from the following general expression (Eq. (1)).

$$L_i = \sum_{j=1}^n (1 - R_{j,i}) L_j + (1 - R) S_i + (1 - R) P_i + (1 - R) D_i + \varepsilon_i \quad (1)$$

where

- L_i = load at the outlet of sub-basin i ;
- L_j = load at the outlet of nearest upstream sub-basin j ;
- $R_{j,i}$ = retention on the way from the outlet of sub-basin j till the outlet of sub-basin i ;
- n = number of sub-basins located to the nearest upstream;
- S_i = total losses from soil to water in sub-basin i ;
- P_i = point source discharges to waters in sub-basin i ;
- D_i = atmospheric deposition on surface waters in sub-basin i ;
- R = retention in sub-basin i ;
- ε_i = statistical error term.

The load at each sub-basin can be divided into contributions from the sources located in sub-basins further upstream (the first term in Eq. (1)) and contributions from the sources located within the sub-basin under consideration (the S_i , P_i and D_i terms). For some sub-basins n can be equal to zero (e.g. the uppermost sub-basin or separate basin without any upstream sub-basin). In this case, Eq. (1) form will be without the first term. The parameterization of the model is flexible and can be study-area specific. The model is fitted by minimizing the sum of squares for the difference in the observed and estimated loads. In current study, P_i and D_i were assumed to be known and S_i was assumed to be a simple function of land use according to $S_i = (\beta_1 a_{1i} + \beta_2 a_{2i} + \beta_3 a_{3i})$. Herein, a_{1i} , a_{2i} and a_{3i} denote the area of arable land (excluded arable land on drained peat soil, ~5% of total arable land), natural areas (other land) that included forest, pastures, natural grass lands, bogs (excluded areas on drained peat soils) and drained peat soils (including all land types located on drained peat soils) in the sub-basins i . β_1 , β_2 and β_3 are unknown export coefficients (i.e. emission coefficients, unit-area loads) for the three land use categories calculated by MESAW.

In the current study, the emission coefficients were assumed to have a non-linear dependency on the land use coverage percentage of the total sub-basin area. For example, the dependence of emission coefficients β_1 was calculated:

$$\beta_{1,i} = \beta_1 - \beta_1 * \left(1 - \frac{\alpha_{1,i}}{\text{drainage area}_i} \right)^\gamma \quad (2)$$

where

- $\beta_{1,i}$ = export coefficient for sub-basin i ;
- β_1 = export coefficient if land use a_1 covers hundred percent of sub-basin i ;
- γ = exponent which can be estimated as a parameter or specified by user.

Similar equation was used for calculating β_2 and β_3 . MESAW user interface and source code were modified in order to use the nonlinear regression for estimating emission of nitrogen in rivers. In the modified program it is possible to use old version of MESAW (emission coefficient does not depend on land cover percentage), user specified exponents and exponents estimated as parameter.

Nutrients are normally retained temporally or permanently in watercourses. In MESAW model, retention is expressed as a summary expression for all hydrological and biogeochemical processes that may retain nutrients. It can be parameterized by any empirical function. In this study, the retention is divided in two – retention in lakes and river retention (i.e. instream retention). It was assumed that retention in lakes is a function of the lake area divided by the drainage area, and riverine retention a function of the drainage area.

Both types of retention can be expressed according to the following formula:

$$R_i = 1 - \frac{1}{1 + \lambda_1 \times \text{drainage area}_i} \times \frac{1}{1 + \lambda_2 \times \frac{\text{lake area}_i}{\text{drainage area}_i}} \quad (3)$$

where λ_1 and λ_2 denote a non-negative parameter and R_i denotes the retention in the i th basin. The first part of the function reflects the in-stream retention whereas the second part reflects the retention in lakes and reservoirs.

Retention from an arbitrary sub-basin m to the river mouth (R_{mouth}) can be derived from:

$$R_{m,\text{mouth}} = 1 - \prod_{j=1}^k (1 - R_j) \quad (4)$$

where

- $R_{m,\text{mouth}}$ = retention from the outlet of the sub-watershed m on the way to the mouth of the whole river;

k = number of sub-basins downstream sub-basin m ;
 R_j = values of retention within the different sub-basin downstream the sub-basin m .

In this study, the data (total nitrogen concentrations and water runoff) of 7 years during the period 2005-2013 was used. Estonian Environmental Agency (EEA) collected hydro chemical parameters from 50 sites in the annual state monitoring programme. In addition, EEA measured water runoff at 46 locations that in some measurement points did not coincide with the hydro chemical monitoring sites. Soil types in the sub-basins were estimated using digital soil map obtained from Estonian Land Board. Digital CORINE land cover map was used to derive land use statistics for each sub-basin. In the model, the emission coefficients were calculated for arable lands excluding peat soils, areas with drained peat soils and other lands excluding peat soils (including forest, pastures, natural grasslands and bogs). Maximal, minimal and average percent of different land types are presented in Table 1.

	Arable	Drained peat	Other	Lakes
max	52.2	44.3	87.9	8.4
min	7.2	3.9	34.2	0.0
average	29.2	14.9	55.3	0.7

Table 1. Land cover in sub-basins (%)

In Figure 2 the percentages of land cover classes in the sub-basins under investigation is shown. It can be seen that arable lands cover 25-30% in 11 sub-basins, 10-15% of drained peat soils are in 22 sub-basins and other lands cover 45-50% in 14 sub-basins. In all sub-basins other lands cover at least 30% of the total area. There are some sub-basins where natural areas are dominant (~90%). Arable lands cover maximally 52% of sub-basins total area while drained peat soils are covering typically 5-25% of total area of a sub-basin. There are no sub-basins where the land use is only arable or soil type is only drained peat.

3. RESULTS AND DISCUSSION

MESAW modified was used to calculate the diffuse source emission coefficients of total nitrogen. The coefficients were estimated for seven single years (2005, 2006, 2007, 2008, 2011, 2012 and 2013, Table 2). The years were selected with different average runoff to cover the whole amplitude of runoff from minimal to maximal values.

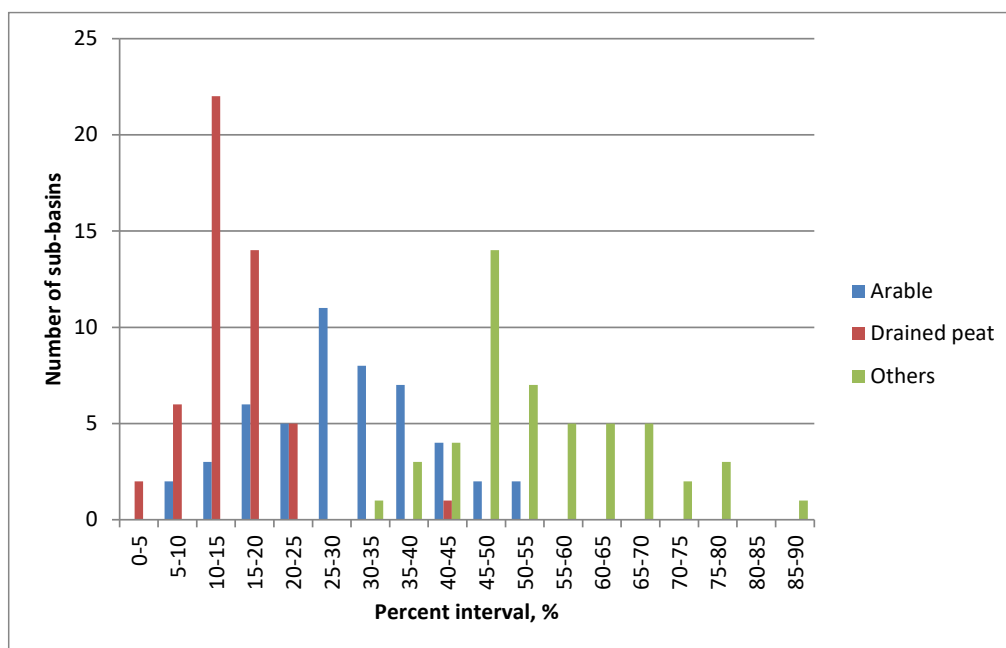


Figure 2. The division of different land use in sub-basins under investigation

Year	Arable, kg/ha	Exponent	Natural areas, kg/ha	Exponent	Drained peat areas, kg/ha	Exponent	Average water runoff, mm
2005	10.9	6.2	1.5	10.2	26.0	10.8	285
2006	9.6	6.1	0.8	9.3	11.1	18.0	160
2007	15.0	9.8	1.2	4.4	21.4	23.5	245
2008	17.6	6.8	2.8	23.1	35.6	30.1	409
2011	19.6	11.5	2.2	15.1	36.4	8.1	338
2012	19.2	7.0	2.7	23.2	29.3	7.2	392
2013	10.5	11.5	1.2	4.0	24.5	18.2	253

Table 2. Results of estimated total nitrogen emission coefficients and exponents.

The exponents in Eq. (2) were calculated for all three land cover class. As an example in Figure 3 the dependency of emission coefficient and the percentage of arable land covering the sub-basins is shown. Results for three years, 2012, 2007, and 2006, are presented as they correspond to years with maximal, minimal and average nitrogen emission. It can be seen that emission of nitrogen from one hectare of arable land increases with the increase in the percentage of arable land covering the sub-basins and reaches its maximum value at 50%. It means that emission for one hectare is the same when arable lands cover more than 50% of sub-basins total area. On the other hand, set of data used in this study does not contain sub-basins where arable lands cover more than 50% of the area. So for validation more data (additional water samples) is needed to define the emission coefficients for sub-basins where more than 50% is covered with arable lands and/or 45% with drained peat soils. In Figure 4 the same dependencies of relative emissions are shown (divided by maximum emission). It can be seen that the relative emissions are similar, indicating that the function shape is not dependent on annual changes in total nitrogen emission. Dependencies for other land uses are similar but variances in different years are higher. For example, emission from drained peat achieves maximal value at 10% in the year 2008 and at 50% in the year 2012. Use of exponent increases the standard

errors of all emission coefficients because in this case the number of parameters is higher. The emission values increase compared to emission coefficients gained using linear areal dependency (Vassiljev et al. 2016). This can be explained by the fact that in this case the emission from one hectare is not constant.

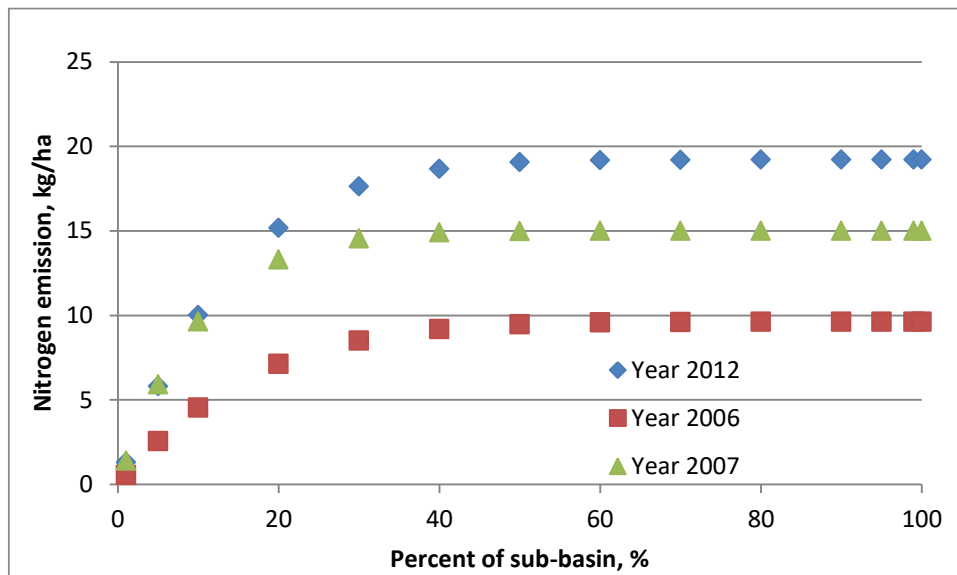


Figure 3. Nitrogen emission from hectare of arable land vs. percentage of arable land in sub-basin

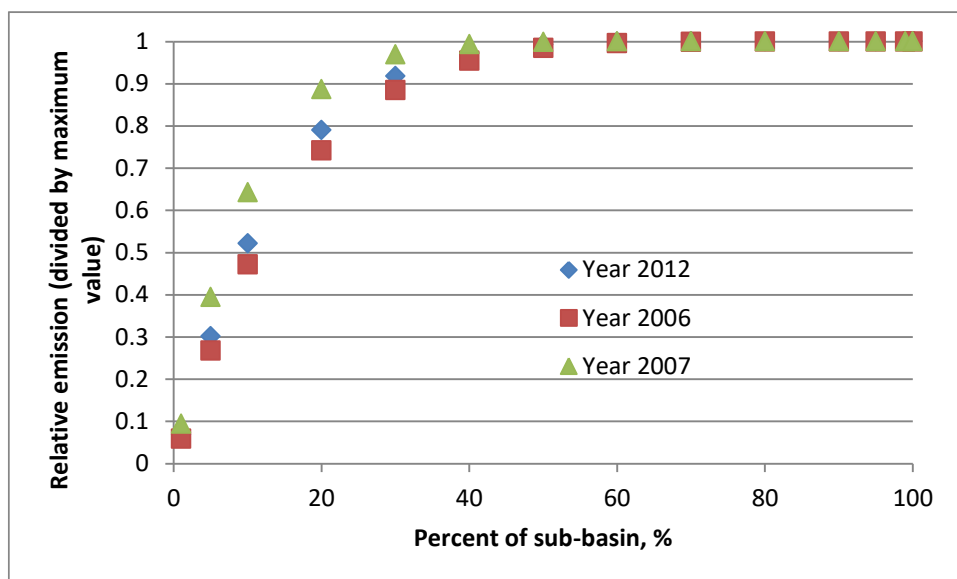


Figure 4. Relative emission vs. percentage of arable land in sub-basin

The emission coefficients gained in this study show that during the study period in between single years the emission from arable land varied up to 2, natural areas 3 and drained peatlands 3.5 times. Emission coefficient from arable lands vary from 9.6 kg/ha to 19.6 kg/ha (Table 2) and are comparable with results gained by Vassiljev et al. (2008). They studied the emission coefficients for the whole of Estonia during 1995-2005 and found that the coefficient of nitrogen from arable lands was 12.3 kg/ha. Stålnacke et al. (2015) have obtained similar results in the Baltic Sea drainage basin. The estimated emission from cultivated areas in the Baltic Sea drainage basin were found to be 14.3 kg/ha. Emissions

form drained peat areas were not investigated in aforementioned studies. A field study in Northern Germany (Scholz & Trepel 2004) revealed that the nitrate concentrations in drained peatlands with unnatural state can be 5 – 60 times higher compared to natural peatlands. The emission coefficients from drained peatlands gained in this study varied from 11.1 kg/ha to 36.4 kg/ha and were up to 2.4 times higher than from arable lands (Table 2). High emission coefficients of drained peat soils obtained in this investigation can be explained by a high percentage of unnatural state of peatlands in Estonia since only 5.5 % of peatlands are in near-natural state (Paal & Leibak 2011).

Annual changes in emission coefficients during the study period were up to 2 times from arable lands, 3.5 times from natural areas and 3.3 times from drained peatlands. Nutrient concentrations are affected by annual changes in the temperature. In Tiemeyer et al. (2007) it was found that in some years NO₃–N concentrations were low (or zero) at the beginning of the discharge season, while low flow rates at the end of the discharge season (April) coincided with comparatively high NO₃–N concentrations. This was explained with different climate conditions during the study period as the net release of nitrogen from peat soils increases with increasing temperature (Koerselman et al. 1993). Additional investigations are needed to find the relations between the temperature and the changes in the emission coefficients. This will allow to take into account the effects of seasonal changes to emission coefficients. More in situ measurements are needed to support the hypothesis.

The comparison of the emission coefficients gained in this study and during previous investigations reveal that the presented modelling approach is reliable. In the current study, the emission coefficients were assumed to be dependent on the land use coverage percentage of the total sub-basin area. A non-linear dependency between the emission coefficient and land cover percentage was introduced. The modelling approach was linked with the MESAW user interface that allows user to self-define the exponents or estimate them as a parameter. The MESAW model user interface is MS Excel based and free of charge.

4. CONCLUSIONS

- MESAW model user interface was modified to add the non-linear emission coefficient calculation tool.
- The MESAW model modified was used to estimate diffuse emission coefficients and their non-linear dependencies on the percentage of the land use cover in a sub-basin.
- Unit-area losses from arable lands varied up to 2 times during the study period.
- Unit-area losses from drained peat soils were estimated to vary between 11-36 kg/ha while the exponent of the emission coefficient varied from 7 to 30.
- Unit-area losses from drained peat soils can be up to 2.4 times higher than from arable land and coincide with results obtained in Vassiljev et al. (2016).
- More in situ measurements are needed to obtain more stable data.

ACKNOWLEDGEMENTS

This study was supported by the Institutional Research Funding IUT19-17 at Tallinn University of Technology, by EEA Financial Mechanism 2009-2014 programme "Integrated marine and inland water management" (project title "Development of data modelling system and the decision support tool for the integrated marine and inland water management"), and by the Central Baltic Programme project Waterchain. The authors would like to thank Estonian Environmental Research Centre for valuable data.

REFERENCES

- EC Eurostat 2011, *Data requirements, availability and gaps in agri-environment indicators (AEIs) in Europe, 2011 edition*. Publications Office of the European Union, Luxembourg.
- Grimvall, A & Stålnacke, P 1996, 'Statistical methods for source apportionment of riverine loads of pollutants', *Environmetrics*, vol. 7, no. 2, pp. 201–213.
- Heikkinen, K 1994, 'Organic matter, iron and nutrient transport and nature of dissolved organic matter in the drainage of a boreal humic river in northern Finland', *Science of the Total Environment*, vol. 152, no. 1, pp. 81–89.
- Hoffmann, M, Johnsson, H, Gustavson, A & Grimvall, A 2000, 'Leaching of nitrogen in Swedish agriculture - a historical perspective', *Agriculture, Ecosystems & Environment*, vol. 80, no. 3, pp. 277–290.
- Iital, A, Pachel, K, Loigu, E, Pihlak, M & Leisk, U 2010, 'Recent trends in nutrient concentrations in Estonian rivers as a response to large-scale changes in land-use intensity and lifestyles', *Journal of Environmental Monitoring*, vol. 12, no. 1, pp. 178–188.
- Kløve, B 2001, 'Characteristics of nitrogen and phosphorus loads in peat mining wastewater', *Water Research*, vol. 35, no. 10, pp. 2353–2362.
- Kløve, B, Sveistrup, TE & Hauge, A 2010, 'Leaching of nutrients and emission of greenhouse gases from peatland cultivation at Bodin, Northern Norway', *Geoderma*, vol. 154, no. 3–4, pp. 219–232.
- Koerselman, W, Van Kerkhoven, M & Verhoeven, JTA 1993, 'Release of inorganic N, P and K in peat soils; effect of temperature, water chemistry and water level', *Biogeochemistry*, vol. 20, no. 2, pp. 63–81.
- Lidén, R, Vasilyev, A, Loigu, E, Stålnacke, P & Wittgren, HB 1999, 'Nitrogen source apportionment using a physically and statistically based model', *Ecological Modelling*, vol. 114, no. 2–3, pp. 235–250.
- Litaor, MI, Eshel, G, Sade, R, Rimmer, A & Shenker, M 2008, 'Hydrogeological characterization of an altered wetland', *Journal of Hydrology*, vol. 349, no. 3–4, pp. 333–349.
- Paal, J & Leibak, E 2011, *Estonian mires: inventory of habitats*, Publication of the Project „Estonian Mires Inventory completion formaintaining biodiversity”, Regio, Tartu, Estonia.
- Scholz, M & Trepel, M 2004, 'Water quality characteristics of vegetated groundwater-fed ditches in a riparian peatland', *Science of the Total Environment*, vol. 332, no. 1–3, pp. 109–122.
- Stålnacke, P, Pengerud, A, Vassiljev, A, Smedberg, E, Mörrth, CM, Hägg, HE, Humborg, C & Andersen, HE 2015, 'Nitrogen surface water retention in the Baltic Sea drainage basin', *Hydrology and Earth System Sciences*, vol. 19, no. 2, pp. 981–996.
- Tiemeyer, B, Frings, J, Kahle, P, Kohne, S & Lennartz, B 2007, 'A comprehensive study of nutrient losses, soil properties and groundwater concentrations in a degraded peatland used as an intensive meadow -Implications for re-wetting', *Journal of Hydrology*, vol. 345, no. 1–2, pp. 80–101.
- Vassiljev, A & Blinova, I, 2012, 'The influence of drained peat soils on diffuse nitrogen pollution of surface water', *Hydrology Research*, vol. 43 no. 4, pp. 352–358.
- Vassiljev, A, Blinova, I & Ennet, P 2008, 'Source apportionment of nutrients in Estonian rivers', *Desalination* vol. 226, no. 1–3, pp. 222–230.
- Vassiljev, A, Margus, G, Annus, I & Stålnacke, P 2016, 'Investigation of possible nutrient sources in Estonian rivers', *Procedia Engineering*, vol. 162, pp. 188–195.
- Verhoeven J. T. A. & Setter T. L. 2010 'Agricultural use of wetlands: opportunities and limitations', *Annals of Botany*, vol. 105, no. 1, pp. 155–163.